Do Films Make You Think? –

Inference Processes in Expository Film Comprehension

Dissertation

der Fakultät für Informations- und Kognitionswissenschaften

der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dipl.-Psych. Maike Tibus

aus Braunschweig

Tübingen

2008

Tag der mündlichen Qualifikation:	05. November 2008
Dekan:	Prof. Dr. Oliver Kohlbacher
1. Berichterstatter:	Prof. Dr. Stephan Schwan
2. Berichterstatter:	Prof. Dr. Dr. Friedrich Hesse

In erster Linie kommt es darauf an, verstanden zu werden.

Theodor Fontane

Danke

An dieser Stelle möchte ich den Personen danken, die mich auf dem Weg zu dieser Arbeit unterstützt haben. Mein besonderer Dank gilt den folgenden Personen:

Prof. Dr. Stephan Schwan - für die Betreuung dieser Arbeit, insbesondere für Inspiration, Ruhe, Gelassenheit und Vertrauen

Prof. Dr. Dr. Friedrich Hesse - für die Übernahme des Zweitgutachtens und sein Engagement für ein wissenschaftlich anregendes Umfeld

Prof. Dr. Peter Gerjets – für Dein intensives Mitdenken und -diskutieren

Prof. Dr. Danielle McNamara - als transatlantische Diskussionspartnerin mit der Textverstehensbrille

Prof. Dr. Rolf Plötzner - als Zweitbetreuer im VGK

den VGK – Mädels, insbesondere Jessica Dehler und Kristin Knipfer – für eine tolle Tübinger Promotionszeit

Sebastian Groteloh - als rettender Anker in sämtlichen technischen Belangen

Eva Brenner, Anke Heier, Martin Merkt, Katharina Nöbels und Sonja Weigand - für die tatkräftige Unterstützung bei der Durchführung der Experimente

Karin Petropoulos - für Notfallkekse und ein offenes Ohr

den Beteiligten des Virtuellen Graduiertenkollegs der Deutschen Forschungsgemeinschaft "Wissenserwerb und Wissensaustausch mit neuen Medien" – für den Blick über den eigenen (Dissertations-)Tellerrand

sowie den Mitarbeitern des Instituts für Wissensmedien in Tübingen.

Zu guter letzt danke ich meinen Herzensmenschen – für Eure emotionale Wärme, Eure aufmunternden Worte, Euer Zuhören und für Eure vielen anderen Gesten der Zuneigung, die mir den nötigen Rückhalt und Stärke gegeben haben, um diese Arbeit fertig zu stellen.

Contents

1 INTRODUCTION ... 1

2 DOES FILM COMPREHENSION EQUAL TEXT COMPREHENSION? ... 4

2.1 Differences between texts and films ... 4

2.2 Similarities between text and film comprehension processes ... 7

3 PRINCIPLES OF TEXT COMPREHENSION ... 10

3.1 The Construction-Integration Model ... 11

3.2 Inferences ... 16

3.3 Coherence and comprehension ... 20

4 BRIDGING INFERENCES IN TEXT COMPREHENSION ... 25

4.1 Causal bridging inferences in narrative texts ... 26

4.2 Causal bridging inferences in expository texts ... 28

4.3 Reasons for the imbalance in inference generation as a function of text genre ... 31

5 INFERENCES IN FILM COMPREHENSION ... 33

5.1 Empirical studies ... 33

5.2 Methodological implications ... 37

6 INFERENCE DETECTION METHODS ... 39

6.1 Online methods ... 39

6.2 Offline methods ... 51

7 SUMMARY AND OVERVIEW OF EXPERIMENTS ... 56

8 EXPERIMENT 1: GLOBAL BRIDGING INFERENCES: ARE THEY GENERATED ONLINE IN EXPOSITORY FILM COMPREHENSION AND HOW DO THEY RELATE TO OFFLINE COMPREHENSION PERFORMANCE? ... 58

8.1 Hypotheses ... 59

8.2 Method ... 60

8.3 Results ... 66

8.4 Discussion ... 73

9 EXPERIMENT 2: LOCAL CAUSAL BRIDGING INFERENCES: ARE THEY GENERATED ONLINE IN EXPOSITORY FILM COMPREHENSION? 79

9.1 Hypotheses 80

9.2 Method 80

9.3 Results 91

9.4 Discussion 94

10 EXPERIMENT 3: LOCAL CAUSAL BRIDGING INFERENCES: HOW DO THEY RELATE TO OFFLINE COMPREHENSION PERFORMANCE? 99

10.1 Hypotheses 100

10.2 Method 101

10.3 Results 107

10.4 Discussion 116

11 GENERAL DISCUSSION 120

11.1 Theoretical considerations 120

11.2 Methodological considerations 125

11.3 Practical educational implications 127

12 SUMMARY 129

ZUSAMMENFASSUNG 133

LIST OF FIGURES 137

LIST OF TABLES 138

BIBLIOGRAPHICAL REFERENCES 140

FILMOGRAPHICAL REFERENCES 155

APPENDIXES 157

1 Introduction

Moving pictures (films, videos) have come a long way since they were first presented by the Lumière Brothers in Paris in 1895. Nowadays, they dominate not only the entertainment sector in the form of big blockbusters, but they are also used for informational purposes. They have entered the internet in the form of digital video databases (http://www.youtube.com), news video clips (http://www.spiegel.de/video) or informational political video podcasts (http://www.bundeskanzlerin.de/Webs/BK/DE/Aktuelles/VideoPodcast/video-podcast.html). As these examples demonstrate, films clearly pursue an informational and educational purpose beyond their entertaining value. For instance, television, videotapes and films are more frequently used as instructional media in school than other media such as newspaper or magazine articles, computers and video cameras (Feierabend & Klingler, 2003). Moreover, informational science shows have now entered the TV landscape: for example, on Mondays, informational science shows are on from 2:00 pm in the afternoon until 11:00 pm at night across German television programming (Lehmkuhl, 2007).

Over the years, moving images have learned to speak and have grown into sophisticated forms of dynamic audiovisual media (e.g., films, digital videos, animations), which allow for a realistic, vivid, experience-driven way of conveying information to a broad public. A crucial precondition for conveying information successfully via films is, however, that film viewers elaborate the presented content deeply. Research on multimedia learning (Mayer, 2001, 2005; Sweller, 1999) suggests that films might lead to superior learning outcomes as they address different modalities (i.e., visual and auditory) and combine verbal and pictorial information, leading to an optimal distribution of processing load across memory subsystems. Contrary, the assumption that films are not elaborated deeply, but rather processed superficially, has dominated common wisdom (cf., “the film watching couch potato”) as well as educational research (e.g., DeFleur, Davenport, Cronin, & DeFleur, 1992; Salomon, 1984; Weidenmann, 1989, 2002). For instance, educational psychologists in the field of instructional dynamic visualizations (e.g., films, animations) have argued that the presentation of dynamic audiovisual media may result in a cognitive overload of the learner and at the same time in suboptimal elaboration processes (e.g., Lowe, 2003, 2004; Tversky, Bauer-Morrison, & Bétrancourt, 2002) or, on the other hand, that film viewers invest deliberately little mental effort because they assume that film easily processed (Salomon, 1984). However, the assumption that films are an ineffective educational medium under all conditions is questionable due to methodological shortcomings of research on film comprehension as well as due to the current exponential growth of instructional and informational films in different contexts like school, TV and the internet.

To yield deeper insights with regard to the educational potential of informational films (also called “expository films” in this dissertation), a more detailed and process-oriented methodological approach to study elaboration processes in film comprehension is advocated in this dissertation. Up to now, research on learning from expository film is characterized, first, by a large variability with regards to the materials used and the measures obtained and, second, by a focus on performance measures instead of process measures.

That is, researchers have inconsistently used different dependent variables (e.g., retention of factual knowledge, ability to answer different types of inference questions, ability to give coherent summaries). Additionally, they used different instructional materials such as news (Furnham, de Siena, & Gunter, 2002; Furnham & Gunter, 1985, 1989; Walma van der Molen, & van der Voort, 1997, 2000), commercials (Furnham, Benson, & Gunter, 1987; Furnham & Williams, 1987), narrative films (Baggett, 1979; Salomon, 1984), silent films (Salomon, 1984), and expository films from different domains (e.g., Furnham, Gunter, & Green, 1990). Accordingly, results from different studies can hardly be compared systematically. Moreover, usually only offline measures of learning and comprehension are obtained subsequently to film processing. For instance, educational psychologists have often focused on measures of learning outcomes obtained after the viewing of films (e.g., DeFleur et al., 1992; Salomon, 1984; Wetzel, Radtke, & Stern, 1994) and, as a second step, drew post-hoc conclusions regarding the elaboration processes that might have taken place during the film reception process. This method of measuring elaboration processes post-hoc is suboptimal to draw unequivocal conclusions referring to elaboration processes during film comprehension.

Hence, expository film reception processes have rarely been directly analyzed. As a consequence, this dissertation aims at developing online measures to investigate elaboration processes during expository film comprehension at a more fine-grained level. A prototypical type of elaboration processes are inference processes that are in focus of this thesis. The term inference is originally borrowed from logic, but is used in a slightly different sense in comprehension research (e.g., Kintsch, 1998). Inferences can broadly be defined as information that is not explicitly stated, but that the recipient generated himself[1] to understand the available information. A type of inference that is crucial for comprehension is that of “bridging inferences”, because bridging inferences semantically relate the current information with previously stated information (Graesser et al., 2002) and thus make the discourse coherent (e.g., Fincher-Kiefer & D’Agostino, 2004; Graesser, Millis, & Zwaan, 1997; Graesser, Singer, & Trabasso, 1994; Kintsch, 1988, 1998; Long, Golding, & Graesser, 1992; Wiley & Myers, 2003). Therefore, if film viewers generate bridging inferences during expository film reception, it can be assumed that expository films are elaborated deeply.

In sum, this dissertation aims at investigating

1) Whether expository films are elaborated deeply in terms of whether bridging inferences are generated during expository film comprehension;
2) Whether the presentation of dynamic pictorial information leads to a cognitive overload of the comprehender and, therefore, to suboptimal elaboration processes;
3) How inference processes during film reception relate to traditional offline measures of learning outcomes, such as retention and comprehension;
4) How inference processes during film reception can be measured.

[1] For better readability, the masculine form is used in this dissertation, but refers to both masculine and feminine.

This dissertation is divided into 12 chapters. Following the introduction in Chapter 1, Chapters 2 through 6 present the thesis' theoretical background. Chapters 7 through 10 describe the three empirical studies conducted in this thesis. Chapters 11 and 12 comprise a general discussion of the found results and a summary.

In Chapter 2, it is argued that theories of text comprehension processes are extendable to film comprehension processes (Section 2.2), although films and texts differ along several dimensions (Section 2.1). This is because discourse understanders in general try to generate a coherent mental representation of the presented content, independent from the presentation mode. Therefore, basic principles of text comprehension, such as the Construction-Integration Model, provide the theoretical framework of this thesis. These basic principles are introduced in Chapter 3. In Chapter 4, bridging inferences that are at the focus of this dissertation are described in more detail. As this dissertation addresses inference processes during film comprehension, existing empirical studies regarding the role of inferences in film comprehension are reviewed in Chapter 5. Moreover, the current thesis investigates how inference processes can be measured during film comprehension. As text comprehension theories are extendable to film comprehension, different inference detection methods derived from text comprehension research are introduced in Chapter 6 and are discussed regarding their suitability to measure inferences in film comprehension. Chapter 7 provides an overview about the empirical studies of this dissertation and states the research questions that are empirically addressed in Chapters 8 through 10, with each describing an empirical study to answer the aforementioned questions. In Chapter 11, the empirical findings of this thesis are discussed in the broader scope of educational, media and cognitive psychology and further implications with regards to empirical research and practice are drawn. Finally, Chapter 12 concludes with a summary of this dissertation.

2 Does Film Comprehension Equal Text Comprehension?

Text comprehension theories have been developed in the context of written texts. However, it is assumed that these theories also transfer to spoken texts, such as audio plays or radio shows, and - at least partially - to film or animation. A transfer from written texts to spoken texts is quite plausible (e.g., Furnham, de Siena, & Gunter, 2002; Kürschner, Schnotz, & Eid, 2006; Marx & Jungmann, 2000; Tompkins, Fassbinder, Lehmann-Blake, & Baumgaertner, 2001), as written and spoken texts just differ in their sensory modality (written texts: visual, spoken texts: auditory) and as both text forms transmit information verbally. In line with this reasoning, verbal information conveyed both visually as well as auditory is processed in one common subsystem of working memory, namely the phonological loop (Baddeley, 1992). In contrast, a transfer of text comprehension theories to film comprehension theories might be more complicated, as written texts and films differ along a variety of dimensions. These differences are discussed in Section 2.1 with regards to their instructional implications. Despite the differences between written texts and films, it has been widely assumed that text comprehension processes follow the same general principles as film comprehension processes (e.g., Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998). The reasons for this assumption are explained in Section 2.2.

2.1 Differences between texts and films

Different Modalities and Codalities

Different media are describable along two dimensions: modality and codality. While modality refers to the channel of sensory perception, such as visual and auditory, codality refers to the semiotic codes or symbol systems used (see Clark & Salomon, 1986), such as verbal and pictorial. Texts and films differ in both sensory modality and semiotic codes. Written texts are presented visually, spoken texts are presented auditorily and films are presented both visually and auditorily. Moreover, written and auditory texts present verbal information, whereas films present both verbal and pictorial information. While verbal information is considered to be mentally represented in propositions (e.g., Kintsch, 1988, 1998), pictorial information is assumed to be mentally represented analogous to the external representation, for instance, by means of mental images (e.g., Barsalou, 1999; Kosslyn, 1980, 1994; Zwaan, 1999, 2004).

Theories of multimedia learning (cf. Chandler & Sweller, 2001; Mayer, 2001, 2005; Sweller, 2005; Sweller, van Merriënboer, & Paas, 1998) have addressed the conditions under which the presentation of information in different modalities and different codalities leads to superior learning outcomes. The Cognitive Theory of Multimedia Learning (Mayer) and the Cognitive Load Theory (Sweller) are based upon Baddeley's (1992) working memory classification in a verbal (phonological loop) and a pictorial subsystem (visual-spatial sketchpad). Mayer's Cognitive Theory of Multimedia Learning is additionally based upon Paivio's Dual Coding Theory (1969, 1990). According to Paivio, long-term memory consists of a verbal and a pictorial subsystem. However, as has recently been critized, Paivio's and Baddeley's assumptions are codality-specific, whereas Mayer and Sweller confuse these codality-specific assumptions with modality-specific interpretations. For instance, Baddeley assumes that only visual pictorial material is processed in the visual-spatial sketchpad but not visual verbal material like written texts. Contrarily, Mayer and Sweller postulate that also visual-verbal material is processed in the visual-spatial sketchpad (for details see Rummer, Schwepper, Scheiter, & Gerjets, 2008). According to Mayer and Sweller, an optimal distribution of processing load across the subsystems is beneficial for learning processes, because the likelihood of a cognitive overload within one of the subsystems is reduced. Furthermore, according to Mayer's theory, a combination of verbal and pictorial information leads to a dual coding and to a more comprehensive mental representation in long-term memory.

Based on these considerations, it could be concluded that films should be beneficial for recall performances and learning outcomes. Mayer and Sweller have derived practical guidelines that address the optimal design of instructional media. For instance, according to the "multimedia principle", the simultaneous presentation of verbal and pictorial information (as in films) should enhance learning outcomes compared to single media, such as spoken or written texts. According to the "modality principle", the simultaneous presentation of pictures and spoken text (as in films) is more beneficial for learning than the simultaneous presentation of pictures and written text (as in some animations).

Superficial elaboration of films?

Although current multimedia theories might suggest that films are beneficial for learning, the assumption that films are not deeply elaborated is widespread (e.g., DeFleur et al., 1992; Salomon, 1984; Weidenmann, 1989, 2002). Five main reasons for that assumption are identified in the following:

Films are easy and print is tough

The assumption that learners consider films as easy and print as tough has been proclaimed by Salomon (1984). Salomon assumes that films are perceived to be easily processed, whereas texts are perceived to be cognitively demanding. These different demand characteristics of texts and films cause learners to invest different amounts of effort ("AIME – amount of invested mental effort"). AIME is defined as "the number of nonautomatic elaborations applied to material" (Salomon, 1984, p. 647) and refers to deliberately invested cognitive resources. The more cognitive resources invested, the better the learning outcomes. AIME depends on a) perceived demand characteristics, i.e., a subjective assumption whether the medium demands few or many cognitive resources and b) perceived self-efficacy, i.e., a subjective assumption on how efficient the personal performance is with regards to the respective medium. Salomon expects an interaction between those two factors with regards to AIME. Learners will invest more AIME when both factors are either low or high. Learners will invest fewer AIME when one factor is either low or high.

The AIME concept has been fairly influential in media psychology. However, its conceptualization is based on a criticizable methodological procedure. Salomon applied a silent film that is not characteristic for the medium film as the verbal commentary is missing. Therefore, the beneficial aspects of films, such as a dual encoding of the presented information in long-term memory (Paivio) or an optimal distribution of processing load (Mayer; Sweller), are not supported by a silent film. Accordingly, Beentjes and van der Voort (1993) found that viewing *audiovisual* television stories resulted in more inferential learning than reading print stories and that viewers' retention was superior to readers' retention in a delayed recall test. Therewith, they could not replicate Salomon's findings when applying audiovisual films. Despite these empirical inconsistencies, the assumption that print is easy and text is tough is still widespread in educational and media psychology.

Transient information presentation causes cognitive overload

Some authors (i.e., Lowe, 1999, 2004; Tversky, Bauer-Morrison, & Bétrancourt, 2002) reasoned that the dynamic presentation of complex dynamic stimuli in dynamic visualizations (film, animation) may cause a cognitive overload (Chandler & Sweller, 1991; Sweller, 2005; Sweller et al., 1998) because complex dynamic information is only available for a limited amount of time. To still process this transient information adequately, the learners need to store information in working memory, while new incoming information needs to continuously be processed. This constant holding, processing and connecting of information demands many cognitive resources and might result in a cognitive overload of the learner, also known as the "overwhelming effect" (Lowe, 2004). In line with this reasoning, Sturm (1984) postulates that there is a "missing half-second" ("fehlende Halbsekunde") in films that prevents a deeper elaboration of the presented content, as an inner verbalization, which is necessary to categorize the presented pictorial information, is not possible due to the transient presentation mode. This may impede elaboration processes.

Cognitive overload leads to suboptimal learning strategies

The transient presentation of information may exceed the learners' cognitive capacity and cause the learners to adopt selective information processing strategies that reduce cognitive load, but that may result in neglecting important information. For example, learners may focus on surface features with high perceptual salience (see Lowe, 1999, 2003, 2004) and on seductive details (see Harp & Mayer, 1998) instead of focusing on relevant contents.

Dynamic pictorial information inhibits mental processes

Some authors assume that the dynamic pictorial information in films or animations hinders a deeper elaboration because the recipients' imaginative elaboration is not addressed (Weidenmann, 1989). This assumption is supported by findings from Greenfield (1984), who found that children generated more pictorial imaginations when listening to audio plays than when watching films. In animation research, this effect has been called “inhibition effect” (Schnotz & Rasch, 2005), because the use of dynamic pictures might inhibit the learners from mentally animating the presented content and thus lead to a shallow processing. Lowe (2004) calls this the “underwhelming effect” because learners trust the illusion of understanding resulting in a superficial elaboration. As a consequence, learners may be able to memorize surface details, but may not generate a coherent mental model (Palmiter & Elkerton, 1993).

Self-regulated text reading versus medium-regulated film viewing

Self-regulated learning has been assumed to foster elaborated processing (Azevedo, 2005; Winne & Hadwin, 1998), because information processing adapts to the learners’ current cognitive needs. While texts can be processed in a self-regulated manner, regular film processing is determined by the medium. For instance, text learners can reread difficult passages, whereas film viewers have to follow the film, unless interactive features are provided (cf. Schwan & Riempp, 2004). Thus, unlike film viewers, text readers are able to exercise some control (Kozma, 1991) over the processed information, which is assumed to foster elaborated learning processes (e.g., Furnham & Gunter, 1985; Gunter, 1987). Film processing, however, is medium-determined and thus may easily cause a cognitive overload and, therefore, prevent learners from elaborating the contents conveyed.

Summary and conclusions

To sum up, written texts and films differ along a variety of dimensions that have been supposed to influence learning processes. Especially the role of the pictorial information for learning has been controversially discussed. While some authors argued that the combination of verbal and pictorial information is beneficial for learning (Mayer, Paivio, Sweller), some researchers argued that the pictorial information hinders learning (Lowe, Schnotz & Rasch). However, when concentrating on inference and comprehension processes, it can be argued that the discussed media differences are rather negligible, because both film and text comprehenders try to construct a mental model by generating inferences. Consequently, it can be assumed that text comprehension theories developed in the context of written texts extend to film comprehension (Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998). The following section justifies this assumption in greater detail.

2.2 Similarities between text and film comprehension processes

Independent from the obvious differences between text and film, it can be assumed that higher order mental processes lead to a medium-independent construction of a mental representation (Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998). This construction process is called “comprehension” (e.g., Kintsch, 1998).

Practical examples for the assumption of medium-independent comprehension processes are film adaptations of literature and informational films on, e.g., popular scientific contents. Literature and popular scientific information exist in both printed (e.g., books, magazines) and filmic formats (e.g., films, TV magazines) and the same contents are expressed and understood medium-independently. This is only possible if higher order mental processes that neglect the aforementioned media differences and focus on the existing commonalities are assumed. For instance, film and text follow the same overarching structural principles (Bordwell, 1985; Ohler, 1994) that lead to the same higher order comprehension processes (e.g., Zwaan & Radvansky, 1998). In other words, film and text can express the same content, which is understood independently from the medium, because both media follow the same structural principles. This assumption is supported by findings from Baggett (1979), who reported that students who saw a short silent film produced structurally similar recall protocols as students who heard an auditory version of the film that structurally equaled the film. This leads to the conclusion that the construction of a coherent mental representation implies higher order medium-independent cognitive processes.

In line with this reasoning, a medium-independent comprehension skill is assumed (Zwaan & Radvansky, 1998). People who are good at understanding texts should be good at understanding films. This was demonstrated by Gernsbacher, Varner, and Faust (1990). The authors had college student comprehend three stories in different formats (written and auditory text and a picture story) and administered comprehension tests. The students' performance on comprehension tests across the different presentation modes correlated significantly. The authors explained these findings with the assumption of higher level medium-independent comprehension processes and skills, such as the construction of a mental model via inferences.

This conclusion is also found in the research realm of event cognition (e.g., Magliano, Miller, & Zwaan, 2001; Newtson, 1973; Schwan & Garsoffky, 2008; Schwan, Garsoffky, & Hesse, 2000; Schwan, Hesse, & Garsoffky, 1998; Zacks & Tversky, 2001). It has been assumed that stories conveyed through discourse contain more information than what is explicitly stated. Rather, they convey a complex set of events that resemble real-world events. To understand events, comprehenders generate a mental model, independent from the "event conveying medium", such as text or film. In line with this reasoning, Magliano et al. (2001) demonstrated that the assumed construction of a so-called situation model along five dimensions, such as, e.g., time and space (Zwaan, Magliano, & Graesser, 1995) in text comprehension, also holds true for film comprehension. Film viewers consistently indicated situation changes on the two most basic dimensions of the situation model, namely time and space. These findings are in accordance with findings from narrative text comprehension where readers consistently identified changes in time and space (Zwaan et al., 1995). The authors interpreted their findings by the assumption of general comprehension mechanisms for event comprehension that operate independently of the event conveying medium.

In line with this reasoning, Graesser et al. (2001) assume similar inference processes in films and texts: As comprehenders make inferences when observing the real world -- and as texts and films reflect parts of the real world – comprehenders, therefore, also make inferences in reading texts and watching films. Correspondingly, Magliano et al. (1996) assume similar inference processes in texts and films, because media producers want the recipients to be caught up in the imaginative world. Accordingly, the authors could show that film viewers generate predictive inferences when triggered by filmic devices (see also Chapter 5).

Summary and conclusions

From this section, it can be concluded that different media, such as film and text, follow the same general comprehension principles. Comprehenders construct a higher-order mental representation that is media-independent. This does not preclude that some aspects of the resulting mental representation may vary depending on the medium used (e.g., in film comprehension, but not in pure text comprehension, pictorial information needs to be integrated with verbal information). Thus, it is hypothesized that in general inferences are generated in film comprehension, as the generation of inferences is crucial for the construction of a mental representation. However, it is an open question whether the film inherent dynamic pictorial information or the transient presentation mode might prevent a deeper inferential processing.

This chapter elaborated the differences between text and film and argued that it is – despite these differences – justified to extend text comprehension theories to film comprehension, because comprehenders of both media try to comprehend the presented content by constructing a medium-independent mental representation via inferences. The details of inference and comprehension processes are explained in the next chapter.

3 Principles of Text Comprehension

As explained in Chapter 2, text comprehension processes are comparable to film comprehension processes in spite of the existing differences between these two media. This chapter introduces the general principles of text comprehension according to the Construction-Integration Model (Kintsch, 1988, 1998, 2004). To comprehend a text, inferences are necessary. Inferences are introduced in Section 3.2. Inferences are needed to establish coherence and coherence is assumed to foster learning. The relationships between inferences, coherence and learning are discussed in Section 3.3.

This chapter starts by introducing different text genres, because it is assumed that text comprehension processes vary slightly as a function of different text genres (e.g., Zwaan, 1994).

Texts are divided into different text genres. Researchers have identified many categories of text genre, such as narrative, expository, persuasive, etc. (see e.g., Biber, 1988; Wolfe, 2005), but unequivocal definitions are lacking. For the purpose of this dissertation, only narrative and expository texts are of interest. Both genres encompass a broad range of actual texts. Brewer (1980, p. 223) defines narrative texts as "events that occur through time and are related through a causal or thematic chain. Narrative texts center on one or more protagonists who carry out different actions in order to satisfy a goal." By contrast, expository texts are defined as texts that describe the structure and processes involved in a system or event. Their primary function is to inform (Brewer, 1980). For the purpose of this dissertation, expository texts and, therefore, expository films are defined as educational, informational texts (films) that explain a set of entities and their interrelations. Their goal is to convey information and make the recipient learn.

In this dissertation, narrative texts are introduced, because most theories in text comprehension have been established in the context of narrative texts. These theories can be transferred to expository texts with slight modifications (see Figure 3.1; Table 4.1). Expository texts are introduced, because it is assumed that expository text comprehension processes extend to expository film comprehension processes that are at the focus of this thesis.

Figure 3.1 illustrates how theories of narrative texts are relevant for expository film. The dotted line indicates that although a direct transfer of theories from narrative texts to expository films is only possible to a limited extent, there are two indirect routes via narrative films and expository texts that can be taken.

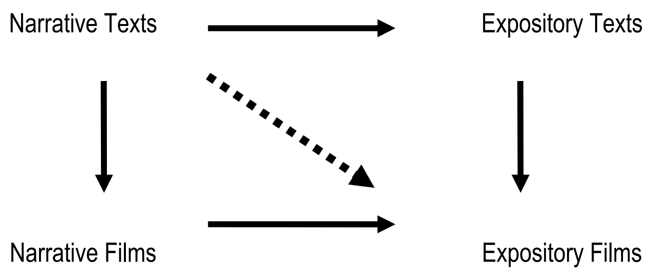


Fig. 3.1: Relationships between narrative texts, expository texts, narrative films, and expository films.

3.1 The Construction-Integration Model

One of the most dominant models in text comprehension research is the model of text comprehension by Kintsch and van Dijk (1978). It was first published in 1978, a revision followed in 1983 (van Dijk & Kintsch, 1983) and, in 1988, the Construction-Integration Model was introduced (Kintsch, 1988).

The Construction-Integration Model is an extended version of the text comprehension model by Kintsch and van Dijk (1978) and van Dijk and Kintsch (1983). It expands upon the earlier model and newly deals with knowledge use in text comprehension. Moreover, the Construction-Integration Model uses production rules to model text comprehension processes computationally. It is beyond the scope of this thesis to go into further detail regarding the computational aspects of the Construction-Integration Model. For an overview, see Kintsch (1988, 1998, 2004). For the purpose of this dissertation, only the structural aspects of the model are considered. Even though this model was originally developed in the context of narrative texts, its framework can be applied to expository texts, as well.

In the current chapter, the structural aspects of the Construction-Integration Model are introduced first. These include different levels of mental representations and the propositions that form the elements of these mental representations. To construct the different levels of mental representations coherently, inferences are needed. However, inferences will only be outlined briefly in this chapter, as the inferences relevant for this thesis will be discussed in more detail in Chapter 4. The relationships between coherence and comprehension are discussed in Section 3.3.

Mental Representations

The Construction-Integration model assumes that readers generate a coherent mental representation of the text they are reading. The coherent mental representation is called *episodic text memory*; the process of building a coherent mental representation is called *comprehension*. Once a reader has established a coherent mental representation of the text, it is assumed that he has comprehended the text. In other words, he has grasped the gist of the text rather than memorized the text.

To construct the episodic text memory, the reader mentally represents the meaning of a text in an interrelated network of *propositions*. A proposition is defined as “a basic unit of language” (Kintsch, 1998, p. 37) that transmits meaning. Additionally, the reader generates *inferences*. Inferences were earlier defined as information that is not explicitly stated in the text, but that the recipient understands by generating the implicit information himself. More specific definitions of inferences will follow in the course of this chapter.

The episodic text memory is a unitary structure, but, for analytical reasons, different components of this structure are distinguished. Most researchers agree on three levels of representation, namely: *surface*, *textbase*, and *situation model* (e.g., Graesser et al., 1994, 1997; Graesser, León, & Otero, 2002). Others add a genre level, a communication level, and a problem level. The mental representational levels are constructed by the reader in order to understand the text. This analytical distinction in several levels is done for research and instructional reasons. It is assumed that the mental representational levels are interdependent and interact to a degree that is not yet well understood (Graesser et al., 1997).

Even though most researchers distinguish three levels of mental representation (e.g., Graesser et al., 1994, 1997, 2002; McNamara et al., 1996), Kintsch (1998) does not consider the surface level to be a genuine mental representational level and, therefore, distinguishes only two mental representational levels: namely the textbase and the situation model. He regards the surface level that represents the exact wording and syntax of the text, to be part of the textbase, next to the propositional structure. The surface level is not considered a true mental representation on the side of the reader, but rather just an image of the text, as mental activity, such as establishing meaning by generating inferences, has not yet occurred.

For each level of mental representation, Kintsch distinguishes between the microstructure, i.e., propositions for the local content, and the macrostructure, i.e., propositions for the global structure of the text.

This dissertation follows Kintsch's distinction and focuses on the textbase and situation model as mental representational levels. When trying to comprehend a text, the ideal result is a locally and globally coherent textbase and situation model (for details see Section 3.3).

The *textbase* comprises those propositions that can be directly derived from the text by means of inferences. These inferences, however, might require some degrees of prior knowledge, especially in expository texts. For instance, referents of pronouns must be identified, synonymous terms must be matched, possible coherence gaps must be closed, etc. Some authors claim that this is the minimal amount of active processing a cooperative and motivated reader performs (McNamara & Kintsch, 1996). On the basis of the textbase, a reader can answer questions about the text, verify statements about the text, recall the text, and summarize it.

In most cases, however, a deep comprehension requires more than the textbase. To understand a text, a reader usually needs to integrate the new information into existing prior knowledge structures, must add propositions to the existent propositional network and establish links from his own prior knowledge and experience. The resulting mental representation is called the *situation model*.

The situation model is based on explicitly stated information in the text, general and/or specific prior knowledge, and inferences generated by the reader. To construct the situation model, various sources of knowledge are relevant: knowledge about the language, general world knowledge, knowledge about the specific communication situation, domain-specific prior knowledge, and personal experiences. In this dissertation, the focus is on the role of the domain-specific prior knowledge and its prerequisite knowledge, that is, the knowledge that is necessary to understand this prior knowledge. Therefore, the term prior knowledge will be used comprehensively for domain-specific prior knowledge and its corresponding prerequisite knowledge throughout this dissertation, unless otherwise stated.

For an expository text on a scientific topic, e.g., a mechanical system, the situation model consists of a sketch of the system components, the processes that occur when the system is operating, the relations between components and functions of the system, and the various uses of the system by humans (cf. Graesser & Bertus, 1998; Maury, Pérez, & León, 2002).

Some authors introduce a 4th level of representation, especially in expository text comprehension, a so-called *Problem Model* (Fischer, Henninger, & Redmiles, 1991; Nathan, Kintsch, & Young, 1992). Good readers create a level of representation called “the problem model”, which is constructed taking into account the formal (mathematical) relations that exist between the elements described in the statement of a problem. For this process, the reader needs -- in addition to his world knowledge -- scientific and mathematical knowledge (prior knowledge) on the relations between the variables involved (Graesser et al., 2002). The (*pragmatic) communication level* specifies the main message that the author wants to convey to the reader, for example, how to repair equipment or how a thermos is constructed (Graesser et al., 1997, 2002). Graesser et al. (1997) also name the *text genre* as a level of mental representation. However, the mental representational levels that have acquired the most attention in research are the textbase and the situation model. These are the two levels the current dissertation focuses on, as well.

According to Kintsch (1983, 1998), the textbase and the situation model construct the *episodic text memory* of the reader. The episodic text memory forms the end product of the comprehension process (Adam & Butler, 1999). In educational psychology, the concept of a *mental model* (cf. Johnson-Laird, 1983) is often used to describe the end product of a comprehension or learning processes. However, whereas situation model and episodic text memory consist of propositions, a mental model in the sense of Johnson-Laird mainly relies on analogous (i.e., mental images) and procedural representations. Despite these conceptual differences, educational psychologists often use the aforementioned terms synonymously and apply the term mental model in a broader context; whereas the term episodic text memory is hardly applied. To be in line with the common usage of the concept of a mental model in the realm of educational psychology, the terms mental model and episodic text memory are used interchangeably in this dissertation when it comes to learning processes. In this terminology, a mental model consists of the textbase and the situation model.

The Construction-Integration Model’s name is based on two successive stages assumed in the comprehension process. The initial phase, i.e., construction phase, is governed by production rules that activate all possible concepts associated with the words in the text. During the following integration phase the concepts with the strongest associations win by receiving the highest levels of activation and are transferred to working memory. It is assumed that the comprehension process comprises several processing cycles of these two phases that permanently update the information already extracted from the text, as well as the new incoming information, and integrate both with prior knowledge structures. Consequently, the text comprehension process requires a permanent exchange process between working memory and long-term memory.

To better understand the comprehension process and the interrelations between text, textbase, situation model, episodic text memory and prior knowledge, the reader is referred to Figure 3.2. Straight lines represent different types of inference processes whereas dotted lines indicate the contribution of prior knowledge. The quantitative degree to which prior knowledge structures matter for a successful inference generation is illustrated by the thickness of the dotted lines.

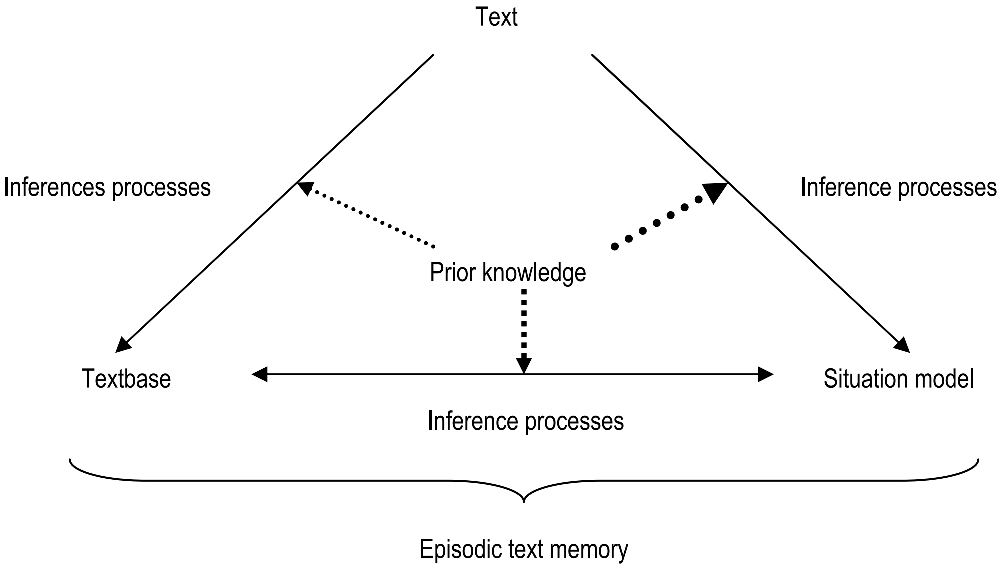


Fig. 3.2: The comprehension process.

Propositions

According to Kintsch and other discourse psychologists (e.g., Kintsch, 1998; Lorch Jr. & van den Broek, 1997; van Dijk & Kintsch, 1983), meaning is mentally represented in propositions. Propositions reflect the meaning of a text and are more closely related to the mental representation of meaning than regular language. Regular language serves other purposes than just representing meaning and is, therefore, often less suited for text comprehension research.

When reading a text, the reader generates a propositional network that includes propositions with various strengths. Propositions can be derived directly from the text (textbase) or can be derived from textbase integration with long-term memory structures, such as prior knowledge (situation model).

The term “proposition” is borrowed from logic, but is used in a different sense in text comprehension research. For the purpose of text comprehension, a proposition is a meaning unit of language with the form of a predicate-argument schema. Each proposition contains a predicate (e.g., main verb, adjective, connective) and one or more arguments (e.g., nouns, embedded propositions). Each argument has a functional role, such as agent, patient, object, or location. Van Dijk and Kintsch (1983) distinguish atomic propositions from complex propositions:

Atomic propositions consist of a relational term, the predicate and one or more arguments. Written as PREDICATE [ARGUMENT, ARGUMENT][2]. The predicate determines the number and kinds of arguments that may fill argument slots. For example

Mary gave the book to Fred
GIVE [MARY, BOOK, FRED]

Complex propositions are compounds composed of several atomic propositions that are subordinated to a core proposition, such as in

Yesterday, Mary gave the old book to Fred in the library.
P1 GIVE [MARY, P2, FRED]
P2 OLD [BOOK]
P3 YESTERDAY [P1]
P4 IN-LIBRARY [P1]

Note that not all expressions and information in the text are represented in the propositional notation (e.g., tense, definite article, passive or active voice).

Empirical evidence for the psychological correspondence of propositions comes from three sources: recall, reading times, and priming studies (e.g., Goetz, Anderson, & Schallert, 1981; Kintsch, 1998; van Dijk & Kintsch, 1983; Wanner, 1975). Wanner (1975), for example, had participants study the following sentence in a cued recall experiment:

The mausoleum that enshrined the tzar overlooked the square.

Propositional notation:
P1 OVERLOOK [MAUSOLEUM, SQUARE]
P2 ENSHRINE [P1, TZAR]

Participants who are given the word *overlooked* as a recall cue are more likely to remember the word *square* than *tzar*. *Overlook* and *square* are arguments of the same proposition; *tzar* and *overlooked* are not. Participants remember the word *square* more often even though *overlooked* and *tzar* are closer together in the surface structure than *overlooked* and *square*.

[2] Note that slightly different propositional notations exist side by side that differ in their format (Kintsch, 1998, 2004).

Thus far, it has been explained that, according to the Construction-Integration Model, there are two levels of understanding, namely the textbase and the situation model. Both levels form the episodic text memory. Meaning is mentally represented in propositions that are the elements of the episodic text memory. To generate the episodic text memory, different types of inferences are necessary. The following section addresses these inferences.

3.2 Inferences

To extract meaning from the simple letters that form the text, the reader needs to generate inferences. Inference researchers have proposed several inference taxonomies (cf. Graesser et al., 1997, 2002, 2007) that share many common features, but can be distinguished by using different names for the same types of inferences or by distinguishing a different number of inference types.

An inference taxonomy derived in the context of scientific texts is exemplarily introduced in Table 3.1 (Graesser et al., 2002). For the purposes of this dissertation, the taxonomy has been slightly modified and shortened. It is important to mention that this taxonomy is not exhaustive and that there are other taxonomies that differentiate a greater number of inference types, mainly for narrative texts (e.g., Graesser et al., 1997, 2007). However, to understand the current dissertation, knowledge of these types of inferences is sufficient. Inference types 1 through 2a are mostly assumed to accompany comprehension and are thus assumed to be generated online (i.e., immediately and/or automatically). Inference types 4 and 5 will not be discussed in this dissertation and are only included in the table for the sake of completeness.

Table 3.1: Inferences generated in scientific text comprehension (modified according to Graesser et al., 2002).

1) Anaphoric references	A pronoun or noun-phrase refers to a previous text constituent or to an entity already introduced in the mental model.
2) Bridging inferences	These inferences are needed to semantically or conceptually relate the sentence currently being read with the previous content. These are sometimes called backward inferences. A special type of bridging inference is the causal bridging inference.
2a) Causal bridging inferences	The current event being read is explained by a causal chain or network of previous events and states. These are sometimes called causal antecedent inferences.
3) Predictive inferences	The reader forecasts what events will causally unfold after the current event being read. These are sometimes called causal consequences or forward inferences.
4) Goal inferences	The reader infers that an agent has a motive that explains an intentional action.
5) Process inferences	These inferences specify the detailed steps, manner, or dynamical characteristics of an event as it unfolds.

Inferences are needed to generate the episodic text memory. As the episodic text memory is distinguished in two mental representational levels, inferences can be analytically distinguished in a) inferences that are necessary to form the textbase and in b) inferences that are necessary to generate the situation model. All inferences are needed to form a coherent episodic text memory. See also Figure 3.1.

Again, the reader should keep in mind that this distinction is mainly an analytical one that may be hard to transfer to real life. However, it helps one understand the human text comprehension processes. Thus, this dissertation follows this distinction, but emphasizes the analytical purpose. In the following passage, inferences necessary to form the textbase are presented first. Next, inferences necessary to form the situation model are presented.

Inferences needed to form the textbase level

To develop the textbase level, the reader needs to draw a) anaphoric references and b) bridging inferences.

Anaphoric references

An *anaphoric reference* is defined as a pronoun or noun-phrase that refers to a previous text constituent or to an entity already introduced in the mental representation (Graesser et al., 2002).

E.g.: *We got the beer out of the car. It was warm.*

The reader will automatically infer that the “it” refers to the beer. It could also refer to the weather, but this makes less sense here. The anaphoric relation between “it” and “beer” is automatically established, while the syntactic possible relation between “it” and “the weather” is automatically suppressed. When you read the two sentences, you probably did not even notice that it is not explicitly stated that it was “the beer that was warm”.

Bridging inferences

The same holds for *bridging inferences* that are needed to semantically or conceptually relate the current sentence being read with the previous sentence (Graesser et al., 2002). Haviland and Clark (1974) state that bridging inferences are generated when anaphoric resolution fails and that bridging inferences preserve coherence by establishing connections among explicit text ideas. Most researchers (e.g., Black & Bern 1981; Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Graesser et al., 1994, 1997, 2002, 2007; Potts, Keenan, & Golding, 1988; Singer, 1994; Singer & Ferreira, 1983; Wiley & Myers, 2003) assume that bridging inferences are generated *online*, i.e., they are generated automatically, within a second or less, without taking up a lot of cognitive resources. Look at the following example:

a) We got the picnic supplies out of the car. b) The beer was warm.

To understand those two sentences you must automatically generate an idea that links the two sentences, such as "the picnic supplies included beer". This linking idea bridges sentence a) and b) and is, therefore, called *bridging inference*.

Bridging inferences are often so automatically generated that the reader does not even notice that he drew a bridging inference. Check for yourself: When you first read those two sentences a) and b), you probably did not notice that it is not explicitly stated that the picnic supplies included beer. If you were asked if the beer was taken out of the car, you would probably answer "yes" without even thinking about it, because you inferred that the picnic supplies included beer. This is what researchers call a bridging inference.

As the understander's working memory resources are limited (Baddeley, 1992), the number and variety of inferences that accompany comprehension (i.e., are generated online) are significantly restricted. As a consequence, one of the most scrutinized questions in inference research is what kind of inferences are generated online and under what conditions (e.g., Graesser et al., 1994, 1997; Singer, Graesser, & Trabasso, 1994).

Inferences needed to generate the situation model

To form a coherent mental representation of the text, the reader needs to generate a situation model. To transform the textbase into a situation model, inferences are again necessary. Often, the inferences that are generated to construct the situation model are comprehensively called *elaborative inferences* (e.g., Durgunoğlu & Jehng, 1991; Long et al., 1992; McKoon & Ratcliff, 1986; Potts et al., 1988; Singer & Ferreira, 1983). Elaborative inferences "embellish the text representation" (Long et al., 1992; p. 635), but are not required for text coherence and are thus not essential for an online (i.e., immediate) comprehension. Their generation takes up more cognitive resources and they are thus not generated online, rather they are more likely to be generated after the reception process or when directly asked a question for which the reader needs the inferred information (McKoon & Ratcliff, 1986, 1990). That means that elaborative inferences are not generated automatically during the comprehension process, but rather need to be computed strategically. Elaborative inferences require going beyond the text and deriving unspecified text aspects, such as consequences of events (McKoon & Ratcliff, 1986; Potts et al., 1988) and specification of ongoing states (Seifert, Robertson, & Black, 1985). Elaborative inferences require an integration of the inferred information with long-term memory structures, such as prior knowledge, more than anaphoric references or bridging inferences do. Therefore, their generation lasts longer and takes up more cognitive resources. In general, inferences that are generated within more than a second are assumed to be generated offline (Graesser et al., 2007), even if they are generated during the reception process.

Examples of elaborative inferences that are considered to be generated offline are *predictive inferences* (Graesser et al., 1994, 1997, 2002, 2007). Predictive inferences answer the question: "What will happen next?" and are defined as the prediction of what events will causally unfold after the current event being read. Predictive inferences are sometimes called causal consequences or forward inferences (Graesser et al., 2002, see also Table 3.1, Section 4.2, and Chapter 5). Even though predictive inferences should occur during the comprehension process, they are not generated automatically, as they are not needed for an immediate (i.e., online) understanding. Moreover, their generation requires more cognitive resources and processing time than online generated bridging inferences (Graesser et al., 2007). Thus, predictive inferences are assumed to be generated offline.

Online versus offline inference processes

The aforementioned considerations made evident that inferences necessary to generate the textbase are considered to be generated online, whereas inferences necessary to generate the situation model are considered to be generated offline. However, the distinction in online versus offline inference processes is not always clear. In inference research, the distinction has often been made whether the inference processes occur during the reception process (i.e., online) or after the reception processes (i.e., offline). But, when looking at the example of predictive inferences that are claimed to be generated offline, it becomes evident that the temporal distinction of inference processes in *during* versus *after* the reception process is too restrictive. While online inferences are always generated during the reception process, offline inference processes might occur during the reception process but also after the reception process. As they are not necessary for an immediate understanding, they are generated during the reception phase only strategically, when taking breaks to think about a passage or when asked a specific question. As stated above, the generation of offline inferences takes up more cognitive resources and requires more processing time. However, their generation is necessary to generate a coherent situation model.

These considerations made clear that a situation model may be generated even during the reception process, but that the generation of a situation model will mostly not occur during an immediate, online processing of the content presented. Rather, during an online processing, bridging inferences are generated necessary to construct a coherent textbase. On the contrary, offline inferences are not generated spontaneously and are considered to be generated optional, depending on the reader's cognitive prerequisites and personal reading goals. However, especially for expository discourse comprehension, it seems as if the distinction in online generated bridging inferences on the one hand and offline generated elaborative inferences on the other hand cannot always hold true in real life. Sometimes closing coherence gaps via bridging inferences relies substantially on prior knowledge and requires thus, a strong involvement of long-term memory structures such as prior knowledge during the generation of the textbase that is supposed to be generated online.

In this dissertation, the term „online” is used for a spontaneous, immediate, automatic processing that occurs during the reception process within a second or less. This kind of processing holds also, if the reader is instructed to follow a certain reading goal, e.g., trying to understand the content presented, and conducts automatic bridging inferences to pursue this goal. If the reception process is interrupted within less than a second for a bridging inference generation, then this inference is still considered to be generated online. The term „offline“ on the other hand, refers to resource-intensive inference processes that are either generated during or after the reception process. Offline inference generation is assumed to last more than a second and thus, when generated during the reception process, the reception process is interrupted for more than a second. An example for offline generated inferences during the reception process are predictive inferences that are generated during the reception process, but they are not necessary for an immediate understanding. In contrast, the generation of a bridging inference does not require the interruption of the reception process and is conducted within a second or less.

In sum, online and offline processes can be described on two dimensions, i.e., time and automaticity. Whereas online refers to an immediate, spontaneous (time) and automatic, not resouce-demanding (automaticity) processing, offline refers to a delayed (time) and strategic, resource-demanding (automaticity) processing.

The distinction in online and offline inference processes is reflected in the distinction of online and offline methods to measure inference processes (see Chapter 6). Even though the distinction in online and offline processes pervade inference research, the distinguishing criteria are not always unequivoval as the aforementioned remarks might have shown. However, one fact that has been widely agreed on is that different classes of inferences are needed to deeply understand the content presented. It is assumed that generated inferences enhance the degree of coherence and thus understanding. Coherence is either achieved by explicit relations in the text or by inference processes. Coherence results in better learning outcomes, such as retention and comprehension. This correlation is explained further in the following section.

3.3 Coherence and comprehension

To comprehend a text means to establish a coherent mental representation. If no coherent mental representation is constructed, comprehension is supposed to be suboptimal. But how is coherence achieved and how does coherence relate to comprehension? The next passage provides answers to these questions.

Levels of coherence

Texts consist of several complex propositions that are related to form a coherent mental representation. By definition, unrelated propositions do not form a text or discourse. Three levels of coherence among propositions are distinguished (van Dijk & Kintsch, 1983; Kintsch, 1998). Coherence is either directly stated in the text or the comprehender generates propositions (inferences) to make the text coherent.

1) *Indirect coherence*: the meaning units are part of the same episode. That is, they share a time, place, or argument.
2) *Direct coherence*: the same as indirect coherence, but, in addition, the coherence is specifically marked by separate clauses or sentences. Sentence adverbials, such as *therefore*, *then*, *so*, *as a result,* etc., compound sentences, or explicit coordinating connectives, such as *because*, *that's why,* etc., establish direct coherence, i.e., link the propositions directly.
3) *Subordination*: one proposition is subordinated to another. This might be the case when one proposition conditions the other, i.e., when one proposition is embedded in another proposition. When the whole proposition (in its entirety) is not linked to the other propositions, but rather just an argument, such as a specification of the agent or action, then restrictive relative clauses are used.

To illustrate how coherence is expressed or generated by the reader, look at the following minitext:

The snow was deep on the mountain. The skiers were lost, so they dug a snow cave, which provided them shelter.

P1 DEEP [SNOW, MOUNTAIN]
P2 LOSE [SKIERS, MOUNTAIN]
P3 DIG [SKIERS, SNOW CAVE, MOUNTAIN]
P4 PROVIDE [SNOW CAVE, SKIERS, SHELTER]

The *coherence* between the first two propositions is *indirect* because they share an argument, the place: the skiers were probably also lost on the mountain. However, this argument overlap is inferred by the reader, it is not explicitly stated in the text. The second and third propositions are linked by *direct coherence*: the adverbial *so*. Additionally, the second and third propositions are linked by argument overlap for agent and place: the skiers (agent) are presumably still on the mountain (place). Moreover, a "*so"* links the second and third proposition directly. The fourth proposition is embedded in the third proposition as a specification for one of its arguments: the snow cave (argument of the third proposition) provides them shelter. This is easily recognizable by the relative clause indicated by "which". Therefore, the third and fourth propositions are linked by *subordination*.

Often, argument overlap has been considered as the primary dimension for establishing text coherence (Kintsch & van Dijk, 1978; McKoon & Ratcliff, 1992). Coherence is achieved if there is an argument in the incoming proposition that overlaps with an argument within any textbase proposition in working memory, stated above as *indirect coherence*.

How does coherence relate to comprehension?

Comprehension requires that the learner constructs a coherent mental representation of the text called the "episodic text memory". To form a coherent episodic text memory, a coherent textbase and a coherent situation model are necessary. Once the reader has established an episodic text memory, it can be concluded that the reader has understood the text at a deeper level.

One reason why a reader fails to comprehend a text might be because he fails to form a coherent textbase and/or a coherent situation model. If coherence is not expressed in the text (as in most cases), inferences are necessary to form new propositions that make the mental representation of the text coherent. To enable an online understanding of the text, usually semantic information is sufficient to draw bridging inferences. However, to enable a deeper understanding, the propositional textbase needs to be integrated with long-term memory structures, such as prior knowledge, to generate a coherent situation model.

How text, textbase and situation model interact is illustrated in Figure 3.3. At the top of the figure, a small passage from an expository text is shown. The propositions in italics found in the upper half of the figure are those propositions that are derived from the text and that together with the propositions from prior knowledge (left upper half, but not directly connected to the propositions derived from the text) form the propositional textbase. When integrated with prior knowledge (the cuboids in the bottom that represent the circuit system, the arrows in between the cuboids, and the proposition not in italics at the bottom), then a coherent situation model is formed as indicated by the vertical arrows. Here, it has to be noted, that Kintsch applies the idea of an analogous mental representation as demonstrated by the depiction of the circuit system. However, the situation model mainly constitutes of propositions.

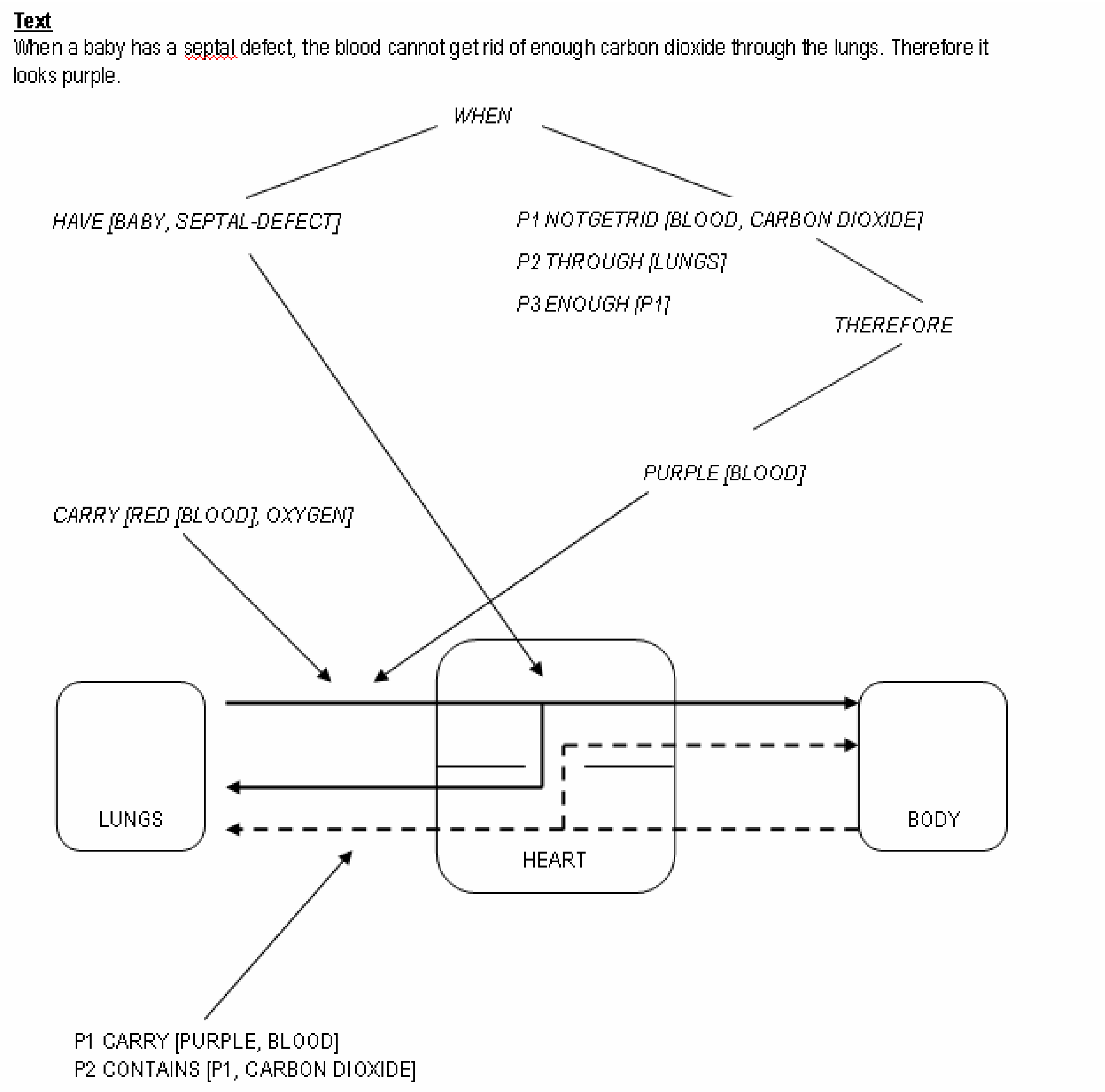


Fig. 3.3: Text, textbase, and situation model and their interrelations according to Kintsch (1994, 1998).

One way of increasing comprehension, and thus learning outcomes, is to increase the coherence of a text (e.g., Kintsch, 1998, 2004). Studies on fifth-grade students (Beck, McKeown, Sinatra, & Loxtermann, 1991) and college students (Beyer, 1990; Britton & Gülgoz, 1991) have shown that revisions that increased the coherence of a text increased recall and comprehension performance. Britton and Gülgöz (1991) found a -.80 correlation between coherence gaps (call for inference) and recall scores. They used a computer simulation of the comprehension process of the original text (Kintsch & van Dijk, 1983; Miller & Kintsch, 1980). The simulation broke down whenever there were coherence gaps from one proposition to the next or from one sentence to the next. These coherence gaps were positions that “called for inferences”.

Whenever the computer simulation broke down, usually the reader would have to generate bridging inferences in order to close the gaps. The authors improved the coherence of the text by adding explicit information to bridge the coherence gaps. This is what the reader would have usually done by generating a bridging inference. Presenting a coherent text version to the readers increased recall performance. Especially readers with low background knowledge benefited from the more explicit text versions, as shown by their improved recall performance. Linderholm et al. (2000) could show similar beneficial effects on comprehension outcomes for causal text revisions: readers benefited from causal revisions of texts, i.e., when causal relations were made explicit, but only for difficult texts.

These considerations demonstrate that one effect of increasing the coherence of a text is to increase comprehension, as readers do not have to infer the implicit information, because it is explicitly stated in the text. However, an increased coherence may result in reduced active processing during reading. This is because the reader doesn’t have to figure out the meaning of the text, which results in a less effective learning strategy, boredom and superficial text processing. For instance, McNamara, Kintsch, Songer, and Kintsch (1996) found that high prior knowledge readers scored higher in the dependent measures *free recall*, *written questions* and *a key word sorting task* in a less coherent text than low prior knowledge readers. Low prior knowledge readers, on the other hand, benefited from a coherent text that didn’t reveal any coherence gaps. This leads us to the conclusion that active text processing – if enabled by prior knowledge – can be fostered if the text is somewhat incoherent, but mainly for high prior knowledge readers. The latter can engage in compensatory processing as shown by higher scores for the dependent measures that measured comprehension, such as problem solving questions, bridging inference questions, elaborative inference questions, as well as for the key words sorting task, when learning from a low-coherent text version.

Summary

In this chapter it was explained that text understanders generate a coherent mental representation of the text they are reading. Two different levels of mental representation are distinguished (i.e, textbase and situation model) that form the episodic text memory (cf. mental model in this dissertation). To construct the episodic text memory, the reader mentally represents the meaning of a text in an interrelated network of propositions and generates different types of inferences. Inferences are needed to form a coherent mental representation and high degrees of coherent mental representations can lead to high comprehension scores. Bridging inferences are crucial for online comprehension, whereas elaborative inferences are often found to be generated offline. As bridging inferences are one focus of this dissertation, they are described in further detail in the next chapter.

4 Bridging Inferences in Text Comprehension

The current dissertation aims at analyzing online inference processes. Bridging inferences are prototypical online inference processes that are crucial for comprehension. Therefore, the current chapter describes bridging inferences in further detail. Two important distinctions are made. First, bridging inferences are divided into local vs. global bridging inferences. Second, causal bridging inferences -- as a special subtype of bridging inferences -- are introduced. Causal bridging inferences can be either local or global and bridging inferences can be either causal or non-causal.

Local versus global bridging inferences

Bridging inferences preserve coherence by establishing connections among explicit text ideas (Haviland & Clark, 1974). Some researchers (e.g., Albrecht & O'Brien, 1993; Graesser et al., 1997; McKoon & Ratcliff, 1992; Myers, O'Brien, Albrecht, & Mason, 1994; Singer et al., 1994; Tapiero & Otero, 2002) distinguish local from global bridging inferences. Local bridging inferences involve pieces of information that are stated closely together in a text, i.e., not more than two or three sentences apart (e.g., Eysenck & Keane, 2005; McKoon & Ratcliff, 1992; see Experiment 2 in this dissertation); whereas global bridging inferences connect pieces of information that are widely separated throughout the text (e.g., Graesser et al., 1997; see Experiment 1 in this dissertation).

Online or offline?

As mentioned above, one of the most scrutinized questions in inference research is what kind of inferences are drawn online and under what conditions. There are contradictory theoretical assumptions as to what type of bridging inference (local vs. global) is generated online. These two theoretical positions can be considered to form the poles of a continuum and will be described briefly in this section.

Whereas the Minimalist Hypothesis (McKoon & Ratcliff, 1992, 1995) assumes that global bridging inferences are *only* generated online when local coherence fails to be established or when global bridging inferences can be based on rapidly available information, the Constructionist Theory (Graesser et al., 1994, 1997, 2007) assumes that global bridging inferences are generated online under a wider range of circumstances.

The Minimalist Hypothesis has been heavily criticized for several reasons (e.g., Albrecht & O'Brien, 1993; Singer et al., 1994; Tapiero & Otero, 2002). One subject of criticism was the definition of the Minimalist Hypothesis' central concepts of automaticity, availability, and local coherence. For instance, regarding the definition of automaticity, McKoon and Ratcliff claim that automatic inferences are constructed in a few hundred milliseconds and that they are computed in the absence of specific strategies and awareness.

Therewith, they define their concept of automaticity in line with Schneider and Shiffrin's (1977) definition of automatic processes. However, it is questionable whether the automatic-strategic distinction according to Shiffrin and Schneider applies to discourse processing. For instance, a classic facilitation effect in the domain of automatic processing is the so-called priming effect, i.e., that the word *bread* immediately preactivates the word *butter*. In discourse processing, however, Sharkey and Sharkey (1992) reported that a target is not facilitated by an immediately preceding prime that completes a sentence, but is facilitated when the same prime completes a scrambled list of the sentence words. Thus, it seems that the profiles of processing word lists and sentences are different (Singer et al., 1994). For instance, it seems that automatic inferences last longer than a few hundred milliseconds as proposed by Schneider and Shiffrin for automatic processes. Nevertheless, McKoon and Ratcliff base their definition of automaticity on Schneider and Shiffrin's concept of automatic processing.

Additionally, McKoon and Ratcliff's theory seems to imply that adopting a specific reading strategy nevertheless involves minimal inferences that are computed automatically. Insofar as automatic and strategic processes form a dichotomy according to Schneider and Shiffrin (1977), this conclusion is paradoxical (cf. Singer et al., 1994).

In sum, the rather hazy distinction between automatic and strategic processes has been considered one major theoretical flaw in the Minimalist Hypothesis. Moreover, there is empirical evidence challenging the fundamental assumption of the Minimalist Hypothesis that global bridging inferences are not generated when the text is locally coherent (e.g., Singer, 1993). For an extensive discussion of McKoon and Ratcliff's Minimalist Hypothesis (1992, 1995) and its shortcomings, the reader is referred to Singer et al. (1994).

Due to the inconsistencies of the Minimalist Hypothesis, this dissertation follows the view of the Constructionist Theory and those researchers who assume that local *and* global bridging inferences are generated online (Albrecht & O'Brien, 1993; Graesser et al., 1994, 1997; Singer et al., 1994). Accordingly, as one focus of this dissertation is online inference processes, global bridging inferences are seen to be within the domain of this dissertation.

4.1 Causal bridging inferences in narrative texts

As stated previously, a subtype of bridging inferences is causal bridging inferences. Causal bridging inferences establish coherence by linking text ideas that are related *causally*. Most researchers agree that causal relations are crucial for the understanding of discourse, narrative texts, as well as expository texts (e.g., Graesser & Bertus, 1998; Millis & Graesser, 1994; Trabasso & Suh, 1993; Trabasso & van den Broek, 1985; Trabasso, van den Broek, & Suh, 1989). Causal bridging inferences refer both to physical causes and motivational causes (Singer, 1994).

As in the context of non-causal bridging inferences, investigators have analyzed whether causal bridging inferences are generated online in the comprehension of narrative texts. In this regard, *local* causal bridging inferences (LCBIs) were at the focus of investigation[3]. Murray Singer contributed a great deal to causal bridging inference research. He and his colleagues (Singer & Halldorson, 1996; Singer, Halldorson, Lear, & Andrusiak, 1992; Singer, Revlin, & Halldorson, 1990) conducted several studies that followed more or less the following procedural pattern:

Participants read either sentence a) or sentence b)[4].

a) *Dorothy poured the bucket of water on the fire. The fire went out. (causal)*
b) Dorothy placed the bucket of water by the fire. The fire went out. (temporal)

After each sequence, participants had to answer a question that probed the knowledge hypothesized to be involved in the construction and validation of the causal bridging inference:

c) Does water extinguish fire?

Answer latencies were shorter when participants read causally related statements compared to the control condition stating temporally related statements (Singer et al., 1990, 1992). This leads one to the conclusion that local causal bridging inferences (LCBIs) are generated online and not at the time the question is asked (i.e., offline), as was the case for participants in the temporal condition. Participants in the causal conditions showed shorter latencies, because at the time they were asked the question, they had already generated the LCBIs (instead of generating the LCBI at the time the question was posed). Participants in the temporal condition had not generated LCBIs online and, therefore, they had to generate them when asked if water extinguishes fire. This prolonged their latencies in answering the question.

The idea that inference generation takes up processing time and that this is reflected in longer latencies, if the inference was not generated online, but needed to be generated offline - or reflected in shorter latencies, if the inference was generated online - underlies many methods that measure inferences online. For an extensive review of various methods in inference research, see Chapter 6.

Singer and his colleagues explained the aforementioned findings by proposing a *validation model* for causal bridging inferences in narrative texts. Let's look at the causal sentence from the above-mentioned example once more:

a) *Dorothy poured the bucket of water on the fire. The fire went out.*

[3] The cited references all refer to local causal bridging inferences (LCBIs).
[4] The preposition was not underlined in the experimental material, but is underlined here to demonstrate the difference between the two sentences.

The validation model states that before a causal bridging inference is accepted, it must be validated on the basis of relevant world knowledge (Singer & Gagnon, 1999; Singer & Halldorson, 1996). First, the reader identifies a mediating idea that links the causal antecedent phrase to the cause, (for the example: a mediating idea that links the information that Dorothy poured the water on the fire, causing the fire to go out). The mediating idea is that water extinguishes fire. Second, the mediating idea must be validated on the basis of relevant background knowledge. For the present case, long-term memory provides the information that water indeed extinguishes fire and the inference is successfully validated. To test this idea, Singer (1994) presented sentences such as *Dorothy poured the gasoline on the bonfire. The bonfire went out.* Readers were able to detect the inconsistency as reflected in longer reading times. This is interpreted as an indicator that validation processes took place.

Prior research in the context of narrative texts has shown that the computation of local causal bridging inferences, as well as the validation of the mediating ideas on the basis of prior knowledge, occurs online (Potts et al., 1988; Singer, 1993; Singer & Ferreira, 1983; Singer & Halldorson, 1996; Singer et al., 1992). It is assumed that the general idea of the validation model also holds true for expository texts and that expository text understanders have to construct and validate their inference processes on the basis of prior knowledge. This is because a) understanding causal relations seems to be crucial for expository text comprehension and b) generating inferences in expository text comprehension involves prior knowledge structures to a great degree (see also Figure 3.3). But do causal bridging inferences also accompany comprehension (i.e., are generated online) when reading expository texts? The next passage addresses this question.

4.2 Causal bridging inferences in expository texts

There is general agreement that readers spontaneously generate local causal bridging inferences when reading narrative text (Singer & Halldorson, 1996; Singer et al., 1992; van den Broek, 1990). However, the empirical situation is less clear for expository texts (e.g., Wiley & Myers, 2003).

There is no uniform empirical evidence as to whether causal bridging inferences are generated online in expository texts comprehension. Noordman, Vonk and Kempff (1992) were among the first who studied causal bridging inferences in expository texts. They found that causal bridging inferences are not generated online (Experiments 1, 2, 3), unless the readers' specific reading goals support causal bridging inference generation. The authors found that if the readers' reading purpose was a) to answer a certain question (Experiment 4) and b) to understand the text and to judge each sentence as to whether it was consistent with the meaning of the previous parts of the text (Experiment 5), the pattern of results demonstrated online causal bridging inference generation. Hence, it can be concluded that under certain circumstances (e.g., a specific reading goal), readers generate causal bridging inferences online when reading expository texts.

This assumption was backed by additional empirical evidence from findings by Singer, Hartness and Stewart (1997). The authors shortened and modified some of the sentences used by Noordman et al. (1992) to make the applied material less complex and found that readers were able to generate causal bridging inferences as long as they had enough time for their reading (Experiment 2 and 4). They concluded that local causal bridging inferences are generated online in expository text comprehension if the reading situation the reader encounters is not too difficult and complex.

Causal bridging inferences are further subdivided into *causal antecedents* versus *causal consequences* (e.g., Millis & Graesser, 1994). A causal antecedent is when the sentence being currently read is linked causally to previously read sentences; a causal consequence is when the reader forecasts what events will causally unfold after the current event being read about (Graesser et al., 2002, see Table 3.1). Causal antecedents belong to the class of bridging inferences, because the information is linked backwards. Causal consequences, on the other hand, belong to the class of predictive inferences (see Table 3.1), because the information currently read is linked forward. In most studies about causal bridging inferences, causal antecedents were investigated. In line with this, this dissertation also refers to causal antecedents when referring to causal bridging inferences unless otherwise stated.

Millis and Graesser (1994) collected lexical decision latencies to test whether causal antecedents and causal consequences are generated online in expository text comprehension. To investigate the time-course of the inference generation, the researchers included two different SOAs (*s*timulus *o*nset *a*synchrony): 540 ms (i.e., 540 ms after the final word of the inference-triggering sentence had been read) and 1040 ms SOA. Lexical decision latencies were facilitated in the case of *causal antecedents* (causal bridging inferences) at both SOAs, but not in the case of *causal consequences* (predictive inferences). The found facilitation effect for the causal antecedents indicated that causal antecedents were generated online that primed the lexical decision in an inference-triggering context and thus shortened the lexical decision latencies in comparison to an unrelated context, in that the same word was presented, but did not evoke an inference (for a more detailed description of inference measuring methods, see Chapter 6). However, in the case of causal consequences a facilitation effect was not found, indicating that causal consequences were not generated online. The found latencies did not differ across the different SOAs indicating that causal antecedents are generated online within approximately 500 ms. Causal consequences on the other hand are not generated online because they were not even found at an SOA of 1040 ms. Therefore, their generation – if they are at all generated - exceeds 1000 ms and is thus considered to be offline (Graesser et al., 2007).

These results in the domain of expository text comprehension lead to the conclusion that causal antecedent inferences are generated online, whereas causal consequences are either generated offline or not at all. This conclusion is in line with predictions for narrative texts where causal bridging inferences (causal antecedents) are assumed to be generated online, but predictive inferences (causal consequences) are assumed to be either generated offline.

Graesser and Bertus (1998) tested the same assumption by using self-paced sentence reading times from younger (M = 21.8 years) and older adults (M = 67.3 years). They additionally measured working memory span, general world knowledge, reasoning ability, and reading frequency (i.e., how much participants read any kind of texts). Multiple regression analyses on reading times (for a discussion of how reading times are applied to measure inferences, see chapter 6) showed that causal consequence inferences were more time-consuming than causal antecedent inferences and noncausal elaborative inferences were not constructed. No differences among readers on age, working memory span, general world knowledge, reasoning ability, or reading frequency regarding the inference generation were found. These results support the findings from Millis and Graesser (1994) who found that causal antecedent inferences are computed online during the reading of expository texts; causal consequences are not.

Similarly, Noordman et al. (1992) found that readers noticed inconsistencies with regard to local causal bridging inferences, but only when they were especially encouraged to look for them (Experiment 5). Using the materials from Noordman et al. (1992) and Singer et al. (1997), Singer and Gagnon (1999), on the other hand, showed that even if readers were not especially instructed to look for inconsistencies, they were able to detect inconsistencies online. Their findings revealed longer reading times for inconsistent target sentences compared to consistent target sentences. The prolonged reading times for the inconsistent target sentences were interpreted as an indicator that the participants noticed the inconsistencies and tried to integrate them with the previous content. However, Wiley and Myers (2003) criticized that the inconsistencies may not have been noticed if the causal connective “because” wasn’t explicitly linking the causal information in the target sentences. Besides, comprehension questions asked at the end of each sentence set tapped the causal bridging inference that could be made in each passage. This may have caused the reader to adopt a strategic reading style for proceeding sentences. In addition, reading times were only measured for the premise and the conclusion simultaneously (premise and conclusion were presented in one sentence). According to Wiley and Myers, it cannot be concluded from the setting presented whether the generation of the inference depended on the premise and conclusion being presented at once.

To clarify that question, Wiley and Myers (2003) measured reading times for the premise and conclusion separately. They measured reading times for conclusions that were either consistent or inconsistent with the premises. In contrast to the Singer et al. (1997) and Noordman et al. (1992) studies, these premises contained either one or two statements. The results indicated that causal bridging inferences were only drawn if participants were given two premises instead of one. The authors concluded from these results that if the necessary information to compute the inference is easily available [either explicitly stated in the text (such as by two premises as in their experiment) or through prior knowledge], the reader computes the causal bridging inference online.

Wiley and Myers (2003) argued that, in the version with two premises, prior knowledge plays a smaller role in the online generation of causal bridging inferences, because the information needed is explicitly given. Experiments 2 and 3 analyzed the conditions under which these inferences are drawn by manipulating the accessibility of the critical information in the text. The authors tested if proximity of the relevant information (and, therefore, availability of the relevant information) affects the online generation of causal bridging inferences. Readers detected inconsistent conclusions only when the conclusion was immediately preceded by a premise. Even with just one sentence of filler information intervening between premise and conclusion, readers no longer detected inconsistent conclusions.

These findings suggest that readers are able to generate causal bridging inferences online in expository text comprehension, but only under certain circumstances. These circumstances include that readers have all the necessary information to compute the inferences either given explicitly in the text or through prior knowledge. Furthermore, general inference encouraging reading goals have been found to support an online inference generation as well as a reading situation that is not too complex and difficult. In other words, it is beneficial for an online generation of causal bridging inferences in expository text comprehension if the reader has sufficient reading time, if the text is not too difficult and if the reader possesses enough prior knowledge to generate the required causal bridging inferences.

4.3 Reasons for the imbalance in inference generation as a function of text genre

Section 4.1 may have made clear that local causal bridging inferences are reliably generated online in *narrative* text comprehension. In contrast, Section 4.2 illustrated that for an online generation of local causal bridging inferences in *expository* text comprehension certain circumstances such as sufficient prior knowledge are necessary to enable an online local causal bridging inference generation. These findings can be transferred to the broader scope of general inference generation. Online and offline inferences are usually generated more reliably in narrative than in expository text comprehension. Possible reasons for this imbalance are introduced in Table 4.1. Table 4.1 provides an overview of why online and offline inference generation may be harder in expository text comprehension than in narrative text comprehension and exemplifies differences in narrative and expository text comprehension on four dimensions.

Table 4.1: Inference-relevant differences between narrative and expository texts.

Dimension	Narrative texts	Expository texts
Typical text structures	People have general knowledge about typical structures of narrative texts in the form of a story schema. This knowledge permits the reader to associate the narrative ideas with story categories, such as theme, setting, etc. (Singer & Gagnon, 1999).	The typical expository text structures are less predictable than narrative text structures. Thus, the reader of an expository text is required to think in abstract categories which may exacerbate an inference generation (Singer & Gagnon, 1999; Wolfe, 2005).
Familiarity of presented information	Information is more familiar; prior knowledge can support inference generation.	Information is often unfamiliar; insufficient prior knowledge might impede inference generation.
Automaticity of inferencing	Prior knowledge enables comprehension to proceed in an automatic and not resource demanding way.	No or insufficient prior knowledge may force comprehension to be very resource demanding and may influence the quality and quantity of inferences.
Information processing strategies	Different information processing strategies may play a role: text content needs to be integrated with prior knowledge to a lesser extent (Wolfe & Mienko, 2007).	Different information processing strategies may play a role: text content needs to be integrated with prior knowledge structures to a greater extent (Wolfe & Mienko, 2007).

What becomes clear in Table 4.1 is that readers are more familiar with narrative texts than with expository texts. These differences in familiarity regard text structures and text content. These differences in familiarity, in turn, lead to different inference and information processing strategies. In narrative text comprehension an automatic and not resource demanding inference generation is enabled, whereas in expository text comprehension an automatic and not resource demanding inference processing is impaired by the lacking familiarity and prior knowledge. It becomes evident that prior knowledge plays a considerable role in a successful inference generation, but that in expository text comprehension often the required prior knowledge is lacking and therefore inference generation might be impaired.

While this chapter discussed empirical studies about bridging inferences in text comprehension, the following chapter describes a few empirical studies analyzing inferences in film comprehension. As this dissertation is about inference processes in film comprehension, the few empirical studies that have been conducted within the context of inferences in film comprehension will be presented in greater detail.

5 Inferences in Film Comprehension

It has been argued that comprehension is a higher level process that occurs medium-independently (Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998; also Section 2.2). Therefore, inferences should not only be generated in text comprehension, but also in film comprehension. Inference processes in film comprehension have primarily been investigated for narrative films - and then - primarily predictive inferences have been examined.

One reason for the lack of empirical studies of inferences in film comprehension might lie with the difficulty of measuring inferences online in film comprehension. Methods derived from text comprehension cannot easily be applied to film comprehension, because text and film differ in some respects (see Section 2.1) that make a direct transfer of methods from text to film difficult. For instance, self-paced measures, such as reading times, can no longer be measured in film comprehension, because regular film does not allow the adaptation of individual processing strategies to cognitive processes, such as slowing down while reading when experiencing comprehension difficulties.

The question as to the right method for measuring inferences in film comprehension is raised again in Chapter 6. In the present chapter, the three existing empirical studies that investigate whether predictive inferences are generated online in narrative film comprehension are presented. As these three studies are the only empirical studies of online inference processes in film comprehension, they are presented in greater detail.

5.1 Empirical studies

Magliano and colleagues (Magliano et al., 1996) as well as Unsöld and Nieding (in press) analyzed whether predictive inferences are generated online in narrative film comprehension. It has been stated in Chapter 3 that predictive inferences belong to the class of elaborative inferences that are assumed to be generated offline in *text* comprehension (Graesser, Haberlandt, & Koizumi, 1987; Graesser et al., 1994, 1997; Potts et al., 1988). Inferences that are generated offline are usually generated strategically or when specifically asked for the respective information and demand significant cognitive resources and processing time. However, it might be that the situation is different in narrative film comprehension as, e.g., *filmic devices* (see Bordwell, 1985; Bordwell & Thompson, 1993; Magliano et al., 1996) that are not relevant in text comprehension, but guide the film viewer's attentional focus might influence the film comprehension process. Magliano et al. (1996) investigated whether film comprehenders generate predictive inferences online when considering film scenes such as depicted in Figure 5.1.

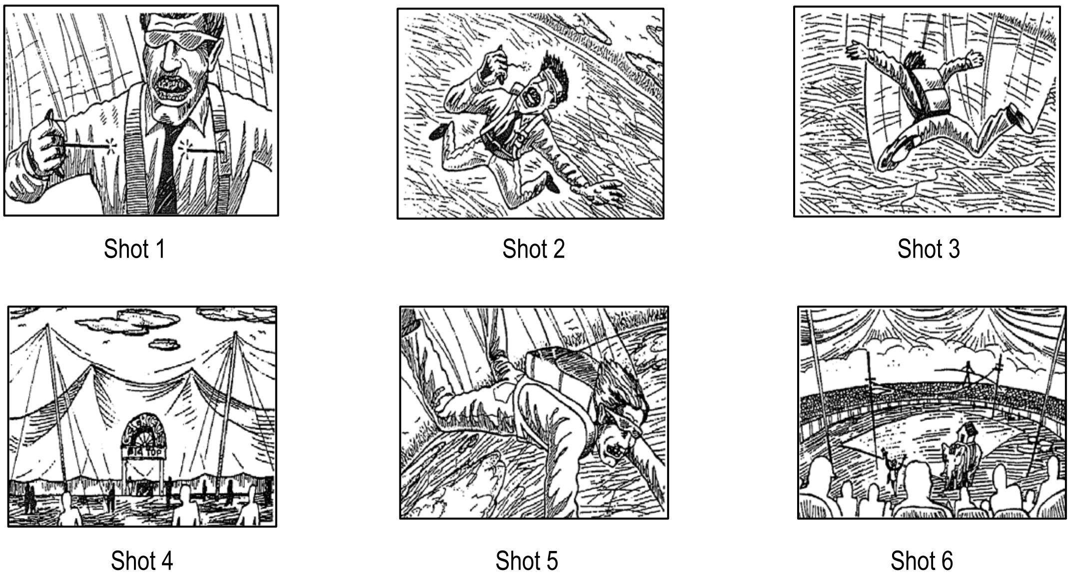


Fig. 5.1: Shots from the James Bond film applied by Magliano et al. (1996).

In Figure 5.1 six shots taken from a James Bond film are shown. Jaws a reoccurring character in Bond films has his parachute fail to open while free falling from an airplane (shots 1, 2). In shot 3, he is shown to fall without the protection of the parachute. Subsequently, in shot 4, the outside of a circus tent is depicted. Jaws' falling and screaming is then again shown in shot 5. Finally, the inside of the circus tent with acrobats and a safety net is shown in shot 6. Even though several outcomes to that scene are possible, most viewers would predict that Jaws will land on the circus tent and be protected by the safety net in a way that nobody gets hurt which is indeed shown in shot 7 (not depicted here).

The example demonstrates the use of filmic devices and how filmic devices influence the film viewer's reception process (e.g., Bordwell, 1985; Bordwell & Thompson, 1993; Magliano et al., 1996). For instance, filmic devices include zooming the camera closer to an object in focus such as in shots 1 and 5 to guide the viewer's attention or a fast editing of the shots and including a parallel montage such as in shots 3, 4, 5, 6. Jaws and the circus tent are shown subsequently, but the viewer knows that, in fact, the events take place simultaneously.

Magliano et al. (1996) analyzed whether filmic devices foster an online generation of predictive inferences in narrative film comprehension in a set of two empirical experiments. The authors assumed that filmic devices would foster an online inference generation of predictive inferences by guiding the recipient's attentional focus. The filmic devices studied by Magliano et al. (1996) were framing, montage, mise en scène and music. The camera position and camera focus belongs to *framing*, the editing of different shots is called *montage*, and, e.g., lighting belongs to the *mise en scène*. Other filmic devices that might influence the recipient's information processing are *music* and *dialogue.*

In a first experiment, participants watched the James Bond film *Moonraker* (Gilbert & Broccoli, 1979) and were instructed to stop the film whenever a prediction ("What will happen next?") came to mind and to write it down. These written protocols were analyzed: Whenever two or more viewers generated the same predictive inference as indicated by the written protocols, the filmic devices that supported the predictions were determined. A further analysis of predictions and the supporting filmic devices revealed that montage was the most frequent source of support, followed by framing, mise en scène, music and dialogue. Furthermore, it was found that those predictions that were supported by filmic devices were more often congruently (more than two viewers) generated than predictions that were not supported by filmic devices. Hence, it can be concluded from Experiment 1 that filmic devices foster an online generation of predictive inferences during narrative film comprehension.

To investigate whether viewers generate the same predictive inferences as in Experiment 1, but when not specifically instructed to do so, a second experiment was conducted by Magliano et al. (1996). The authors selected positions in the film where predictive inferences were congruently generated in Experiment 1 and divided them into four categories: predictions that were supported by several sources of information (i.e., mise en scène, montage, framing, music and/or dialogue), by a single source of information, or by no source of information. Control positions were also selected. Participants watched the same Bond film as in Experiment 1. The film was stopped at the selected positions and participants were instructed to think aloud at these positions. The authors hypothesized that film viewers generate predictive inferences online even if not especially instructed to do so and that the generation of predictive inferences is supported by filmic devices.

Empirical data supported this assumption. The likelihood that viewers in Experiment 2 generate predictions, even if not especially instructed to do so, increased as a function of the amount of informational support, i.e., filmic devices. The mean proportion of participants who generated a prediction at positions with multiple filmic devices was greater than those with one filmic device, no filmic device and control positions.

The authors concluded from this set of experiments that predictive inferences are generated online in narrative film comprehension even if not instructed to do so (Experiment 2) and that filmic devices support the online generation of predictive inferences (Experiment 1 and 2).

The finding that filmic devices have an impact on the generation of predictive inferences is a crucial finding for film comprehension research; however, it cannot be concluded from the experiments presented by Magliano et al. (1996) that predictive inferences are generated online in narrative film comprehension. It cannot be concluded from the applied methods, i.e., stopping the film when prediction comes into mind and writing it down (Experiment 1) and thinking aloud at predefined predictive inference positions (Experiment 2) that the predictive inferences were generated online. The term online was defined earlier as a spontaneous, immediate, automatic processing that occurs during the reception process within a second or less (see Section 3.2). The fact that the film was stopped either by the participants (Experiment 1) or by the experimenters (Experiment 2) interrupted the reception process in an obtrusive way and provided participants with more time to think about the respective predictive inference positions that they would have had under natural viewing conditions such as no interruption of the film. However, the film needed to be interrupted for a decent amount of time to enable thinking aloud or writing the protocols. Moreover, the completion of the written or think-aloud protocols probably lasted longer than one second after the inference triggering unit and the revealed predictive inferences can thus not be considered online. The fact that think-aloud protocols are not considered valid online methods is discussed again in Section 6.1.

It could be argued, however, that at least for Experiment 1 participants must have generated the predictive inference before they stopped the film for the written protocol. Therefore, the argument that the inference was generated offline at the time writing the written protocol seems too restrictive for Experiment 1. However, it is questionable that the content of the written protocols corresponds solely to online inference processes as the written protocols were formulated rather off- than online. Hence, the predictive inferences expressed in the written protocols do not exclusively reflect online inferences processes.

For these reasons, it cannot be unequivocally concluded from the reported findings that predictive inferences are generated online in narrative film comprehension.

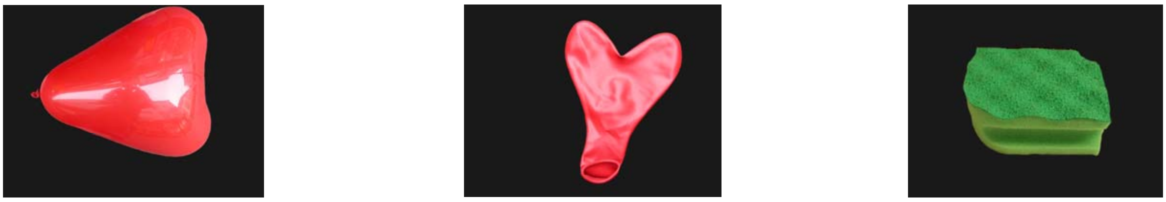
Unsöld and Nieding (in press) applied a more unobtrusive method to measure predictive inferences online that did not interrupt the film as obtrusively and as long as Magliano et al. (1996); namely the naming task (see Chapter 6). The authors adopted a developmental psychological perspective and examined if participants of different ages (6, 8, 13, 22 years old) are able to generate predictive inferences online while watching different narrative film clips. The film clips were interrupted at predefined positions that left room for predictions. For instance, one film clip Unsöld and Nieding used in their experiment showed that Tobi, the main character, buys a balloon that he wants to blow up (see Figure 5.2). The film clip ends with a shot that depicts Tobi blowing his cheeks (shot 2) while he holds the balloon in his hands (shot 1).

Shot 1 Shot 2

Fig. 5.2: Shots from a film clip applied by Unsöld and Nieding (in press).

Tobi's blowing his cheeks is shown as a still for one second. After this second a short auditive signal is presented and a black screen is shown for two seconds. The black screen is followed by a pictorial object.

The pictorial objects either demonstrates a) an object matching the prediction, b) an object shown in the film, but not matching the prediction, or c) an irrelevant object (see Figure 5.3). The participants have to name the depicted object as fast as possible and the naming latencies are recorded. After the participants name the object the film continues.



a) object matching the predictive inference   b) object not matching the predictive inference   c) irrelevant object

Fig. 5.3: Pictorial stimuli in the naming task applied by Unsöld and Nieding (in press).

The results revealed that only the youngest group showed shorter naming latencies for a), than for b) and c). This finding was interpreted as the fact that only the youngest group generated predictive inferences online and that older viewers did not generate predictive inferences online, because they processed the film clips more strategically and did not generate any predictive inferences online because they are not necessary for an online understanding (Fincher-Kiefer, 2001; Unsöld & Nieding, in press).

However, the findings from Unsöld and Nieding (in press) and the conclusion that the youngest children generate predictive inferences online can be criticized, because the way the authors applied the naming tasks does not meet the requirements for a valid online method. As online inferences are assumed to be generated within a second or less (Graesser et al., 2007) the naming task applied by Unsöld and Nieding with a SOA of three seconds (one second = still, two seconds = black screen) does not measure an online inference generation, but an offline inference generation. Therefore, their argumentation that the youngest group of children generated predictive inferences online fails as the for that age group found facilitation effect - that led to the conclusion of online generated predictive inferences - is based on offline generated predictive inferences. Thus, instead of finding online generated predictive inferences for the youngest group of children Unsöld and Nieding found offline generated predictive inferences. This is in line with results from narrative text comprehension that found that predictive inferences are generated offline (e.g., Graesser et al., 1987, 1994, 1997; Potts et al., 1988).

5.2 Methodological implications

These three presented empirical experiments investigated whether predictive inferences are generated online in narrative film comprehension and revealed that predictive inferences are not generated online in narrative film comprehension. Rather, predictive inferences are generated offline in narrative film comprehension. This is in line with findings in narrative text comprehension (e.g., Graesser et al., 1987, 1994, 1997; Potts et al., 1988). However, in contrast to text comprehension, filmic devices have been found to trigger the generation of predictive inferences in narrative film comprehension. But the generated predictive inferences that were triggered by filmic devices need to be considered as generated offline.

What becomes clear from these considerations is that the applied inference detection method has a crucial impact on the interpretation of the results obtained. To measure online inference processes, the method used needs to be as unobtrusive as possible in order not to cause an unnatural inference generation such as think-aloud protocols might. It should be noted, however, that a film interruption per se does not disqualify a method to be a valid online method. What disqualifies a method to be a valid online method, on the other hand, is if the film interruption is disruptive (such as an interruption for think-aloud protocols would be) and if the interruption – or the SOA as in naming tasks - exceeds app. 1000 ms as inferences that are generated within more than a second, are considered offline inferences (Graesser et al., 2007).

In text comprehension, researchers have focused on online methods to measure which inferences are generated online. The next chapter introduces different online methods derived from text comprehension and discusses their transferability for measuring online inference processes in film comprehension. Although it is assumed that inference processes follow the same basic principles in film as in text comprehension, it soon becomes clear that online inference detection methods are not as easily transferable, as text and film differ in some aspects. For instance, film viewers need to adapt their comprehension strategies to a medium-determined presentation, whereas text readers are able to adapt their comprehension strategies according to their current cognitive needs. For more differences between text and film, see Section 2.1. Due to these differences between text and film, some methods derived from text comprehension cannot be transferred to film comprehension, whereas others need only to be slightly modified to be applicable in film comprehension research. In addition to online methods, researchers have identified several offline methods to measure comprehension outcomes. These offline methods are also introduced in the next chapter. A distinction between offline measures for text and for film is not necessary for the most part, as these measures are mainly administered after the reception process.

6 Inference Detection Methods

Researchers in text comprehension have identified different on- and offline methods to assess inference processes at different levels of understanding. For instance, those inferences necessary to form the textbase are assumed to be generated, and thus determined, online; whereas inferences necessary to form the situation model are supposed to be generated offline and are thus measured offline. Remember that *offline* does not necessarily mean *after* the reception process even though it *can* mean exactly this. Inferences are considered to be generated offline when their computation takes longer than approximately one second, when they are generated only strategically, or when the comprehender is asked for that specific information either during or after the reception process. By contrast, inferences are generated online when they are generated automatically, within a second or less (for a definition of on- and offline inference processes see also Section 3.2). Online generated inferences are necessary for an immediate understanding and are thus always generated during the reception process. This distinction is important to have in mind when talking about the different inference detection methods, as they are distinguished as on- or offline methods, accordingly. It is assumed that inferences needed to form the textbase are generated online, because without them, no understanding is possible. Elaborative inferences, on the other hand, refine the mental representation of a text, thus allowing a deeper level of understanding, but are not necessary for an immediate understanding. They can occur during the reception process, but also after the reception process is completed.

This chapter presents on- and offline methods in inference research. First, in Section 6.1, online methods, mainly derived from text comprehension research, will be presented and, for each method, its main critique points will be discussed. Moreover, each set of online methods will be evaluated concerning its transferability to film comprehension. As text and film differ along several dimensions (see Section 2.1), not every method applied rather successfully in text comprehension is suitable to measure online inferences in film comprehension. For that reason and to account for pictorial information in films, think-aloud protocols and a method derived from event cognition research are also discussed concerning their suitability to measure online inference processes in film comprehension. Finally, in Section 6.2, offline methods are introduced. However, as most offline methods are applied after the reception process, they can be used for both text and film comprehension processes.

6.1 Online methods

One critical question in inference research on text comprehension is whether certain classes of inferences are generated online. Lively debates over the proper measures and experimental designs to test this question have dominated research in this field (Graesser & Bower, 1990; Keenan et al., 1990; Magliano & Graesser, 1991; McKoon & Ratcliff, 1989, 1990, 1992; Potts et al., 1988; Singer, 1988). However, no clear agreement on the best method has been found (Graesser et al., 1994). According to Graesser et al. (1994) there is no perfect measure and task, there are just compromises with certain advantages and shortcomings. An ideal dependent measure would tap those processes that occur *during* comprehension as compared to reconstructive processes that are measured after the reading process is completed such as summarization, recall etc. Additionally, an ideal dependent measure would track the time-course of cognitive processes such as response times (Graesser et al., 1994).

In reading research two categories have been distinguished for online methods: those methods that *do not interrupt* the natural reading process and those that *interrupt* the natural reading process. The basic idea underlying methods of the first category is that drawing inferences takes processing time, but facilitates comprehension of information that is based on the inference. These methods include self-paced reading times and eye tracking. The basic idea underlying methods in the second category is that online generated inferences prime the response to inference representing probes and therefore latencies to the inference representing probe are shorter in an inference context compared to an irrelevant context. For a comprehensive review of all reading methods, see Haberlandt (1994).

The present chapter presents the uninterrupted and the interrupted online inference detection methods in reading research and discusses their transferability to film comprehension research. To additionally account for pictorial information that is characteristics in films but that is neglected by text comprehension methods, think-aloud protocols are additionally discussed. This method has been mentioned in Chapter 5 because Magliano et al. (1996) applied think-aloud protocols to measure predictive inference processes in narrative film comprehension. However, it was argued in Chapter 5 that think-aloud protocols, even if applied during the reception process, cannot be considered valid online inference detection methods, because they interrupt the reception process in an unnatural way and mainly provide the understander with more time than one second to generate the inference. Therefore, to measure online inferences and still be able to include the pictorial information, a method derived from event cognition research is additionally presented, namely the Newtson paradigm.

Uninterrupted methods

Self-paced reading times

Self-paced reading time methods are based on the assumption that a reader reads a passage at a pace that corresponds to the internal comprehension process and, therefore, an analysis of the reading times will uncover the underlying comprehension processes (Just & Carpenter, 1980). Two additional hypotheses dominate the interpretation of reading times and, moreover, eye-tracking methods: the immediacy and the eye-mind hypothesis.

According to the immediacy hypothesis, the reader tries to comprehend a text unit, usually a word, as soon as possible rather than waiting until the end of the sentence. The eye-mind hypothesis assumes that the reader processes the text unit, usually a word, currently fixated by the eye. In other words, there is no delay between the fixated word and the cognitive process devoted to that word.

To measure self-paced reading times, participants read a text that is presented text-segment after text-segment (words, clauses, sentences) on a computer monitor. Self-paced means that participants press a key to indicate that they are finished reading the text-segment presented on the screen. After having pressed the key, a new text-segment is presented on the screen.

The basic idea is that inference generation during comprehension changes the processing and, therefore, the reading time of text-segments. Thus, if the participants show modified reading times at a theoretically predicted position, as compared to a base line, researchers infer that participants generated an inference at this location (Cirillo & Foss, 1980; Haviland & Clark, 1974). For example, when readers detect an inconsistency that does not fit in well with a former section in the text, then their reading times are prolonged, because readers are first surprised and then try to integrate the text segment currently being read with the former context. In other words, they establish global coherence by generating a global bridging inference. This cognitive process if reflected in longer reading times.

Reading-time methods have been criticized, because they have certain interpretational problems (Haberlandt, 1994). For instance, the eye-mind assumption may not always hold: readers may process a word before fully fixating on it (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Additionally, readers may continue to a new text segment while still thinking about the previous one.

Are self-paced reading times transferable to film?

Self-paced reading times are not transferable to non-interactive films as focused on in this dissertation, because the recipients do not have the possibility to influence the film's pace according to their current cognitive needs. To transfer this method to films, understanders would have to have the possibility to indicate when they are done processing a film-segment as analogue to a text-segment. Therefore, they would have to interact with the film (cf. interactive videos; see Schwan & Riempp, 2004) and change the natural presentation speed according to their own rate of processing.

In the present dissertation, recipients do not have the option to interact with the film. Thus, self-paced reading times are not suitable to measure inference processes during the comprehension of non-interactive films. Besides, the aforementioned interpretational problems provide another drawback regarding the application of self-paced reading methods.

Eye-tracking

In eye-tracking methods, a text is exposed on a computer screen and the participant's eye movements and fixations are recorded during reading (Calvo, 2001; Garrod, O'Brien, Morris, & Rayner, 1990; Haberlandt, 1994; O'Brien, Shank, Myers, & Rayner, 1988; Poynor & Morris, 2003; Rayner, 1997, 1998). Eye-tracking recordings reveal what the reader is looking at and for how long. The underlying idea – like in self-paced reading methods – is that eye movements reflect cognitive processing. As in self-paced reading times, the immediacy and eye-mind hypotheses, as well as the associated interpretational problems, hold also for eye-tracking methods. It is assumed that during an inference generation, the eye movement pattern changes compared to a base line when no inferences are generated. For instance, understanders fixate on inference triggering words longer or more frequently or show regressive instead of forward saccades. Typically, gaze durations (sum of all fixations) on inferential words plus first fixations on inferential words are analyzed (Garrod et al., 1990; O'Brien et al., 1988).

Critics refer mainly -- as in reading time methods -- to the interpretational problems. As there are several fixation measures, it is difficult to identify "the" eye fixation. Is it okay to sum all fixations to a given word or does that distort the actual reading process and, therefore, the interpretation of the cognitive process? Eye-tracking researchers recommend interpreting several eye-tracking measures.

Are eye-tracking studies transferable to film?

This method is not directly transferable to film either. In contrast to text, which is a static, stable medium, film is a dynamic, transient medium (e.g., Lowe, 1999, 2003, 2004; Schwan, 2007; Sturm, 1984) and film immanent features may influence eye movements and aggravate their interpretation. For instance, the parameter "pupillary dilation" may change if there is more cognitive processing (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Van Gerven, Paas, van Merriënboer, & Schmidt, 2004), but may also change if the amount of light in the film changes.

In sum, there are too many variables that need to be controlled to truly attribute the eye-tracking parameter to an inference generation. It would be difficult – at least currently – to use eye tracking as an indicator of generated inferences in film comprehension. Besides, the definition of dynamic units of analysis, such as dynamic areas of interest (Rötting, 2001), is still a problem and very time-consuming. Until these problems are resolved (for a first attempt, cf. Fischer, Papenmeier, & Huff, 2007), the use of eye-tracking methods for measuring online inferences in film comprehension is not recommended.

Self-paced reading and eye-tracking procedures have been useful to track the time-course of processing during reading, but these methods usually cannot measure *what* information is processed by the reader at a certain point of time (Lorch Jr. & van den Broek, 1997). For instance, slowing down during reading may reflect an inference generation, but does not reveal which memory structures are activated to draw this inference or what cognitive processes occur at that time and how they are related. To acquire this information, interrupted methods are used.

Interrupted methods

To acquire further evidence about exactly what the reader is processing, so-called "probe techniques" have been developed (Lorch Jr. & van den Broek, 1997). In a probe technique, reading is disrupted at a theoretically driven, a priori defined position in the text and the reader responds to a critical word, question, or sentence. The probe is supposed to reflect the information that needs to be inferred by the reader to understand the content read. The most typical probe techniques include recognition, lexical decision, and naming. Sometimes these techniques are called "priming methods" (Singer, 1994), because the generated inference is supposed to prime the presented probe word by spreading activation in the semantic network and, therefore, facilitates the participant's response to it. Sometimes, the recognition and lexical decision techniques are called "decision methods" (Haberlandt, 1994), because the reader is forced to make a speeded decision to target stimuli. The naming task is excluded from the decision methods (but not from the probe or priming methods), because the naming task does not require a decision, it requires only the naming of the probe. In decision -- as well as naming -- tasks, response latencies are the dependent variable.

In probe techniques, often, an *inference encoding score* (Millis & Graesser, 1994) is calculated, which compares the response latencies in an inference context to the response latencies in a control context (Long et al., 1992; Millis & Graesser, 1994; Potts et al., 1988). It is assumed that if the respective inference is generated, the response latencies in the inference context are shorter, because they were facilitated by the generated inference that activated the respective concepts in the semantic network.

Probe techniques have been widely used in reading and inference research, because they allow experimenters to exert control over choice, placement and timing of probes (Lorch Jr. & van den Broek, 1997). The recognition, lexical decision and naming tasks are presented in the following sections.

Recognition tasks

In recognition tasks, participants have to determine as fast and as accurately as possible whether the probe presented occurred in the text passage read. For inference detection, especially negative targets play an important role. Negative targets are words that did not occur in the text passage; therefore, the correct answer would be "no". However, negative targets are words that could be inferred by the reader. Thus, the latencies of hitting "no" would be increased, because the reader inferred the respective information and is no longer sure whether the word occurred in the text (Haberlandt, 1994; McKoon & Ratcliff, 1986). Moreover, false alarms (incorrectly hitting "yes") to negative targets may reflect that the reader generated the respective inference.

The recognition paradigm has been heavily criticized (Haberlandt, 1994). Critics mainly address the selection of the inference representing probes and control words. For example, it is assumed that, in several empirical studies that used the recognition paradigm (e.g., Dell, McKoon & Ratcliff, 1983; McKoon & Ratcliff, 1986, 1989), the probe words were more compatible with the text than the control words were and, therefore, latencies to positive probe words were facilitated, even if the required inferences were not generated (e.g., Forster, 1981; McKoon & Ratcliff, 1989). Haberlandt (1994) calls this a compatibility check, because the reader checks whether the probe is compatible with the context or not. Several researchers claim that latencies obtained in the recognition paradigm are thus ambiguous (O' Brien et al., 1986; Potts et al., 1988; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982).

Lexical decision tasks

In lexical decision tasks, participants have to decide whether or not a presented test string is a word of the respective language or a nonword, i.e., not an existing word in the respective language (Keenan et al., 1990; Long et al., 1992; McKoon, Ratcliff & Ward, 1994; Potts et al., 1988; Seidenberg et al., 1984). Half of the presented words are words, half of the presented words are nonwords; participants don't know in which sequence the words are presented. The idea is that decision latencies concerning a word that reflects the drawn inference are shorter than latencies to control words.

The lexical decision task has been criticized for its decision component (Forster, 1981; Keenan et al., 1990; Potts et al., 1988; Seidenberg et al., 1984; Swinney, 1979). The main concern is that lexical decision latencies reflect the time it takes to assess some sort of lexical memory located in long-term memory, as well as the time it takes to make a decision. Therefore, the latencies are comparatively long. As the reader is required to check long-term memory structures to decide whether the probe is a word or a nonword, decision latencies might be facilitated, even if no inference was generated, because the broader context might have caused the facilitation effect. Keenan et al. (1990) call this effect the context-checking effect. To exclude the possibility that a context-checking effect biases the results obtained, the authors recommend using naming tasks to measure online inferences, because they neither require nor encourage context checking.

Naming tasks

In naming tasks, participants have to name a word that is either presented in an inference or in a control context. The word is supposed to represent the generated inference. The underlying idea is that the online generated inference primes the response to the naming word, because the respective concept is activated in the semantic network. Therefore, naming latencies should be facilitated by the generated inference and are thus shorter in the inference context than in a control context (Long et al., 1992; Murray & Burke, 2003; Potts et al., 1988; Unsöld & Nieding, in press).

Critics have brought up the same reproaches for naming tasks as for lexical decision tasks, namely that the found facilitation effects rely not on inference generation but on context effects (cf. Keenan et al., 19900. However, naming tasks might be less susceptible to context checking than lexical decision tasks, because in naming tasks, no decision is required and the facilitation effect obtained is mainly based on the fact that the respective inference is generated in working memory and thus the respective concept is activated in working memory, too. Therefore, no long-term memory structures are necessarily required for the naming task. The effect of context checking for lexical decision and naming tasks is discussed more detailed in the following section.

Inference generation or context checking?

It has been criticized that probe techniques are influenced not only by the generated inference, but also by the context (Forster, 1981; Haberlandt, 1994; Keenan et al., 1990; McKoon & Ratcliff, 1989; Potts et al., 1988).

However, this criticism applies mainly to lexical decision tasks but not to naming tasks. Lexical decision tasks require the participants to check whether or not a presented test string is a word. To make this decision, participants search long-term memory structures that might have been activated by the broader context of the content presented. Naming tasks, on the other hand, do not require the involvement of long-term memory structures because the generated inference is generated in working memory and thus the concept of the probe that represents the inference is also activated in working memory. Here, the broader context should have a smaller impact on activation of concepts in working memory as in long-term memory as working memory resources are limited and only a small amount of information can be held and activated in working memory at the same time.

The assumption that naming tasks are less influenced by context checking than lexical decision tasks because for lexical decision tasks long-term memory structures that are more susceptible for context effects than working memory structures play a greater role is also reflected in shorter latencies for the naming tasks: approx. 400-600 ms in naming tasks versus approx. 900-1000 ms in lexical decision tasks. It is plausible to assume that these short naming latencies and the corresponding facilitation effects that depend mainly on the activation of concepts in working memory may not have been caused solely by the activation of the broader context but by the generation of an inference in working memory, whereas the longer lexical decision latencies and corresponding facilitation effects may have been influenced by long-term memory structures that were activated by the broader context.

To investigate which method is more indicative for online generated inferences, Potts et al. (1988) compared lexical decision tasks and naming tasks as online measures to detect elaborative inferences in narrative texts. As elaborative inferences are assumed to be generated offline (e.g., Graesser et al., 1994, 1997, 2002, 2007), the applied method should not indicate a facilitation effect. However, the lexical decision task did indicate a significant facilitation effect, but the naming task did not. Therefore, the authors concluded that the via lexical decision task found facilitation effect cannot be based on online generated elaborative inferences, because they are generated offline. Hence, the via lexical decision task found facilitation effect must be based on other mechanisms, e.g., context-checking. In line with this reasoning is that the naming task did not indicate a facilitation effect and did thus not indicate an online generation of predictive inferences.

The authors concluded that the naming task is more indicative of online inference processes than the lexical decision task and that the facilitated response latencies found by means of the lexical decision task were caused by context checking. For these reasons, Keenan et al. (1990) and Potts et al. (1988) recommend using naming tasks that do not require nor encourage context checking.

Are probe techniques transferable to film?

Yes, probe techniques are transferable to film. Probe techniques allow the experimenter to exert control over the choice, placement, and timing of probes (Lorch Jr. & van den Broek, 1997). The general idea when applying probe techniques to film is that the film's audio trace is transcribed and that the probe techniques that were derived from text comprehension research are applied to this text transcript. The text transcript needs to be analyzed carefully to identify theory-based inference positions. Therefore, probe techniques are methods to validate inferences empirically, but not to empirically detect inference positions that had not been determined a priori. After the determination of inference positions, adequate probes need to be developed that, on the one hand, reflect the respective inferences and, on the other hand, serve as adequate and comparable control probes. Then, inference and control probes can be inserted into the film and response latencies can be compared in an inference context compared to a control context.

The fact that probe techniques interrupt the film and, therefore, might influence the natural reception process has been discussed by Haberlandt (1994). However, it is assumed that the interruption, especially for the naming task where no decision is required and where naming latencies are so short, is so unobtrusive that its influence on the reception process is somewhat insignificant (Keenan et al., 1990; Potts et al., 1988). However, to measure online inferences validly via a probe technique, that is – for the aforementioned reasons - most favourably the naming task, it is important to keep in mind that the overall SOA may not exceed one second as online generated inferences are generated within a second or less. This important constraint was not accounted for by Unsöld and Nieding (in press) and thus, their found facilitation effects cannot be considered as a valid indicator for online generated inferences (see Chapter 5).

Recapitulatory evaluation of online text comprehension methods

The previous section discussed methods derived from text comprehension research to measure online generated inferences in film comprehension. It can be concluded that not all online inference detection methods in text comprehension research simply transform into online inference detection methods in film comprehension. Self-paced reading times are not applicable in film comprehension, because the dynamic medium-determined presentation speed of film does not allow the film viewer to adapt cognitive processes to his current cognitive needs. Eye-tracking measures seem to be of limited use for film comprehension, because too many confounding variables might interfere. What remains are probe techniques, which seem to be extendable to measuring inferences in film comprehension. In particular, it seems as if naming tasks are the most appropriate method derived from text comprehension research to detect online generated inferences in film comprehension. But, when measuring inferences in film comprehension, the visual pictorial information may additionally to the verbal information trigger inference processes as shown by Magliano et al. (1996) for filmic devices. Naming tasks that are based on possible inference positions identified in the film's text transcript (i.e., verbal information) might neglect the visual pictorial information. Therefore, two methods that allow for an inference generation based on visual pictorial information are presented in the following section, namely think-aloud protocols and the Newtson paradigm.

Think-aloud protocols

Magliano et al. (1996) used *think-aloud protocols* (see also Ericsson & Simon, 1980, 1993; Schellings, Aarnoutse, & van Leeuwe, 2006; Trabasso & Magliano, 1996) to detect predictive inferences in film comprehension. In think-aloud protocols, the understander is instructed to continuously talk about what is being processed in his working memory while executing a certain task such as reading. Think-aloud protocols are used in reading research (Schellings et al., 2006), but also in other domains, such as learning from animations (Fischer, Lowe, & Schwan, in press).

Opponents of think-aloud protocols claim that not all thoughts can be verbalized and that there are interindividual differences in the ability to verbalize thoughts (e.g., Nisbett & Wilson, 1977). Furthermore, it is unclear whether the reported thoughts would even be generated without the demand to think aloud or if they are triggered by the method (Graesser et al., 2007; Nisbett & Wilson, 1977). However, think-aloud protocols have received quite some attention in the past years.

To apply think-aloud protocols for inference detection in film comprehension, their application would need to be somewhat modified, because the film's audio trace would interfere with thinking aloud. One way of avoiding this shortcoming is to instruct the participants to stop the film when the respective inference comes to mind and then to think aloud as applied by Magliano et al. (1996) for predictive inferences (see also Chapter 5). However, online inferences such as bridging inferences are - in contrast to predictive inferences - supposed to be generated automatically (Graesser et al., 2007) and their generation is thus not necessarily noticed consciously by the understander. In other words, it might be difficult for the understander to comprehend the content, monitor his inference processes and explain his inference processes verbally because the understander is not necessarily aware of his own bridging inferences – in contrast to predictive inferences that are supposed to be generated offline and might thus be easier detected by the understander, because they are not generated automatically. For these reasons, think-aloud protocols cannot be applied successfully to detect online generated inferences in film comprehension.

To still account for the films' pictorial information when measuring online generated inferences, a method derived from event cognition research is discussed regarding its potential for measuring online inference processes in film comprehension, namely the Newtson paradigm.

The Newtson paradigm: key presses

In event cognition research, the work of Darren Newtson had a huge impact. He developed a general model of the perceptual and cognitive processes in the segmentation of events that were presented via films. Variations in segmentations have been shown to have a substantial effect on higher cognitive processes like memorizing, causal attribution, decision-making, or experience of time (e.g., Newtson, 1973; Schwan & Garsoffky, 2008; Schwan et al., 1998, 2000). Newtson developed a simple method: recipients viewed film clips of simple and familiar human activities like sorting a number of paper sheets or completing an assembly task. While watching the film clips, the recipients had to press a hand-held button whenever – in their opinion – a meaningful change occurred (i.e., when one meaningful action ended and a different action began). Through this procedure, each subject produced a number of key presses by which the entire observed event was segmented into a number of discrete steps. In other words, this segmentation indicated the position of subjectively defined event boundaries. The subjectively defined event boundaries reflect cognitive mental representations.

To derive conclusions as to what circumstances caused the indication of event boundaries, i.e., a substantial number of key presses; those intervals whose mean key press frequency lies above a certain threshold are analyzed further. These intervals are called breakpoints in the terminology of Newtson. Newtson used a threshold that was defined as 1.65 standard deviations above the mean key press frequency. Furthermore, to determine which circumstances did not indicate event boundaries, i.e. did not cause a substantial number of key presses, also non-breakpoints are defined. Non-breakpoints are those intervals that demonstrate a number of key presses that are 1.65 standard deviations less than the mean key press frequency.

From the segmentation of the event stream, two basic measures can be derived: first, the number of segmentations; second, their respective positions in the event stream. Newtson considers the number of segmentations to be an indicator of the amount of information that a recipient extracts from an event stream. Therefore, the segmentation can be influenced by recipient characteristics, such as prior knowledge. On the other hand, the segmentation can have an impact on the mental representation and thus the recipient's cognitive processing of the observed event.

Although the key press method has not been applied to directly measure online inference processes in film comprehension yet, it has been successfully applied to demonstrate that narrative film understanders notice temporal as well as spatial changes online (Magliano et al., 2001). Time and space have been assumed to be two crucial dimensions in the situation model (Zwaan et al., 1995). The fact that the key press technique has been successfully applied in narrative film comprehension research to online detect temporal and spatial changes that are integrated into a situation model brings one to the idea that it could also detect online generated inferences in film comprehension that are as well integrated into a situation model.

It has been stated earlier that probe techniques do not allow for pictorial information that might also trigger inference processes. In contrast, think-aloud protocols account for this type of information, but think-aloud protocols were identified not to be valid online methods in film comprehension research, because mainly the understanders had more time to think about the content when the reception process was interrupted in such an obtrusive way. The key press technique, on the other hand, allows for the pictorial information and would not interrupt the reception process, because understanders can easily watch the film and press the key when needed without interruptions. Furthermore, in contrast to probe techniques, the key press technique could be used not only to validate inferences whose positions had been determined theoretically, but also to detect generated inferences empirically by analyzing those inference positions that were identified congruently by a crucial number of participants. These empirically identified inference positions could be matched with a priori determined inference positions. If a high correlation between empirically and a priori determined inference positions is found this could eventually mean that, in future research, inference positions can be identified empirically and that the tedious and time-consuming theoretical identification of inference positions becomes obsolete.

In sum, the key press technique has three advantages: First, inferences based on the pictorial information are accounted for. Second, the film is not interrupted. Third, possible inference positions are not only identified a priori, but also empirically.

However, when measuring online inferences, the key press technique shares a drawback with the think-aloud protocols: It is assumed that online inferences such as bridging inferences are generated automatically (e.g., Graesser et al., 2007; see also Chapter 3 and 4). Hence, the understander might not be aware of their generation. The key press technique, however, requires some conscious processing of the incoming information and some monitoring of one's own inference processes, because the understander needs to monitor when he generated the inference to indicate this cognitive process via the key press.

To apply the key press technique to measure online inferences in film comprehension, it seems as if the online inference measured needs to be somewhat near to conscious awareness. Thus, local bridging inferences that have been assumed to be generated automatically (e.g., Haviland & Clark, 1974) drop out. One possible type of online generated inferences that could be closer to conscious detection seems to be global bridging inferences. It sounds plausible that understanders could be aware when they connect pieces of information across greater distances (global bridging inferences) instead of just connecting pieces of information across small distances such as two or three sentences (local bridging inferences) which they are not aware of. This seems to be particularly conceivable in expository domains when understanders are instructed to learn and thus process the content more carefully and attentively and are likely to pay more attention to their reception processes. As a consequence, it could be that the key press technique successfully detects *global* bridging inference processes in expository film comprehension.

Summary and conclusions

Section 6.1 discussed different online methods for inference detection derived in the context of text and film comprehension. It should now be clear that choosing a suitable method to measure online bridging inference processes is not a trivial task. Based on the aforementioned considerations, we can conclude that the naming task is the most suitable method derived from text comprehension to measure inferences unobtrusively and thus seems suitable to measure automatic cognitive processes. However, the naming task does not consider inferences based on pictorial information. Moreover, it does not allow detecting inference positions empirically that were not determined a priori. To measure inference processes based on pictorial information, the key press seems more appropriate. Additionally, it allows identifying inference positions empirically. However, the key press may not reliably indicate automatic inference processes such as local bridging inferences, even though the key press technique might be valid to detect global bridging inferences. In Table 6.1, naming tasks and the key press technique (i.e., the Newtson paradigm) are contrasted on five dimensions that are crucial criteria for online inference detection methods in expository film comprehension.

Table 6.1: Contrasting naming tasks and Newtson paradigm at a glance.

	Naming tasks	Newtson paradigm
Does not cause inferences that were otherwise not generated	+	+
Unobtrusive interruption	+	+
Detect automatic generated inferences	+	-
Takes into account pictorial information	-	+
Allows an empirical identification of inference positions	-	+

Table 6.1 illustrates that no single method has a positive evaluation on all five dimensions. However, the Newtson paradigm has been identified as the most appropriate method to account for the pictorial information in films and to detect online inferences that might be closer to awareness such as global bridging inferences. The naming task has been identified to be the most appropriate method to detect any kind of online generated inferences, except those that are based on pictorial information. The situation in film comprehension research seems to be as Graesser et al. (1994, p. 385) claimed for text comprehension research: “There does not appear to be a perfect measure and task; there are merely trade-offs, with each enjoying some benefits and some shortcomings”.

This dissertation studies next to online inference processes in expository film comprehension, the relations between online inference processes and offline measures of comprehension outcomes. Therefore, the next section briefly introduces different offline methods to measure offline comprehension and learning outcomes. As these measures are administered after the reception process, they can be applied to both text and film comprehension.

6.2 Offline methods

In general, for learning, both memory and comprehension processes are necessary (see Figure 6.1). Even though a clear distinction between memory and comprehension processes is hardly possible, as comprehension processes involve memory processes, researchers distinguish different methods that address either retention or comprehension measures (e.g., Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996). Retention is measured by e.g., recognition tasks, recall measures or direct retention questions that address mainly the surface level of text comprehension (e.g., Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996). To measure comprehension processes, researchers in educational psychology have -- for the most part -- relied on direct assessment methods such as comprehension test questions (e.g., Kintsch, 1998; McNamara & Kintsch, 1996; McNamara et al., 1996). Direct comprehension test questions exist in different formats such as open-ended questions, multiple choice questions or verification tasks. Indirect methods, such as key word sorting tasks, have been developed, but have received less attention in research than direct methods and are not at the focus of this dissertation. For this reason, they won’t be described further here. For more information on indirect methods, see, e.g., Kluwe (1988) or McNamara and Kintsch (1996).

Figure 6.1 illustrates different offline measures of learning outcomes and their relations to retention and comprehension processes graphically.

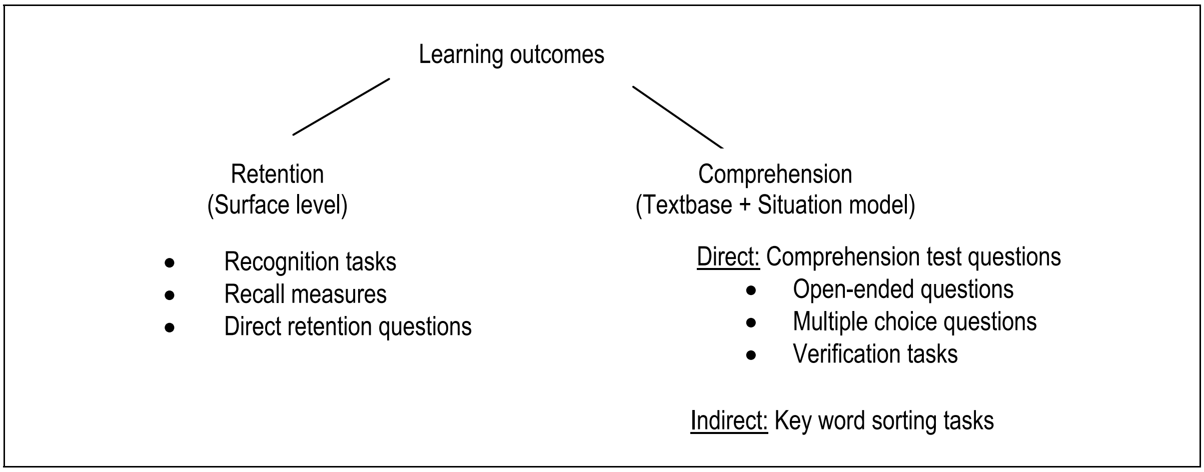


Fig. 6.1: Offline measures of learning outcomes in inference research.

The following passage briefly describes retention measures. Then, the different formats of comprehension test questions are described. As comprehension processes, and not pure memory processes, are at the focus of this dissertation the different comprehension question formats are described in more detail.

Retention measures

Recognition Tasks

Recognition tasks usually tap into memory processes that do not require inferencing. The participants judge whether a picture or a phrase was presented in the experimental material. Test probes may vary depending on the experimental material. Often, text statements, such as phrases or words, are used, but -- for the special case of films -- pictorial stimuli might be plausible.

Recognition tasks were already introduced as online inference detection methods (e.g., McKoon & Ratcliff, 1986; see also Section 6.1). However, in recognition tasks as online methods, false alarm rates and recognition latencies to negative targets are mainly analyzed. The main feature of recognition tasks as offline measures, on the other hand, is that positive targets are at the focus and that recognition latencies are no longer of interest. Moreover, the test statement is presented *after* a larger text body and *not within* a text as in online measures. By presenting the test statement after a larger text body, offline retention performance is measured. In offline recognition tasks, the answers must be retrieved from long-term memory; whereas in online recognition tasks the answers should be activated in working memory structures.

Recall measures

Recall or retention performances have often been used as an indicator for learning processes, even though they mainly reflect memory processes that do not require inferencing. Nevertheless, generated inferences may influence recall performances. Among recall tasks, free recall has been distinguished from cued recall. In free recall, participants state everything they can remember regarding a certain text passage (e.g., Linderholm & van den Broek, 2002). In cued recall, participants receive a cue that helps them remember certain information. Sometimes, the cue for recall represents a certain inference that needs to be generated (McKoon & Ratcliff, 1986).

Direct retention questions

Direct retention questions ask for specific information that has been stated explicitly in the learning material and that the learner needs to retrieve from long-term memory. To answer direct retention questions usually no inference processes are necessary. Often the type of knowledge asked for with direct retention questions is referred to as declarative knowledge.

Comprehension measures

The most often applied offline measures in inference research are comprehension test questions that address - corresponding to the theoretical distinction of different levels of mental representations – the textbase and the situation model (e.g., McNamara & Kintsch, 1996). Comprehension test questions are posed in different formats, such as open-ended questions, multiple choice questions, and verification tasks that differ either in the question format, the answer format or both.

Open-ended questions

An open-ended question is a form of question, opposite to the closed-ended one, which cannot be answered with a simple "yes/no" or a specific piece of information. Open-ended questions are answered by formulating whole sentences (e.g., Gilabert, Martinez, & Vidal-Abarca, 2005). While the construction of open-ended questions is comparatively easy, the open-ended answer analysis takes up many resources and is open to bias. A coding scheme needs to be developed and at least two coders are needed to interpret the given answers individually. To avoid subjective interpretations, the coders have to agree on a sufficient inter-rater reliability. Conflicts are resolved by discussion. Analyzing open-ended questions requires time resources, such as for developing the coding scheme, and training for the coders. Moreover, the interpretation of open-ended answers is less standardized than with multiple choice questions. However, open-ended questions might cover some aspects of knowledge that other question formats would ignore.

Multiple choice questions

One of the most applied formats in research and in the field (e.g., schools) is the multiple choice format (Haladyna, 2004). A multiple choice format asks a closed-ended question and provides a set of answers that comprises the correct answer and distractors (incorrect, but often plausible sounding, answer alternatives). The number of correct answers varies (at least one) and the number of distractors varies. Often, there are at least two distractors.

An advantage of multiple choice questions is that they are handy to analyze. Moreover, they allow a high standardization, because only one answer is correct and no interpretation of open answers is necessary. The drawback is that they require more resources to develop than open-ended questions, because incorrect, but plausible, distractors need to be created. The latter requires prior domain knowledge to design plausible sounding distractors. Furthermore, it can be criticized that multiple choice questions lead to a more somewhat restrictive knowledge elicitation process.

Verification tasks

Verification tasks are not applied as often as the two former formats, but have been shown to successfully measure inference processes in the past. A verification task refers to a statement that addresses text understanding and that needs to be identified as correct or incorrect (e.g., Durgunoğlu & Jehng, 1991). To correctly decide on the validity of the utterance, the reader has to generate inferences. The type of generated inference depends thereby on the relation between true/false statement and text and the text itself. Sometimes, the learner needs to give the rationale for his decision.

As with the multiple choice format, verification tasks allow a comparatively simple analysis. However, while it is easy to analyze whether the statement was correctly identified as either correct or incorrect, the justification for the answer requires some more analysis. It has to be noted though, that the analysis of verification tasks requires fewer resources than the open-ended question format.

Summary and conclusions

Section 6.2 discussed offline inference detection methods that can predominantly be applied for both text and film comprehension research because they are usually applied after the reception process. As for general learning outcomes both memory and comprehension processes play a role, offline methods are distinguished accordingly in retention and comprehension methods. This dissertation focuses on comprehension outcomes that are mainly assessed by direct comprehension questions. Comprehension questions are asked in different formats that each have their advantages and disadvantages. It can be concluded from the aforementioned considerations that open-ended questions and those verification tasks that demand for a rationale of the answer require more resources for the answer analysis than multiple choice questions, but that the latter demand more resources in designing answer alternatives. However, multiple choice questions allow a higher standardization than open-ended questions and verification tasks and are less susceptible to interindividually different interpretations. But they are demanding to create and might be too restrictive in covering the learners' mental representations. Depending on the goals of the researcher the direct comprehension questions seem to be equally suitable to measure comprehension outcomes.

7 Summary and Overview of Experiments

This thesis investigates whether expository films are elaborated deeply in terms of inferences being generated in expository film comprehension. Inferences have been found to be crucial for text understanding. It was argued that text comprehension theories are extendable to film comprehension processes, because higher-order comprehension processes that are medium-independent and disregard existing media differences are assumed. In line with this reasoning, it is hypothesized that inferences are also generated in expository film comprehension, unless the film's inherent dynamic pictorial information or the transient and medium-determined presentation mode impairs elaboration.

To analyze inference processes, the Construction-Integration Model, an influential text comprehension model, was introduced as the theoretical framework that stresses the generation of different levels of understanding, such as the textbase and the situation model, by drawing inferences. For each level of understanding, different kinds of inferences are necessary. For instance, the textbase requires local and global bridging inferences, whereas the situation model requires -- in addition to bridging inferences -- elaborative inferences that are generated by requiring long-term memory structures, such as prior knowledge.

To measure whether bridging inferences are generated online (i.e., immediate and automatically) or offline (i.e., delayed and strategically), different online inference detection methods were distinguished. It was concluded that there is no perfect method, because each method has advantages and disadvantages. However, Newtson's key press technique and the naming task seem to be the most promising online methods to measure bridging inferences in expository film comprehension. To investigate how online bridging inferences relate to offline comprehension outcomes, offline comprehension methods were introduced. Most often direct comprehension questions that can be asked in different formats are applied. It was concluded that the different formats seem to be equally suitable to measure comprehension outcomes and that it depends on the goals of the researcher which method is applied.

By conducting three experiments, this dissertation investigates 1) whether bridging inferences are generated online in expository film comprehension, and 2) how they relate to offline comprehension measures. It has been hypothesized that films are not elaborated deeply, because, among other things, the film's inherent pictorial information hinders a deeper elaboration. To study the effect of the pictorial information on inference and comprehension processes, an audio condition was compared to a video condition throughout all three experiments.

Experiment 1 analyzes global bridging inferences via Newtson's key press. It is assumed that, if participants process films elaboratively, then they should generate global bridging inferences that are necessary for comprehension. However, if the film's inherent pictorial information hinders elaboration, then only participants in the audio condition should generate global bridging inferences, because, with audio only, the pictorial information is missing and participants might thus better concentrate on the relevant contents. Moreover, it was studied how global bridging inferences relate to offline comprehension measures. Offline measures should benefit from an online generation of global bridging inferences, because a more coherent textbase is created by global bridging inferences.

Based on the results of Experiment 1, the naming task was implemented in Experiment 2. To allow a precise localization of the bridged information, Experiment 2 studied local causal bridging inferences instead of global bridging inferences. It is assumed that, if participants generate online local causal bridging inferences in expository film comprehension, then naming latencies are shorter in an inference-related context compared to an inference-unrelated context. If the film's inherent pictorial information hinders this type of elaboration, then only participants in the audio condition should demonstrate the predicted pattern of results.

To investigate the effect of online generated local causal bridging inferences on offline measures of comprehension, Experiment 3 tested whether participants generated a coherent mental model as indicated by better offline comprehension performances than a control group that did not receive any experimental material. Furthermore, it was tested whether inferred information is integrated as coherently into a mental representation as explicitly stated information. Experiments 1 through 3 are shown in overview in Table 7.1.

Table 7.1: Overview of Experiments 1 through 3.

	Research questions	Method
Experiment 1	Global bridging inferences: Are they generated online in expository film comprehension? How do they relate to offline comprehension performance?	Key press (online) plus comprehension test (offline)
Experiment 2	Local causal bridging inferences: Are they generated online in expository film comprehension?	Naming task (online)
Experiment 3	Local causal bridging inferences: How do they relate to offline comprehension performance?	Comprehension test (offline)

8 Experiment 1: Global Bridging Inferences: Are they Generated Online in Expository Film Comprehension and how do they Relate to Offline Comprehension Performance?

The current experiment investigates whether global bridging inferences (GBIs) are generated online in expository film comprehension. Global bridging inferences establish semantic coherence across sentences that are more than two or three sentences apart (Eysenck & Keane, 2005; Graesser et al., 2002). A globally coherent mental representation is essential for an immediate comprehension (e.g., Durgunoğlu & Jehng, 1991; Kintsch, 1988, 1998) and therefore, global bridging inferences are assumed to be generated online (e.g., Albrecht & O'Brien, 1993; Graesser et al., 1997; Myers et al., 1994; Singer et al., 1994; Tapiero & Otero, 2002). An online generation of global bridging inferences has so far never been shown for expository film comprehension and it has been questioned whether films are elaborated in terms of inferences (e.g., Salomon, 1984; Weidenmann, 1989, 2002). One reason why inferences might not be generated online in film comprehension could be that the film's inherent dynamic pictorial information hinders elaboration as claimed by some authors (e.g., Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002).

In the previous chapters it has become clear that applying the most appropriate method that validly measures online inferences is crucial in inference research. Therefore, the current experiment aims at applying an adequate method that, on the one hand, detects online generated GBIs in film comprehension, but that, on the other hand, does not evoke the generation of GBIs that would have not been generated otherwise. To measure online GBIs in expository film comprehension, the Newtson paradigm (1973), which has been previously used in research on segmentation processes in filmic events, is applied. Originally, participants had to press a key when they perceived a meaningful change in a presented filmic event. Magliano et al. (2001) could successfully apply the key press method to film understanders, who indicated online changes in time and space as predicted by the Event Indexing Model (Zwaan et al., 1995) derived from text comprehension. Therefore, it is assumed that the key press might be a valid indicator to detect GBIs in film comprehension as well. In this experiment, participants press the space bar on the computer keyboard to indicate a GBI generation while watching an expository film. The key presses the understanders generate are matched with a priori determined inference positions. These a priori determined inference positions are positions in the film transcript that have been determined by means of an objective analysis. It is assumed that at these positions a global bridging inference generation is necessary to establish a globally coherent textbase, necessary for a globally coherent mental model.

The advantage of the Newtson paradigm over traditional text-based methods is that it takes into account the pictorial information that is characteristic of films. Additionally, it does not seem to be too obtrusive, as the key press does not interrupt the reception process. However, it has been argued that bridging inferences are generated automatically in text comprehension (e.g., Graesser et al., 2007; see also

Chapter 4). Hence, the understander might not be aware of their generation. As the key press requires some monitoring of one's own inference processes, the key press technique might not be able to indicate GBI processes. Yet, it has been questioned in this dissertation whether these assumptions also hold true for global bridging inferences in expository film comprehension. For instance, it can be assumed that understanders are aware when they connect pieces of information across greater distances (global bridging inferences), especially in expository films, when they are instructed to learn from the film and thus process the content more attentively. Consequently, it is hypothesized that the key press successfully detects global bridging inference processes in expository film comprehension.

As a global coherent mental representation is essential for comprehension (e.g., Kintsch, 1998), it is assumed that GBI generation at a priori defined objective inference positions, where GBIs need to be generated to establish a coherent mental representation, results in better comprehension. To test this hypothesized relation, an offline comprehension test is also administered in the current experiment.

This experiment aims at answering the following research questions:

1) Are global bridging inferences generated online in expository film comprehension?
2) Does the presentation of dynamic pictorial information hinder online comprehension and thus deteriorate offline comprehension scores?
3) How do online generated global bridging inferences relate to offline-measured comprehension performances?

8.1 Hypotheses

Based on the theoretical assumptions outlined in the previous chapters, the following hypotheses were derived.

H1: Online generation of global bridging inference generation and its assessment

If participants generate online global bridging inferences (GBIs) and the Newtson paradigm can successfully measure these inferences, then empirically identified global bridging inference positions should match a priori determined global bridging inference positions. If a sufficient congruence between a priori determined global bridging inference positions and key presses is found, then the following conclusions can be drawn: a) participants generate global bridging inferences online in expository film comprehension and b) the Newtson paradigm can successfully measure online generated global bridging inferences in expository film comprehension. In case no sufficient congruence is found, then it can be concluded that one or both of the aforementioned conclusions are not justified.

H2: Relation between online generated GBIs and offline comprehension scores

The more global bridging inferences generated at a priori determined global bridging inference positions, the better the offline comprehension performance. If global bridging inferences are successfully generated in expository film comprehension, then the comprehender generates a globally coherent mental textbase. In most cases, a globally coherent textbase is essential for a coherent situation model -- and thus for successful text understanding. Past empirical results in text comprehension research (Beck et al., 1991; Beyer, 1990; Britton & Gülgoz, 1991; see also Section 3.3) showed that if the text coherence was increased, better recall and comprehension scores could result. It is assumed that, if participants increase the mental global coherence of the film by generating global bridging inferences online, then better offline comprehension scores result.

H3: Effect of the dynamic pictorial information

If the dynamic pictorial information in films prevents elaborated processing of the content presented (e.g., Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002), then film recipients should generate less global bridging inferences than audio tape recipients that received the same verbal information but no pictorial information. Moreover, film recipients' offline comprehension scores should be inferior to audio tape recipients' comprehension scores, as the former group would not have elaborated the film content as deeply as the latter group.

8.2 Method

To analyze the effect of the pictorial information in films on a) global bridging inference generation and b) offline comprehension scores, the presentation mode (video vs. audio) was varied as a between factor across participants. Table 8.1 illustrates the design graphically.

Table 8.1: Design.

Presentation mode	N
Video	33
Audio	33

Independent variables

The independent variable was the *presentation mode* (video vs. audio), in which the participants received the experimental material (video vs. audio tape), for analyzing the effect of the pictorial information on global bridging inference processes and offline comprehension scores. The audio condition comprised the same verbal information as the video condition, but lacked the pictorial information.

Dependent variables

Two dependent variables were implemented in this experiment: *key presses* and *comprehension scores*. The key presses were interpreted as an indicator for online generated global bridging inferences. The offline comprehension scores were an indicator of how well participants understood the experimental material.

Control variables

To control for the effects of *age*, *prior knowledge* and *interest* on the obtained results, these variables were included as control variables.

Participants

Sixty-six undergraduate students (50 female, 16 male) from the University of Tuebingen, Germany, participated in this study. Their average age was 22.74 years (SD = 3.48) and ranged from 18 to 35 years. 30.30% of the participants were psychology majors. The remaining 69.70% of the participants majored in other academic domains, such as economics, computer sciences, etc. To prevent a high level of prior knowledge, students of German, History and Empirical Cultural Studies were excluded as participants. Participants received 10€ for participation. Psychology majors also had the possibility of choosing course credit instead of the payment. German as the native language was required. All participants had normal vision or visual aids (glasses, contact lenses). The experiment lasted approximately 80 minutes. Participants were randomly assigned to the conditions.

Apparatus

The experimental procedures were controlled by a HP Compaq laptop and programmed using MediaLab, an experimental software. The film was presented at full screen on a 17" monitor.

Materials

The expository film "Post-War Germany under Allied Occupation"[5] was presented in this study. It is a DVD produced by the "FWU" ("Film für Wissenschaft und Unterricht - Film for Science and Classes"). Films by FWU are often applied in German high schools. This particular DVD was designed for students 16 years and over and adults. The film contains original news reel footage and documentary film material from the years 1945-49 and is commented by a native speaker in German. At times illustrated maps of the divided Germany and the occupational zones are presented. The film contains 10 segments that treat different topics related to post-war Germany, such as goals of the Allies, the American, British, French, and Russian occupational politics, the divided Germany and the German reunification. The length of the film is 16.35 min. Refer to Figure 8.1 to get an impression of the film.

[5] The film can be viewed at http://www.iwm-kmrc.de/intern/staff/groteloh/down/

Fig. 8.1: Stills from the experimental film "Post-War Germany under Allied Occupation".

This film was chosen, because it represents a typical domain for learning with films, namely, history (Baumann, 1995; Schwan, 2007). In history classes, films are often applied to more vividly exemplify a content that has already passed, along with the original source material (Baumann, 1995). By presenting original source material, it is assumed that students get a better impression of the past events and better understand the respective facts and circumstances, which are no longer available for them to experience in real life (Baumann, 1995; Krammer, 2006). The domain of history has been shown to be particularly suitable for the implementation of expository films, because films enable a concrete, demonstrative, and experience-driven access that is supposed to foster the students' attention and receptiveness (Baumann, 1995).

To analyze online generated global bridging inferences (GBIs), the film was divided in intervals that required GBIs (called "*objective inference intervals*") and in intervals that did not require GBIs (called "*non inference intervals*"). To determine in which intervals bridging inferences were objectively needed, i.e., necessary to establish a globally coherent textbase, the original audio trace of the film was transcribed. A propositional analysis as proposed by Kintsch (1998) and Bovair and Kieras (1985) was conducted on the film transcript by two analyzers individually. The analyzers were trained in the propositional analysis beforehand. The analyzers agreed with regard to 92% of the propositions. Conflicts were resolved by discussion. See appendix for the German film transcript and the German propositional network.

The goal of the propositional analysis was to obtain a propositional network that was then used to determine objective global bridging inference intervals in the material. As stated in Section 3.1, propositions are more related to the mental representation than regular language (e.g., Goetz, Anderson, & Schallert, 1981; Kintsch, 1998; van Dijk & Kintsch, 1983). To determine objective global bridging inference intervals that correspond to human inference processes, a propositional network seems more appropriate than the regular film transcript.

Global bridging inference intervals were defined based on positions in the propositional network that referred to earlier mentioned ideas, concepts, names, etc., and where new propositions should thus be generated to establish global coherence. By definition, GBIs connect the current information with earlier information and establish global coherence by generating new information that bridges information from at least two sources that are at least two sentences apart (Eysenck & Keane, 2005; Graesser et al., 2002; Kintsch, 1998). As films and audio tapes are both dynamic media, the idea that triggered a GBI was expressed in the course of a *time interval* that could last several seconds. The interval, in which the GBI triggering idea was conveyed, was defined as an *objective inference interval (OII)*. The 42 OIIs determined by this method differed in their duration -- from three to twelve seconds -- corresponding to the different durations needed to convey the different GBI triggering ideas. OIIs were used to analyze whether participants generated the respective GBIs. In particular, it was investigated whether a congruence between OIIs and key press frequency can be observed and whether such a congruence results in better comprehension scores.

To acquaint participants with the experimental situation and task, a *trial film* was presented. Participants were instructed to press the key while watching the trial film to indicate global bridging inferences. The trial film was a feature from a FWU film about the biological habitat “the lake” that investigated the food chain in the environment of the lake. The trial film lasted approximately five minutes.

A *comprehension test* was constructed to assess offline comprehension. The comprehension test was constructed in cooperation with an historian. The comprehension test consisted of 39 items concerning the topics that were handled in the film. All questions were answerable based on the verbal information. The items comprised 18 retention questions, 14 global bridging inference questions designed to measure whether understanders constructed a globally coherent textbase, and seven elaborative inference questions to test whether understanders constructed a coherent situation model. This resulted in 39 comprehension test questions overall (see appendix). These comprehension test questions were asked in different formats. See Table 8.2 for more specific definitions of the question types and Table 8.3 for number and formats of the different comprehension test questions.

Table 8.2: Definition of question types.

Question type	Definition
Retention	Information that is explicitly stated in the film
Global bridging inference	Inferences that require combining information across the film that is more than two sentences apart
Elaborative inference	Inferences that require applying the studied information to a novel situation or predictions

Table 8.3: Number and format of different comprehension test questions.

Question type	Number of open-ended questions	Number of verification tasks	Overall number
Retention	15	3	18
Global bridging inference	4	10	14
Elaborative inference	0	7	7

In the *open-ended question* format, participants had to state their answer in an open format. See the following example of an open-ended question:

What were the main goals of the allies that were agreed upon at the conference of Jalta in February 1945? Correct answer: The main goals were the 4 D's (i.e., De-Nazification, Demilitarization, Dismantling, Democratization for Germany).

In the *verification tasks* format, participants read a statement and had to indicate whether this statement was correct or incorrect. In addition, participants had to justify their answers by providing a few explanatory sentences. For the verification tasks, a point was only awarded if the statement was correctly marked "correct" or "incorrect" and if the justification for the answer was correct. If either one was incorrect, the point was not awarded. This rationale was chosen to prevent points being awarded for only guessing the right answer. See the following example of a verification task:

Before 1949, it could have been foreseen that Germany was going to be divided into two separate countries. Correct answer: Correct. Several activities of the allies before 1949 led to the separation of the Eastern part of Germany, for example, the rejection of the Marshall plan by the Russian occupational zone in 1947 and the assignment of Russian communists to key state positions in the Russian occupational zone.

To measure participants' *historical interest,* participants answered two self-rating questions. One question asked about general historical interest; one asked about specific historical interest in the post-war period 1945-49 in Germany. Questions were answered on a five-point Likert scale (1 = very low, 5 = very high). Overall interest was averaged across these two questions.

To measure participants' *prior knowledge,* participants answered two self-rating questions. One question asked about general historical prior knowledge; the second asked about specific historical prior knowledge regarding the post-war period 1945-49 in Germany. Questions were answered on a five-point Likert scale (1 = very low, 5 = very high). Overall prior knowledge was averaged across these two questions.

To measure the *cognitive load* participants experienced while watching the film, participants answered four questions, taken from the NASA TLX (Hart & Staveland, 1988), a standard cognitive load questionnaire. The questions measured cognitive load on three scales (effort, confidence, stress). For the effort scale, the participants answered two questions (e.g., "How much mental activity was required for film content retention?"); for the confidence scale, one question was asked ("How successful were you in connecting different scenes across the film?") and, for the stress level scale, there was also one question ("How much time pressure did you feel while watching the film?"). The effort and confidence questions were asked for two cognitive tasks, respectively: a) the key presses and b) the retention activity. Questions were answered on a five-point Likert scale (1 = very low, 5 = very high). Cognitive load for the effort scale was averaged across the two questions.

Procedure

Participants either viewed the film or listened to an audio tape with the same audio trace as the film, but without the visual trace. Participants listened to all films or audio tapes with individual headphones. For a better readability, the term "film" is used for both the film and the audio tape in the following context, because the procedure was comparable for both conditions. The experiment was conducted individually, as well as computer-controlled. The experimenter was present during the experiment. He welcomed the participant and led him to a desk where a computer and a monitor were located. Participants started the experiment by reading the instructions and answering questions regarding their own historical interest and historical prior knowledge. Afterwards, participants stated their demographical data. Then, the instructions to the trial film and the trial film were presented. Instead of the visual trace, participants in the audio condition looked at a black screen. Subsequently, the instructions for the experimental film were presented, followed by the experimental film. To measure global bridging inferences during film comprehension, the Newtson paradigm was applied. In line with Newtson's research on film segmentation (1973), participants were instructed to press the spacebar whenever they spontaneously thought of a previously viewed scene or moment in the film. After the film, participants answered cognitive load questions, followed by the comprehension test. Table 8.4 shows the procedure at a glance.

Table 8.4: Procedure.

	Video (N = 33)	Audio (N = 33)
1	Overall instruction	Overall instruction
2	2 items asking about historical interest	2 items asking about historical interest
3	2 items asking about historical prior knowledge	2 items asking about historical prior knowledge
4	Instruction & trial film	Instruction & trial audio tape
5	Experimental film	Experimental audio tape
6	4 cognitive load items	4 cognitive load items
7	39 comprehension test items	39 comprehension test items

8.3 Results

Control variables

To ensure that the control variables (interest, prior knowledge, age) were equally distributed across the presentation modes, three one-factorial ANOVAs were conducted. The analyses revealed that all control variables were distributed equally across conditions [all *F*s < 1, except for age ($F(1,64) = 2.84$, $MSE = 11.80$, $p > .10$, $\eta^2 = .04$)]. See Table 8.5 for means and standard deviations.

Table 8.5: Means and standard deviations for control variables as a function of presentation mode.

Presentation mode	Interest	Prior knowledge	Age	N
Video	3.15 (.80)	2.67 (.75)	23.46 (3.97)	33
Audio	3.15 (.67)	2.62 (.59)	22.03 (2.79)	33
Overall	3.15 (.73)	2.64 (.67)	22.74 (3.48)	66

Further data analysis was divided into four main branches: a) global bridging inferences, b) comprehension scores, c) cognitive load measures and d) the relations between global bridging inferences, comprehension scores and cognitive load measures.

Global bridging inferences

To investigate whether global bridging inferences were generated online, the pattern of online generated key presses was analyzed. Overall, participants demonstrated a mean key press frequency of 26.79 (SD = 18.17) across the expository material. In particular, film viewers demonstrated a mean key press frequency of 24.58 (SD = 15.86), whereas audio tape listeners demonstrated a mean key press frequency of 29.00 (SD = 20.22). However, a one-factorial ANOVA revealed that this difference was not significant ($F < 1$).

To evaluate whether the key presses are a valid indicator for GBIs, it was analyzed whether the empirical likelihood for key presses is higher in the objective inference intervals (OIIs) compared to non inference intervals (NIIs). NIIs are complementary to OIIs and define intervals in that no GBIs were objectively necessary. If key presses are a valid indicator for online generated GBIs, then the empirical likelihood to press a key in the OIIs should be higher than to press a key in the NIIs.

To calculate the empirical likelihood for a key press in the OIIs, two scores were calculated. One score was the *number of key presses across all OIIs* per participant. To correct for different durations of OIIs, this number was divided by the *overall duration of all OIIs* (= 4.73 minutes). For example, participant 1 pressed the key eleven times across all OIIs. Eleven divided by 4.73 equals 2.33. This "*OII hits per minute rate*" (also called *"standardized OII hit rate"*) was calculated for all 66 participants. The same procedure was additionally applied to the non inference intervals (NIIs). Per participant, the *number of key presses across all*

NIIs was divided by the *overall duration of all NIIs* (= 11.62 minutes). For example, participant 1 pressed the key 18 times across all NIIs. Eighteen divided by 11.62 equals 1.55. This "*NII hits per minute rate*" (also called *"standardized NII hit rate"*) was calculated for all 66 participants. The two standardized hit rates (OII hits per minute vs. NII hits per minute) were included as a repeated measure in a two-factorial ANOVA with presentation mode as between factor. The analysis yielded a significant difference for the key presses in favor of the OIIs ($F(1,64) = 29.08$, $MSE = .56$, $p < .001$, $\eta^2 p = .31$). No significant main effect of presentation mode ($F(1,64) = 1.21$, $MSE = 2.55$, $p > .10$, $\eta^2 p = .02$) or an interaction ($F < 1$) were revealed. See Table 8.6 for means and standard deviations.

Table 8.6: Number of OII hits per minute and NII hits per minute.

Presentation mode	Number of OII hits per minute	Number of NII hits per minute	N
Video	1.86 (1.32)	1.18 (.87)	33
Audio	2.19 (1.55)	1.46 (1.14)	33
Overall	2.02 (1.44)	1.32 (1.02)	66

In other words, participants demonstrated more key presses in OIIs per minute compared to standardized NIIs. That means that the empirical likelihood for a key press in standardized OIIs is higher than for a key press in standardized NIIs. Thus, at first sight, the obtained pattern of results is in line with the idea that participants generated GBIs at objective inference intervals (OIIs) and that participants were able to indicate their GBI generation by means of key presses. However, it has to be noted that key presses also occurred during NIIs, although actually no global bridging inferences were objectively needed during these intervals. Therefore, it seems as if participants either pressed the key even when no global bridging inferences were generated or generated global bridging inferences when not necessary according to the objective propositional analysis. These considerations question the idea that understanders generate GBIs only at predetermined objective inference intervals (OIIs) and/or indicate these GBIs via key presses.

To explore in greater detail whether the key press is a valid indicator for GBIs and to test whether a sufficient congruence between a priori determined global bridging inference intervals and key presses can be observed, further analyses were conducted. It was investigated whether some OIIs triggered systematically more key presses than other OIIs and whether specific circumstances exist that might be responsible for this potential difference. However, there was a significant correlation between number of key pressing participants per OII and the respective OII's duration ($r = .70$, $p < .01$, $N = 42$). Therefore, it was not suitable to compare the absolute number of participants that pressed the key during the respective OIIs, because the different OIIs had different lengths and the duration of the OIIs influenced the probability to press the key. Thus, the data were adjusted according to the following rationale: The number of participants that pressed the key during each OII was divided by the respective OII's length in seconds and then multiplied by the mean interval duration (6.74 sec.). The results represent the number of participants that would have pressed the key during the respective OII if this OII would have had a standardized duration of 6.74 seconds. This score is called *adjusted number of key pressing participants* and is included in the following analyses. A distribution of the adjusted number of key pressing participants per OII is shown in Figure 8.2 for the video sample and in Figure 8.3 for the audio sample.

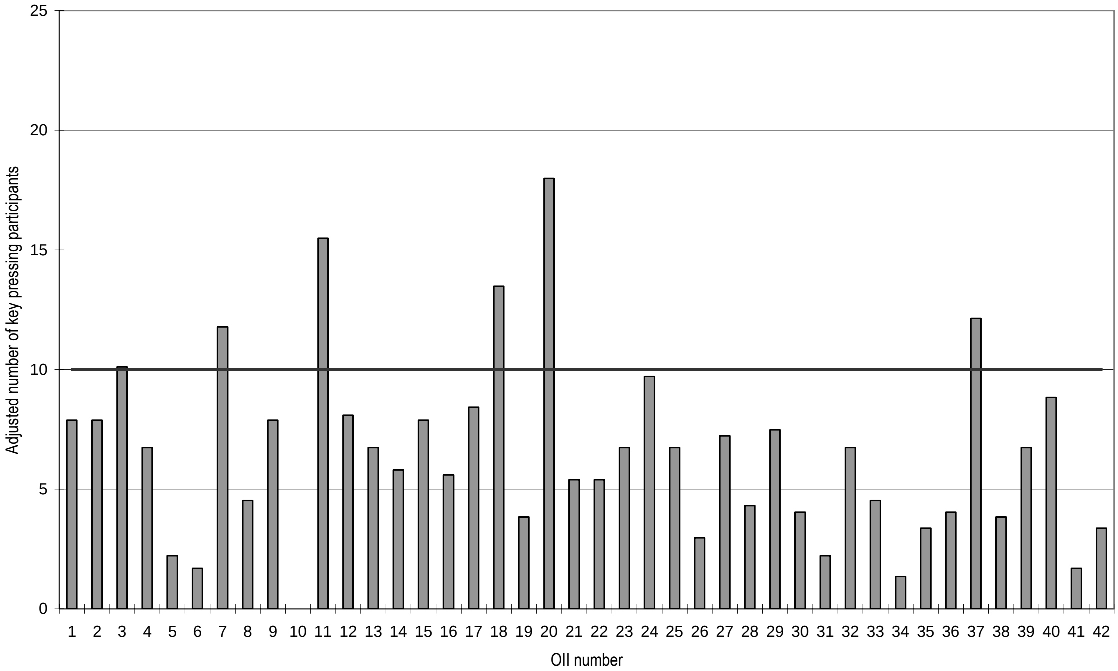


Fig. 8.2: Adjusted number of key pressing for each OII in the video sample.

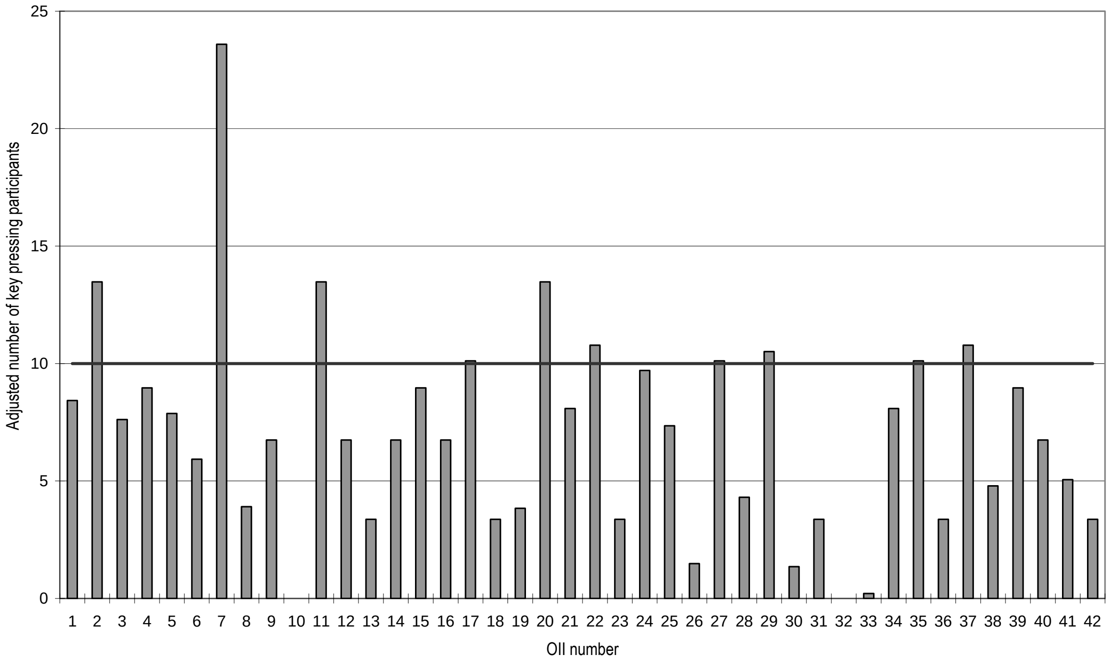


Fig. 8.3: Adjusted number of key pressing for each OII in the audio sample.

When looking at these distributions, it becomes evident that in both samples some intervals triggered substantially more participants to press the key than other intervals. These intervals shall be analyzed further regarding the contents that might have triggered understanders to generate GBIs. To determine which intervals triggered a substantial number of understanders to press the key, a threshold was defined. By looking at the empirical distribution of the adjusted number of key pressing participants, ten is defined as a suitable threshold to distinguish intervals that triggered a substantial number of participants to generate GBIs congruently from those intervals that did not. Intervals in which the adjusted number of key pressing participants was ten or higher were classified as *empirical inference intervals* (EIIs). These EIIs can be considered as having been empirically identified to trigger GBIs. The EIIs that resulted from the analysis are listed in Table 8.7. The EIIs are distinguished in EIIs in the video sample, audio sample and those EIIs that were identified across both samples (i.e., overlapping).

Table 8.7: EIIs across conditions.

	Video	Audio	Overlapping
OII number	3, 7, 11, 18, 20, 37	2, 7, 11, 17, 20, 22, 27, 29, 35, 37	7, 11, 20, 37

To answer the question which contents triggered GBIs, the empirical identified intervals (EIIs) were further analyzed.

Content analysis of empirical identified intervals (EIIs)

As illustrated in Table 8.7, the OIIs with the numbers 7, 11, 20, and 37 were identified as overlapping EIIs in both presentation modes. In other words, in these intervals, a substantial number of participants generated a global bridging inference independent from the medium that conveyed the information. A content analysis of these four overlapping EIIs provided the following insight: a substantial number of participants pressed the key when a key concept of the film, namely, the 4 D’s (i.e., Germany’s De-Nazification, Demilitarization, Dismantling, and Democratization) was mentioned. The 4 D’s were first mentioned in the film from 3.15 min. until 3.25 min. when the 4 D’s were introduced as the allies’ agreement from the conference of Jalta. The subsequent film was organized around the 4 D’s and provided details as to what measures each occupying power undertook regarding these 4 D’s. At first sight, it is possible that participants adopted a strategic reception attitude and hit the key whenever a connection to the 4 D’s was possible. Perhaps, they anticipated that the 4 D’s might be a potential test item based on similar experiences from taking tests in school. Thus, they paid particular attention to the 4 D’s. However, the hypothesis that participants established global coherence across the expository material by systematically bridging the key concept of the 4 D’s cannot be supported, because most OIIs that included a global bridging inference regarding the 4 D’s were not classified as EIIs. In other words, not a substantial number of participants hit the key during those OIIs that referred to one or all of the 4 D's. Therefore, it seems as if understanders did not systematically follow the proposed strategic reception attitude and did not generate GBIs whenever one or all of the 4 D's were mentioned.

Looking at the condition-specific EIIs reveals that two video-specific and six audio-specific EIIs resulted. In the video-specific EII with the number 3, a substantial number of participants might have generated a GBI, because three politicians (Churchill, Stalin, Roosevelt), who played a crucial role in that time period and were depicted several times throughout the film, were shown in that interval. It would seem natural that the picture of these central politicians triggered a GBI only for participants in the video sample and not in the audio sample. Similarly, the video specific EII with the number 18 showed a map that illustrated the Russian occupational zone in contrast to the Western occupational zones (US, Great Britain and France). Thus, it sounds reasonable that the video-specific EIIs were based on pictorial information that was not received by participants in the audio condition. However, the video-specific EIIs with the number 3 and 18 were preceded by two audio-specific EIIs, namely the EIIs 2 and 17. Hence, it could have been that the audio tape recipients pressed the key congruently earlier than the video recipients when GBIs were triggered and that, for instance, the additional pictorial information led to a delayed key press. Yet, the intervals 2 and 3 and 17 and 18 were not directly adjacent, but separated by several seconds in which no GBIs were objectively reasonable. Therefore, a delayed key press for video recipients cannot be assumed.

From these considerations, it could be concluded that the pictorial information triggered understanders to generate GBIs. However, for this conclusion the number of video-specific EIIs seems too small. It could be argued that OIIs based on an objective analysis of the verbal information (i.e., the film transcript) are not valid to measure GBIs based on pictorial information. Therefore, an analysis of the pictorial information in the film was conducted. Seven OIIs were determined that exclusively relied on pictorial information. However, when looking at the adjusted number of key pressing participants for these solely pictorial OIIs, only one OII lay above the threshold of ten adjusted key pressing participants. From this pattern of results, it cannot be concluded that the pictorial information triggered online global bridging inferences systematically.

From the set of OIIs that were determined according to the film transcript, i.e., the verbal information, participants in the audio condition identified four more EIIs than participants in the video condition. EIIs can be interpreted as an indicator for homogeneity in key presses, because an adjusted number of key pressing participants of at least ten is necessary per OII to be classified an EII. This could mean that participants in the audio sample showed a greater homogeneity in pressing the key, because more EIIs were identified. One reason for this greater homogeneity in key presses could have been that audio recipients received only verbal information that could trigger GBIs; whereas participants in the video condition received verbal and pictorial information that could have triggered GBIs independently and thus lead to a greater heterogeneity in GBI generation. However, the difference in number of EIIs was not significant across conditions (χ^2 (1) = .61; p > .10).

Although different analyses regarding the key press pattern and the key presses' validity to measure online generated global bridging inferences have been conducted, no systematic pattern of key presses could be identified that allowed for sound conclusions as to when or under which circumstances online global bridging inferences were generated and/or indicated by the key press.

Comprehension scores

To analyze whether participants differed in comprehension scores across presentation modes, a MANOVA was conducted first. The comprehension scores for the three subtests -- retention, global bridging inferences and elaborative inferences – were included as dependent variable. To furthermore analyze differences between these subtests, they were additionally included as a within factor in the design, yielding a 2 (presentation mode - between) x 3 (comprehension subtests - within) MANOVA. The analysis revealed that comprehension scores did not differ across presentation modes ($F(1,64) = 1.33$, $MSE = 686.56$, $p > .10$, $\eta^2 = .02$). In other words, participants across presentation modes comprehended the content equally well, no matter whether they received the information in a video or an audio format. The separate analyses for the three subtests did not reveal any effects for presentation mode either [all *F*s < 1, except for comprehension scores on global bridging inference items ($F(1,64) = 2.54$, $MSE = 266.23$, $p > .10$, $\eta^2 = .04$)]. The within comparison of the comprehension scores for the three subtests yielded that the comprehension subtest scores differed significantly ($F(2,128) = 28.71$, $MSE = 210.72$, $p < .001$, $\eta^2 = .31$). Post-hoc Bonferroni analyses revealed that retention and global bridging inference performances did not differ from one another ($p > .10$), but that both retention and global bridging inference performances were significantly superior to elaborative inference performances ($p < .05$). Refer to Figure 8.4 for a graphical illustration of the data.

The result that global bridging inference questions were equally well answered as retention questions could mean that global bridging inferences were indeed generated online, but that the key press was not valid to indicate the generated global bridging inferences.

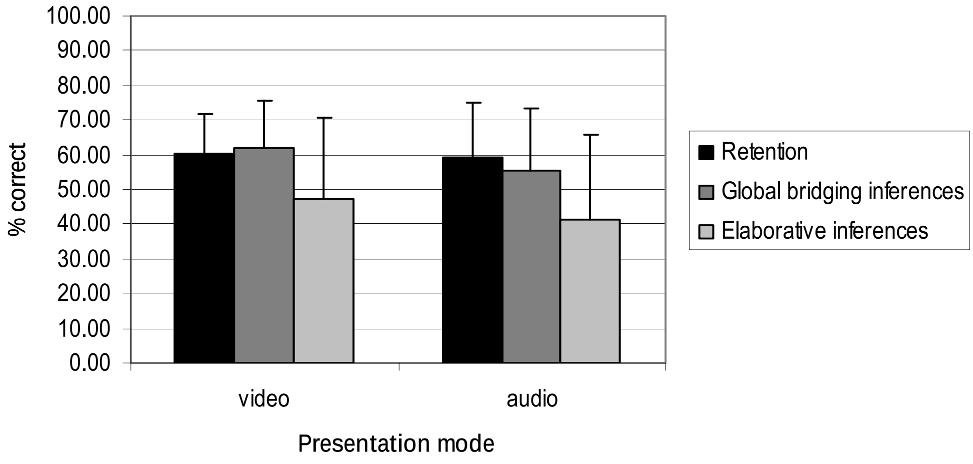


Fig. 8.4: Means and standard deviations (in % correct) of the comprehension scores as a function of presentation mode.

Cognitive load measures

To analyze the effects of presentation mode on participants' experienced cognitive load measures, five one-factorial ANOVAs were conducted. The analyses revealed no differences for all cognitive load measures across presentation modes [all *F*s < 1, except for effort memory ($F(1,64) = 2.22$, $MSE = .49$, $p > .10$, $\eta^2 = .03$)]. In other words, participants experienced about the same cognitive load independent from the presentation mode the information was presented in. See Table 8.8 for means and standard deviations.

Table 8.8: Means and (standard deviations) for cognitive load measures as a function of presentation mode.

Presentation mode	Effort key press	Effort retention	Confidence key press	Confidence retention	Stress	N
Video	2.64 (.83)	3.26 (.71)	2.79 (.86)	3.12 (.60)	1.67 (.99)	33
Audio	2.74 (.73)	3.00 (.70)	2.73 (.72)	3.06 (.70)	1.70 (.68)	33
Overall	2.69 (.78)	3.13 (.71)	2.76 (.79)	3.09 (.65)	1.68 (.84)	66

Relations between key presses, comprehension, and cognitive load

To analyze whether the number of objective inference interval hits per minute (standardized OII hit rate) is related to comprehension test performances and cognitive load measures, correlations were calculated. It sounds reasonable that a high standardized OII hit rate would lead to high comprehension test performances. Additionally, a correlation between cognitive load measures and the standardized OII hit rate might be assumed.

The analyses revealed no significant correlations between comprehension test performances and the standardized OII hit rate ($p > .10$). In other words, it did not influence the comprehension test performances whether participants hit the key when appropriate according to the propositional analysis.

In contrast, the correlation between the standardized OII hit rate and effort regarding key presses was significant ($r = .27$, $p < .05$, $N = 66$). The higher the number of OII hits per minute, the more effort regarding key presses the participants experienced. No other significant correlations were found (all *p*s > .10).

To examine the relations between cognitive load measures and comprehension performances, additional correlations were calculated. The correlation between comprehension performance on global bridging inference questions and effort regarding key presses was significant ($r = -.29$, $p < .05$, $N = 66$). The higher the performance on global bridging inference questions, the lower the invested effort regarding the key presses. Additionally, the correlation between performance on elaborative inference questions and confidence regarding key presses was significant ($r = -.37$, $p < .01$, $N = 66$). The higher the performance on elaborative inference questions, the lower the experienced confidence regarding the key presses. No other significant correlations were found (all *p*s > .10).

8.4 Discussion

Are global bridging inferences generated in expository film comprehension?

Experiment 1 investigated whether global bridging inferences are generated in expository film comprehension and how they relate to offline comprehension scores. Moreover, a new method was applied to measure global bridging inferences online in expository film comprehension.

The question of whether global bridging inferences are generated in expository film comprehension cannot be clearly answered by the current experiment. The empirical results of the current experiment did not reveal a clear pattern as to whether or under what circumstances global bridging inferences (GBIs) were generated by a substantial number of participants. Even though a few empirical inference intervals (EIIs) were identified, many objective inference intervals (OIIs) did not become EIIs, because a substantial proportion of the participants indicated no global bridging inferences congruently at these predetermined inference intervals. This could mean that only a few online global bridging inferences were generated homogeneously or that understanders were not able to detect their online global bridging inferences. The latter assumption seems to be more plausible, as *local* bridging inferences are assumed to be generated automatically without the understanders' awareness (e.g., Graesser et al., 2007). It was questioned whether this assumption also holds true for global bridging inferences in expository film comprehension, but the obtained unsystematic pattern of key presses leads to the assumption that global bridging inferences are as well generated without the understander's awareness. For that reason, the key press turned out to be not a valid indicator for online global bridging inferences, because the key press requires some monitoring of one's own cognitive processes.

In line with the assumption that global bridging inferences were generated online, but not detected by the key press, is the fact that the standardized OII hit rate did not correlate with the offline comprehension scores. This result suggests that a relation between online generated GBIs at a priori determined GBI positions, on the one hand, and offline comprehension performances, on the other hand, does not exist. However, this would be contradictory to well-established models of text comprehension and is thus not a suitable conclusion.

Another indicator for the assumption that global bridging inferences were generated online, but not detected by the key presses, are the offline comprehension scores (see Figure 8.4). If no GBIs were generated, then the performances for global bridging inference questions should have been inferior to retention questions, whose answers were directly given in the experimental material. However, no differences were found between those two scores. Therefore, it can be concluded that global bridging inferences were generated to answer the global bridging inference questions. It could be argued, however, that the global bridging inferences necessary to answer the GBI questions were generated by the time the question was asked, i.e., offline. Yet, it seems somewhat unlikely that the answers for 14 questions that asked for information that was not explicitly stated in the material, but that had to be inferred by the understander were all generated offline, because understanders would have to have integrated those information necessary to answer the GBI questions in long-term memory structures and then, in a next step, generate the GBIs necessary to answer the GBIs questions. It does not seem very probable that the latter procedure led to the same results as the pure retention questions as the latter procedure is more cognitive demanding. Therefore, it is concluded that GBIs were generated online and that the newly generated information was integrated in long-term memory structures to answer the GBI questions correctly, but that generated GBIs were not detected by the key presses.

Offline comprehension performances on elaborative inference questions were inferior to both retention and global bridging inference performances. In other words, elaborative inferences that are necessary to generate a coherent situation model were generated less than global bridging inferences that are necessary to generate a globally coherent textbase. This is in line with the assumption that the generation of a situation model is optional and depends among other things on reader characteristics such as prior knowledge or reading strategies. Furthermore, this result is in line with findings from other research fields such as problem solving psychology showing that problem solving performances decrease as soon as transfer (i.e., elaborative inferences) is involved (e.g., Reed, 1999).

Methodological considerations

As this experiment revealed, the use of the Newtson paradigm for inference detection seems restricted. Although it was successfully applied to indicate event boundaries in filmic presented events (Magliano et al., 2001; Newtson, 1973; Schwan & Garsoffky, 2008; Schwan et al., 1998, 2000), it could not be successfully used for inference detection in expository film comprehension. It was concluded that probably not only local bridging inferences, but also global bridging inferences are generated automatically without the understander's awareness. Therefore, GBIs might not have been systematically indicated by the key press that requires some monitoring of his own inference processes. However, as the results demonstrated a very heterogeneous pattern of key presses, it cannot be clearly concluded that participants were homogeneously not aware of their inference generation either. One possible explanation for the pattern of results found is that the understanders varied to a high degree in their metacognitive skills (Veenmann, 1993, 2005) and, therefore, their ability to monitor their own inference generation while still processing the experimental material varied accordingly.

Monitoring is supposed to be an important metacognitive skill that is claimed to be beneficial for learning (Veenmann, 1993, 2005; Winne & Hadwin, 1998). In this experiment, confidence ratings with regard to key presses can be interpreted as an indicator for monitoring activities. However, confidence ratings correlated negatively with performance on elaborative inference questions. Thus, participants who were more engaged in monitoring their cognitive processes had lower comprehension scores for elaborative inference questions than participants who were less metacognitively active. It could have been that monitoring one's own inference processes required many resources that in turn were not available for generating a coherent situation model, as measured by offline elaborative inference questions.

In line with this reasoning are the found negative correlations between performance on global bridging inference questions and invested effort regarding the key presses. In other words, understanders who invested more effort regarding the key press answered the GBI questions worse than understanders who invested less effort regarding the key press. It could be assumed that understanders, who invested much effort to decide when to press the key, might not have had sufficient cognitive resources left to generate the online GBIs. Hence, their GBI performances were inferior to those that did not pay much attention to the key press online and could thus generate online GBIs more easily.

This finding can as well be interpreted as another hint that GBIs were generated online, but that the key press was not a good indicator to measure online generated GBIs, because a negative correlation between online invested mental effort and GBI performance seems only reasonable if the GBIs necessary to answer GBI questions were generated online, rather than offline.

Effect of the pictorial information on online global bridging inferences and offline comprehension

It has been argued that the pictorial information in dynamic visualization such as films hinders a deeper elaboration (e.g., Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002). However, the current pattern of results did not support this assumption. The empirical inference intervals (EIIs) were identified for each condition separately. Participants in the audio condition identified four more EIIs than participants in the video condition, but this difference was not significant. In other words, film and audio tape understanders did not differ in their key press pattern, indicating that the pictorial information in films did not strongly affect inference processes and that probably most global bridging inferences were triggered by verbal information. In line with this reasoning is that the attempt to define objective inference intervals exclusively on pictorial information did not provide evidence for the assumption that pictorial information triggered substantially global bridging inferences. These findings are somewhat contradictory to findings from Magliano et al. (1996) who concluded from two experiments in the context of narrative James Bond films that filmic devices foster a predictive inference generation. Nevertheless, given the fact that the key press was not considered a valid indicator for online generated global bridging inferences, no clear conclusions regarding the pictorial effect on online inference processes can be drawn.

By contrast, clear conclusions regarding the effect of the pictorial information on offline comprehension scores can be drawn. Film and audio tape understanders did not differ in offline comprehension scores on any level of understanding. This is in favor of the claimed general higher order comprehension processes (Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998) that makes understanders comprehend contents independent from the content-conveying medium.

From these considerations, it can be concluded that the pictorial information in films did not substantially affect comprehension. Therefore, the assumption that the pictorial information in films hinders elaboration (e.g., Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002) cannot be supported by the current pattern of results. In line with this reasoning, no cognitive load differences were found across presentation modes in the current experiment, contradicting the assumption that the pictorial information causes a cognitive overload resulting in suboptimal elaboration processes.

Summary

Experiment 1 tried to apply a method derived from cognitive film psychology (i.e., the Newtson paradigm) to online global bridging inference (GBI) detection in expository film comprehension. A principle advantage of this method over text-based methods is that it takes inferences based on the pictorial information in films into account, which, however, yielded no systematic results in the current experiment. Based on the found pattern of results, it cannot be concluded that pictorial information triggered systematically online GBIs. However, this might have been due to the fact that the Newtson paradigm could not be successfully applied to indicate systematically online generated GBIs, independent of whether they are based on verbal or pictorial information. In line with the assumption that local bridging inferences are generated automatically without the understanders' awareness, it might be concluded from this experiment that global bridging inferences seem to be generated without the understanders' awareness as well and that understanders are not necessarily aware when they conduct online GBIs. To detect automatically generated bridging inferences more successfully, a method that does not require the understanders to monitor their own inference processes needs to be implemented. A method that fulfils this requirement is applied in Experiment 2. Moreover, inference type and learning domain were additionally modified from Experiment 1 to Experiment 2 to further improve the methodological approach.

Implications for Experiment 2

Based on the findings of Experiment 1, several conclusions were drawn as to which criteria a new experiment needs to meet in order to successfully investigate online inference processes in expository film comprehension. These criteria are discussed in the following paragraphs, structured along four criteria: method, material, inference type and learning domain.

Method

To detect online generated bridging inferences that are assumed to be generated automatically (e.g., Graesser et al., 2007), a method seems to be necessary that does not require the understander to be aware of his own inference generation. Looking back at text comprehension methods (see chapter 6.1), it is one of the main advantages of probe techniques that they do not require the understander to monitor his own inference processes. Moreover, they are in principle transferable to film comprehension research. Among the probe techniques, the naming task seems to be the most promising technique, as the naming task does not require nor encourage context-checking (Keenan et al., 1990).

Material

The material used in Experiment 1 was a comparatively long expository film. Especially for participants in the audio condition, listening to the audio trace and looking at a black screen for almost 17 minutes might have been fatiguing. Moreover, the long duration of the film made the analyses of the global bridging inference positions complex and tedious. Thus, the new experimental material should be shorter and less complex than the film applied in Experiment 1.

To better match the verbal and the pictorial information and to be able to interpret the obtained results unambiguously, the two sources of information should be informationally equivalent in such a way that the pictorial information supports the verbal information. As the film applied in Experiment 1 depicted original historical material, it was probably hard to find original pictorial material that supported the verbal information meaningfully. Thus, the pictorial information often did not mach the verbal information and a text-picture gap ("Text-Bild-Schere", Wember, 1976) resulted. To allow a better interpretation of the obtained results, the verbal and the pictorial information should be related to one another in such a way that the pictorial information optimally supports the verbal information and vice-versa.

Inference type

The material used in Experiment 2 should allow for clearly definable inference positions. This is due to a difficulty the current experiment encountered, namely that it was not clear what positions understanders were referring to when hitting the key. This was unclear in two respects: a) it was unclear if the key press occurred delayed regarding the generated inference and b) it was unclear what former position in the film the key press was referring to and, therefore, which information was bridged to form new connecting information. These considerations reveal another specification for the new experiment: in order to clearly identify the bridging inferences that are to be generated, the bridging inference positions need to point unequivocally to two clearly defined sources of information.

That said, the bridging inference type should be modified from Experiment 1 to Experiment 2. In line with the aforementioned reasoning, local bridging inferences should be tested instead of global bridging inferences. As local bridging inferences can be more clearly defined and two unique sources of information that need to be bridged can be exactly located, it seems easier to analyze local bridging inferences in expository film comprehension to answer the question of whether understanders elaborate expository films. It can be expected that the modifications introduced up to now ensure that the method in Experiment 2 is suitable for detecting bridging inferences online. If the new method reveals no clear pattern of results, then this can clearer than in Experiment 1 be attributed to inference processes and not to the method. However, local causal bridging inferences are only generated online under restricted circumstances in expository text comprehension, such as understanders having sufficient prior knowledge and encountering a task that fosters inference generation (Noordman et al., 1992). Therefore, these circumstances need to be accounted for in Experiment 2.

Learning domain

The aforementioned considerations lead to the conclusion that the learning domain should also be modified from Experiment 1 to Experiment 2. To investigate local causal bridging inferences, a learning domain that makes the understanding of causal connections necessary is needed in Experiment 2. Causal connections are often found in scientific learning domains, such as physics or chemistry.

Summary of implications for Experiment 2

Experiment 1 did not allow clear conclusions as expository films are elaborated online in the form of global bridging inferences. However, this lack of conclusion was mainly attributed to the applied method. Experiment 2 takes into account the methodological considerations derived from Experiment 1 and tries to implement the following methodological steps:

1) Method: less obtrusive and not requiring conscious awareness of bridging processes. Thus: naming task;
2) Material: shorter, less complex, complementary pictorial and verbal information, clearly definable, unambiguous inference positions;
3) Inference type: local causal bridging inferences;
4) Learning domain: physics or chemistry, to demonstrate causal correlations.

9 Experiment 2: Local Causal Bridging Inferences: Are they Generated Online in Expository Film Comprehension?

Experiment 1 made clear that a successful paradigm for measuring online bridging inferences in expository film comprehension needs to include an unobtrusive method that detects the generation of global bridging inferences without requiring the understander's awareness of his own inference processes. As the Newtson paradigm turned out to be unsuitable in this respect, well-established methods of inference detection in text comprehension research are reconsidered, even if that means that inference processes may only be analyzed on the basis of verbal information and not additionally on that of pictorial information. However, to establish a first successful paradigm, this seems the only path forward.

The most suitable text comprehension method to successfully measure online bridging inferences in expository film comprehension is the naming paradigm. In the naming paradigm, participants need to name an inference-representing word in an inference-related context and in an inference-unrelated context. Naming latencies are compared. It is assumed that if the inference is generated online, then the naming latencies should be facilitated by the generation of the inference by priming the respective concept in the semantic network in the inference-related context compared to the inference-unrelated context.

The advantage of this method is that it does not require any conscious awareness of the online inference processes. Moreover, it is assumed that naming does not require or encourage context-checking, which has been the major drawback of other probe techniques like lexical decision tasks.

Experiment 1 also made clear that the inference type should be modified. Instead of global bridging inferences, the current experiment analyzes local causal bridging inferences (LCBIs). It is assumed that the information necessary for LCBIs is more unambiguously definable than, e.g., information needed for global bridging inferences. The sharp definition of LCBI information enables the clear interpretation as to whether the required information was bridged online. Moreover, specific offline comprehension questions can accurately ask for the respective information. Furthermore, the experimental material and the content domain have been modified from Experiment 1 to 2. Instead of one long history film, two short films in the domain of physics are now used as the experimental material. The material is described in further detail in the material section of Section 9.3. The applied modifications should lead to a successful paradigm that is able to detect LCBIs in expository film comprehension.

Experiment 2 aims at answering the following research questions:

1. Are LCBIs generated during expository film comprehension (i.e., online)?
2. Can the naming paradigm be successfully applied to film to measure local causal bridging inference generation during expository film comprehension (i.e., online)?
3. Does the presentation of dynamic pictorial information hinder online comprehension in forms of local causal bridging inferences?

9.1 Hypotheses

Based on the theoretical assumptions outlined in the previous chapters, the following hypotheses were derived.

H1: Online generation of local causal bridging inferences and its assessment

If participants generate local causal bridging inferences (LCBIs) during expository film comprehension, then naming latencies for words that represent the respective inferences are shorter in an inference-related (IR) context than in an inference-unrelated (IU) context. When presented in the IR context, the naming word represents the required LCBI. If participants generate the required LCBI during the comprehension process, then the generated LCBI activates the respective concept in the semantic network. This activation serves as a prime and facilitates naming latencies for a word that represents the generated inference. When presented in the IU context, the same word does not correspond to a generated LCBI and naming latencies should, therefore, not be facilitated.

H2: Effect of the dynamic pictorial information

If the pictorial information in films prevents elaborated processing (DeFleur et al., 1992; Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002), then film recipients should not generate local causal bridging inferences. Therefore, the expected facilitation effect in naming latencies in different contexts (see H1) should only be observed in the audio condition, but not in the video condition. If the facilitation effect is additionally observed in the video condition, then expository films are elaborated in the form of local causal bridging inferences and the pictorial information does not impair elaboration.

9.2 Method

The variation of the two factors resulted in a 2 (context - within) x 2 (presentation mode - between) mixed factorial design. Table 9.1 illustrates the design graphically. Moreover, different naming task word sets were varied across participants, but solely for methodological reasons. Therefore, word set is not considered as an experimental between-factor. However, to control its effect statistically, it was included in the respective analyses.

Table 9.1: Design.

Presentation mode (between)	Context (within)	
	Inference-related	Inference-unrelated
Video	N = 33	N = 33
Audio	N = 30	N = 30

Independent variables

Two independent variables were included in this experiment: *context* and *presentation mode*. The context (IR vs. IU), in which the naming task was presented, was varied as a within-factor to obtain naming latencies for the same word twice by the same person: once, when the respective LCBI is required (IR context), and once, when the respective LCBI is not required (IU context). Furthermore, two presentation modes (video vs. audio) were administered in this experiment as a between-factor to analyze the effect of the pictorial information on local causal bridging inference processes. The audio condition comprised the same verbal information as the video condition, but lacked the pictorial information.

Dependent variables

Dependent variables were the *naming latencies* for the naming words presented in an IR context and in an IU context.

Control variables

To control for the effects of *age*, *prerequisite knowledge, interest* and *working memory span* on the obtained results, these variables were included as control variables.

Participants

Sixty-three undergraduate students (47 female, 16 male) from the University of Tuebingen, Germany, participated in this experiment. Their average age was 24.32 years (SD = 3.56) and ranged from 20 to 35 years. 39.68% of the participants were psychology majors. The remaining 70.32% of the participants majored in other academic domains, such as business administration, law, etc. Participants received either course credit or 11€ for participation. Psychology majors also had the possibility to choose course credit instead of the payment. German as the native language was required. All participants had normal vision or visual aids (glasses, contact lenses). The experiment lasted approximately 75 minutes. Participants were randomly assigned to the conditions.

Apparatus

The experimental procedures were controlled by a Microsoft computer and programmed using MediaLab and directRT. Film clips were presented at full screen on a 17” monitor.

Materials

As the former chapter revealed, the experimental material needed to be changed to answer the research questions of this dissertation. The new material was implemented in Experiment 2 and 3. Experiment 2 analyzes whether local causal bridging inferences (LCBIs) are generated online during the comprehension of expository films. Experiment 3 studies how online generated LCBIs relate to offline comprehension measures.

The new material should express a short duration, less complexity, an overlap between pictorial and verbal information, and several causal steps to make LCBIs necessary. Causal steps are often found in physical content domains. Several film clips were analyzed to meet the aforementioned criteria. The resulting experimental material was then taken from “Sendung mit der Maus” (“The Mouse Show”), an award-winning German educational TV program. The show is well-known in Germany and includes, in addition to funny stories, expository film clips that explain different topics. The show was originally designed for children, but adults often watch it, too. Target audiences for the “Sendung mit der Maus” are older kindergarten children and primary school students. However, the average age of its viewers is 39 years (“Sendung mit der Maus”, 2008).

Three expository film clips from “The Mouse Show” were implemented in the current experiment. One film clip (“rear view mirror”; 4.17 min.) served as a trial movie to acquaint participants with the task and situation. The film clips on the “construction of a thermos” (7.19 min) and on the “formation of lightning” (6.05 min.) served as experimental films. See Figures 9.1 and 9.2 to get an impression of the films[6].

Fig. 9.1: Stills from the expository film “Construction of a thermos”.

Fig. 9.2: Stills from the expository film “Formation of lightning”.

[6] The films can be viewed at http://www.iwm-kmrc.de/intern/staff/groteloh/down/

The new experimental material met the criteria very well. The films were short (between six and eight minutes), were less complex (they were originally designed for children and explain theoretically complex issues in a concise and plausible way), demonstrated a complementary relationship between verbal and pictorial information, and contained several causal steps.

In the expository film "the construction of a thermos", the protagonist tries to keep his tea warm over time. He starts by discovering that different materials have different heat conduction. He then detects that air has an isolating effect on heat and that it is helpful to put a lid on top of a glass container to keep the tea inside the glass warm. Next, he infers that aluminum foil reflects the heat and is, therefore, a good heat containing enclosure. At the end of the film, he discovers that a real thermos is built according to the heat preserving ideas he came up with.

The expository film "formation of lightning" starts by demonstrating the effect of generating electricity by rubbing a sweater on a balloon. Then, the effect that a balloon loaded with electricity attracts hair is demonstrated. Subsequently, it is shown that a balloon has a limited storage capacity. These principles are then transferred to the formation of lightning and exemplified by photographing real lightning with a special camera.

Experimental manipulation of film transcripts

In order to measure LCBI generation, the audio traces of the two films were manipulated in a way that made LCBI generation necessary. The manipulation was only applied to the experimental films (thermos, lightning). The trial film was presented in its original version. Nevertheless, there was a voice-over for the trial film to keep the speaker constant across films.

First, the original audio traces for each experimental film (thermos, lightning) were transliterated. Propositional analyses as proposed by Kintsch (1998) and Bovair and Kieras (1985) were conducted for the film transcripts by two analyzers individually. The analyzers were trained in the propositional analysis beforehand. The analyzers agreed with regard to 88% of the propositions for the thermos film and with regard to 90% of the propositions for the lightning film. Conflicts were resolved by discussion.

The goal of the propositional analyses was to identify possible positions in the transcript that made LCBI generation on the side of the recipient necessary to establish a coherent textbase and to understand the films' content. The propositional analysis revealed that the original transcript was coherent and all possible LCBI phrases were explicitly stated. An LCBI phrase linked two sentences causally by stating the explanatory causal bridging information explicitly. If an LCBI phrase was omitted, the understander would have to generate the LCBI to understand the film. To make LCBI generation on the side of the understander necessary, the audio trace was manipulated by excluding explicit LCBI phrases. By excluding the explicit LCBI phrases, the understander had to infer the missing information to establish a coherent textbase and thus to understand the content.

The positions where LCBIs were possible are called *LCBI positions* in the following. Overall, the “thermos” film resulted in eight LCBI positions and the “lightning” film resulted in six LCBI positions (i.e., 14 naming positions across both films). Two procedures were possible at these LCBI positions. If the LCBI phrase at the LCBI position was excluded, then an LCBI generation became necessary on the side of the understander to generate a coherent textbase and to understand the content. If the LCBI phrase was left in the transcript, the recipient did not need to generate the LCBI himself, because the connecting causal information was explicitly stated in the transcript. Both versions were included in the material.

At the LCBI positions, naming tasks were inserted to measure whether the understanders generated the respective LCBI. When an LCBI phrase was excluded at an LCBI position and, therefore, the understander was required to generate a LCBI to generate a coherent textbase, the naming word was presented in its inference-related context (IR context). When an LCBI phrase was left in the transcript at an LCBI position and, therefore, the understander did not need to generate the respective inference himself, then a naming word was presented in its inference-unrelated context (IU context).

It was assumed that if participants generated the required LCBI online, then their naming latencies for the same naming word were shorter at a position that requires the LCBI (IR context) compared to a position that did not require the LCBI (IU context). Both contexts were presented as a within factor to all participants. Thereby, the two films served as mutual control contexts. In other words, each participant saw each film once and each naming word twice. When one naming word was presented in its IU context in the thermos film, it was presented in its IR context in the lightning film and vice-versa.

The factor context was presented within participants to be able to compare the naming latencies obtained in the IR context to a measure that was generated by the same person and evoked by the same naming word. In this manner, it was ensured that interindividual and word-related differences (such as word frequency, number of syllables, etc.) could not bias the facilitation effect caused by an LCBI generation.

The examples in Table 9.2 (translated from the original German material, see appendix) exemplify the material manipulation. The asterisk (*) represents the naming task. Instead of the asterisk, the naming word was presented in the experiment. The respective naming word in its respective context is presented below the text in each cell. The original naming word was in German and is presented in parentheses next to the English equivalent. The phrases that are printed in italics and underlined are the explicit LCBI phrases that were omitted to make LCBI generation necessary in the IR context (left column) and left in for the unrelated context (right column). In Figure 9.2, the two naming words in their respective contexts were cross-mapped. This was done for explanatory reasons to demonstrate the principle of the material manipulation. In the experimental material, however, cross-mapping did not work for all cases, because, for instance, the naming word was explicitly presented in the regular text before the naming task in its unrelated context was presented.

The respective naming words were supposed to best reflect the respective LCBI that needed to be generated in order to make the film coherent. The naming words were based on the propositional analysis of the original material (see appendix). To ensure that participants had enough prior knowledge to generate the respective inferences, two prior studies were conducted (see next section).

Table 9.2: Examples from the experimental material.

Film	Context Inference-related (IR)	 Inference-unrelated (IU)
Thermos	…the tea has turned cold. Mmm. How can you prevent that? There must be a solution, Christoph thinks. With a pot full of hot water? The heat is still in there, of course. Then, he hangs something in the water. And starts a stopwatch. After 30 seconds – the finger test. The metal spoon...hot. Glass...mmm, a little colder. Hardly any heat went to the rubber. And the wood stayed cold. That's it. * * to conduct (leiten)	...the tea has turned cold. Mmm. How can you prevent that? There must be a solution, Christoph thinks. With a pot full of hot water? The heat is still in there, of course. Then he hangs something in the water. And starts a stopwatch. After 30 seconds – the finger test. The metal spoon...hot. Glass...mmm, a little colder. Hardly any heat went to the rubber. And the wood stayed cold. That's it. *Wood conducts the least heat.* * * to unload (entladen)
Lightning	…. And when is there lightning? For that, we take our balloon again. But this time, we don't rub electricity onto it, we fill it with water. That's fine for a while, but, at some point, the balloon's capacity is reached and the balloon bursts. You can also say that the balloon bursts "lightning-fast". And if the sky can't store any more electricity, there is lightning. * * to unload (entladen)	…. And when is there lightning? For that, we take our balloon again. But this time, we don't rub electricity onto it, we fill it with water. That's fine for a while, but, at some point, the balloon's capacity is reached and the balloon bursts. You can also say that the balloon bursts "lightning-fast". And if the sky can't store any more electricity, *it "unloads" electricity in the form of lightning.* * * to conduct (leiten)

The presentation order of the two films was counterbalanced. Each experimental film contained naming words in the IR context on the one hand and naming words in the IU context on the other. See the original German material in the appendix.

For methodological reasons, only at half of the a priori defined LCBI positions for one film did the naming task present naming words in its IR context. At the other half of the a priori defined LCBI positions, the naming task presented naming words from the second film in its IU context. The LCBI positions were split in half to use the two films as reciprocal control contexts and as to not overburden the participants by having too many naming tasks inserted into the material. For the thermos film, eight LCBI positions were a priori defined. For the lightning film, six LCBI positions were a priori defined. This resulted in 14 a priori defined LCBI positions across both films. The 14 corresponding naming words (eight thermos, six lightning) were split in half so that two film versions presenting two sets of naming words (called "word set" in the following) for each film were generated (thermos 1 and thermos 2, lightning 1 and lightning 2). This procedure resulted in four (eight divided by two) LCBI positions in the thermos film version 1 and three (six divided by two) LCBI positions in the lightning film version 1 and seven (four + three) LCBI positions (naming tasks) in film 1 version 1. At each LCBI position, a naming task was inserted. Depending on the film, either three (lightning) or four of these naming tasks (thermos) presented a word in its IR context, whereas the naming task at the remaining positions presented the word from the second film in its IU context. The remaining half of the LCBI positions was analyzed in version 2 of the two films (thermos film version 2, lightning film version 2).

Version 1 and 2 of each film (called "word set" in the following context) were varied across participants for methodological reasons. It was not expected that the different word sets had an effect on the naming latencies. However, the factor word set was included in the statistical analyses.

In sum, the material manipulation included the following steps:

1) Propositional analyses of the original material;
2) Identification of LCBI positions;
3) Exclusion of explicit LCBI phrases to make LCBI necessary on the side of the recipient;
4) Theoretical identification and empirical validation of the naming words;
5) Insertion of naming tasks in the experimental material in two word sets.

Prior studies

In one prior study, twelve participants watched the experimental films without the audio trace. Participants were instructed to think aloud during reception. Their think-aloud protocols were analyzed with regards to the 14 prior defined naming words. All 14 words were at least mentioned once by the participants. Ten words were mentioned by at least half of the participants, whereas six words were even mentioned by three quarters of the participants.

In a second prior study, eight participants watched the experimental films with the original audio trace. The films were interrupted at the predefined LCBI positions before the explicit LCBI phrase was stated. The participants were again asked to provide think-aloud protocols, but only at the required positions. The think-aloud protocols were analyzed with regards to whether the predefined naming words were mentioned by the participants. Again, under these more specific conditions, all 14 naming words were at least mentioned once by the participants. Thirteen words were mentioned by at least half of the participants, whereas ten words were even mentioned by three quarters of the participants. Thus, it can be concluded from both prior studies that participants had enough prior knowledge to understand the causal relations. Hence, the prior defined naming words were included in the experiment.

To measure participants' *interest in physics and chemistry,* participants answered five interest questions (e.g., "Do you like to read science books in your leisure time?"). Items were answered on a four-point Likert scale (1 = very low, 4 = very high). The scale had an even number of answer alternatives to avoid "a tendency to the mean value". Overall interest was averaged across these five items.

To measure the *cognitive load* participants experienced while watching the films, participants answered six questions after each film, all taken from the NASA TLX (Hart & Staveland, 1988), a standard cognitive load self-evaluation questionnaire. The questions measured cognitive load on three scales (effort, confidence, stress); for each scale, participants answered two questions: experienced effort (e.g., "How much mental activity did the film require? "), experienced confidence level (e.g., "How relaxed did you feel while watching the film?"), and experienced stress level (e.g., "How much time pressure did you experience while watching the film?"). Items were answered on a four-point Likert scale (1 = very low, 4 = very high). Cognitive load for each scale (effort, confidence, stress) was averaged across the respective set of two questions.

To ensure that participants tried to understand the content of the films (important for an LCBI generation in expository domains; see Section 4.2), participants answered two *comprehension test questions* regarding the film content after watching the film. Participants knew before starting the film that their comprehension scores were measured subsequent to the film. The comprehension scores were not analyzed further. This is in line with other experiments in inference research (e.g., Albrecht & O'Brien, 1993). The comprehension test consisted of two questions for each film: one question asked for the main concepts presented in the film ("Please explain the main concepts regarding the construction of a thermos?" / "Please explain the main concepts regarding the formation of lightning?"); one question asked a transfer question ("How are liquids heated faster?" / "Under which circumstances does no lightning occur?").

To prevent carry-over effects from the first to the second film, participants rated 10 more or less well-known art paintings regarding their appearance, the artistic composition, etc., on seven dimensions (e.g., "What do you like about the painting?/ How well do you think the colors were chosen? / Does the artistic composition reflect the overall theme of the painting?"). The *picture rating test* was not further analyzed.

To evaluate the obtrusiveness of the naming task, participants answered two questions as to *how disruptive* they experienced the *naming task* to be ("How disruptive did you think the words were throughout the film? / Do you think that the task asking you to name the word affected your comprehension of the film content?").

Items were answered on a four-point Likert scale (1 = very low, 4 = very high). Overall naming task disruption was averaged across these two questions.

To assess participants' prior knowledge, participants answered six multiple choice questions measuring *prerequisite knowledge.* The test did not ask about any concepts that were presented in the film, but related concepts. Therefore, the test is not called prior knowledge test, but prerequisite knowledge test in the following context. It was assumed that, if students scored high on this test, they had a high relevant domain specific knowledge to generate the inferences required in this experiment. Participants received one point per correct answer. Prerequisite knowledge was averaged across the six questions and included as % correct in the analyses. The test was administered after the participants viewed both films in order to avoid priming effects of the test questions that could possibly influence participants' naming latencies when administered beforehand. See the following for an example of the prerequisite knowledge test referring to thermos and lightning.

Thermos
In high altitude areas in Asia, the summers are short. In the spring, farmers sprinkle ashes on their snow-covered fields. Why?

a) Because the hot ashes make the snow melt.
b) Because the ashes absorb the heat coming from the sun rays, transfer the heat to the ground and the snow melts. *(correct answer)*
c) Because white snow reflects the heat coming from the sun, but the ashes prevent the heat from escaping into the air through its isolating effects and thus causes the snow to melt.

Lightning
Which answer explains how a dynamo functions?

a) The movement of the coil in a magnetic field causes electrons in the coil to move, as well. This is what generates electricity. *(correct answer)*
b) The friction of the dynamo head on the bike's tire generates an electrical charge.
c) The fast rotation of the coil in the inner dynamo generates centrifugal forces. This separates protons from electrons and generates electricity.

To assess whether *working memory span* (e.g., Engle, 2002; Daneman & Carpenter, 1980) affected LCBI generation, participants conducted a computer-presented reading span test (Hacker, Handrick, & Veres, 1998) that lasted 25-30 minutes. The test was in German and based on the "reading span task" of Daneman and Carpenter (1980). It is assumed that the reading span serves as a measure for the working memory span. As working memory resources are limited (Baddeley, 1992), only a restricted amount of information can be simultaneously processed in working memory. Hence, if the understander has a high working memory span, he can process more information simultaneously, and might thus be able to generate more inferences. In general, the reading span is a valuable predictor for text comprehension (Hacker et al., 1998).

Participants read an ascending number of computer-presented disconnected sentences for each trial, starting with two sentences and ending with five sentences. The respective number of sentences was presented as a whole before participants noted the last word of each sentence and two key words describing the sentences' content on a separate sheet of paper. Each sentence was presented for five seconds. Participants received one point when the last words of the sentences *and* the key words were correct. This is illustrated in the following example.

(1) Nobody accomplished luring him away from the oven in order to take him out of the hut.
(2) The woman was pregnant and gave birth to a healthy boy.

Last words: hut (1), boy (2)
Content: him, hut (1); woman, son (2)

Scores were added and transformed into reading span values (range 1.5 (low) – 5.5 (high)). The higher the reading span value, the higher the reading (working memory) span.

Procedure

The experiment was run computer-controlled and individually. The experimenter was present during the experiment. He welcomed the participant and led him to a desk with a computer and a monitor. Participants started the experiment by reading the instructions and answering questions regarding their interest in physics and chemistry (science). Afterwards, participants stated their demographical data. Then, the instructions to the trial film and the trial film were presented. Participants listened to all films or audio tapes via headphones. For a better readability, the term "film" is used for the film *and* the audio tape in the following context, because the procedure was the same for both conditions. Subsequently, the instructions to the experimental film were presented, followed by the first experimental film. A word naming task was inserted at predefined LCBI positions that participants were not aware of. Naming words were supposed to reflect LCBI in the IR context, and served as a control condition when presented in the IU context (see material section). The film was interrupted by showing a blank screen for 250ms, followed by an asterisk for 500ms. Then, the naming word (e.g., "to conduct") was presented. Figure 9.3 illustrates the procedure of the naming task graphically.

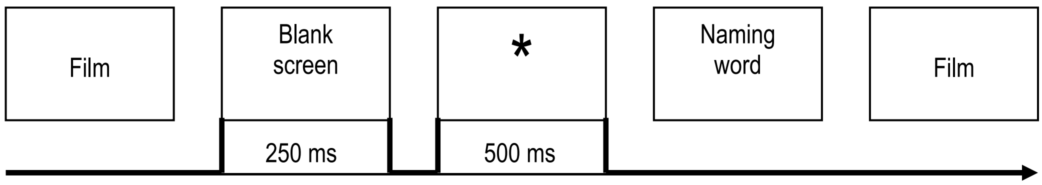


Fig. 9.3: Naming task procedure.

The participants’ task was to name the word aloud as quickly as possible. A voice-activated relay was used to determine the latencies between the presentation of the word and the participants’ response. In addition, an experimenter monitored the session and recorded any pronunciation errors. When participants named the word aloud, the word disappeared and the film restarted where it left off. It was expected that, if participants generated the respective LCBI (“wood conducts the least heat”; see Table 9.2) online, then the LCBI had primed the naming of the word “to conduct” and had facilitated, therefore, the naming latencies for the word “to conduct”. Audio tape listeners looked at a black monitor screen while listening to the audio tape. The experimenter ensured that participants focused on the screen throughout the experiment as the naming tasks appeared randomly and were not announced auditorily. Subsequent to the film, all participants answered the cognitive load items, followed by the comprehension test questions. Afterwards, participants conducted the picture rating test that lasted approx. 10 minutes. Then, participants proceeded with the second film. The procedure was identical to the first film. After answering the comprehension test questions regarding the second film, participants rated how disruptive the word naming task was overall. Subsequently, participants conducted a prerequisite knowledge test. Finally, participants executed a working memory span test. Refer to Table 9.3 for an overview of the procedure.

Table 9.3: Procedure.

	Video (N = 33)	Audio (N = 30)
1	Overall instruction	Overall instruction
2	5 items asking about interest in physics & chemistry	5 items asking about interest in physics & chemistry
3	Demographic data	Demographic data
4	Instructions + trial film	Instructions + trial film
5	Experimental film 1	Experimental film 1
6	6 cognitive load items & 2 comprehension test items regarding film 1	6 cognitive load items & 2 comprehension test items regarding film 1
7	Picture rating test	Picture rating test
8	Experimental film 2	Experimental film 2
9	6 cognitive load items & 2 comprehension test items regarding film 2	6 cognitive load items & 2 comprehension test items regarding film 2
10	2 items regarding the overall disruptive effect of the word naming task	2 items regarding the overall disruptive effect of the word naming task
11	6 items asking about prerequisite knowledge	6 items asking about prerequisite knowledge
12	Working memory span test	Working memory span test

Each participant saw one version of each film (either version 1 or version 2). If participants saw thermos version 1, they always saw lightning version 1 (thermos version 2 → lightning version 2). The word sets presented in the naming tasks belonged together for either version 1 or version 2 across both films.

The presentation order of the films and the word sets were counterbalanced. This resulted in eight conditions; see Table 9.4. Participants were assigned randomly to the conditions. The eight conditions were applied due to methodological reasons. For content purposes, only the presentation modes (video vs. audio) and context (IR vs. IU) were of interest. However, the effect of the factor word set was controlled statistically in the following analyses.

Table 9.4: Condition set-up.

Condition	Video (N = 33)	N	Condition	Audio (N = 30)	N
1	Thermos 1, Lightning 1	9	5	Thermos 1, Lightning 1	8
2	Lightning 1, Thermos 1	9	6	Lightning 1, Thermos 1	7
3	Thermos 2, Lightning 2	8	7	Thermos 2, Lightning 2	8
4	Lightning 2, Thermos 2	7	8	Lightning 2, Thermos 2	7

9.3 Results

Control variables

To ensure that the control variables (interest, prerequisite knowledge, age, working memory span) were equally distributed across the conditions, four 2 (presentation mode) x 2 (word set) ANOVAs were conducted. The analyses revealed that all control variables were equally distributed across conditions [all *F*s < 1, except working memory span across presentation modes ($F(1,59) = 2.59$, $MSE = .45$, $p > .10$, $\eta^2p = .04$) and age across word sets ($F(1,59) = 1.19$, $MSE = 12.76$, $p > .10$, $\eta^2p = .02$)], except that *working memory span (WMS)* was not distributed equally across word sets ($F(1,59) = 14.99$, $MSE = .45$, $p < .001$, $\eta^2p = .20$). Participants who received word set 2 in the naming task had a higher WMS than participants who received word set 1 in the naming task. See Table 9.5 for means and standard deviations.

However, no correlations between WMS and naming latencies' differences were found ($r = -.01$). Hence, no further analyses were conducted. Additionally, *interest* was tentatively not equally distributed across presentation modes ($F(1,59) = 3.86$, $MSE = .35$, $p < .10$, $\eta^2p = .06$). Participants in the audio condition had slightly more interest in physics and chemistry domains than participants in the video condition. However, no correlations between interest and naming latencies' differences were found ($r = .00$). Hence, no further analyses were conducted.

Table 9.5: Means and (standard deviations) for control variables as a function of presentation mode and word set.

Presentation mode	Word set	Interest	Prerequisite knowledge in % correct	WMS	Age	N
Video	1	1.81 (.59)	42.59 (19.99)	2.73 (.59)	24.83 (3.55)	18
	2	1.84 (.60)	53.33 (19.11)	3.67 (.79)	24.60 (3.92)	15
	Overall	1.82 (.58)	47.47 (20.04)	3.16 (.83)	24.73 (3.67)	33
Audio	1	2.11 (.47)	53.33 (21.08)	2.73 (.67)	24.73 (4.32)	15
	2	2.13 (.69)	52.22 (23.46)	3.11 (.65)	23.00 (2.10)	15
	Overall	2.12 (.58)	52.78 (21.92)	2.92 (.68)	23.87 (3.45)	30
Overall	1	1.95 (.55)	47.47 (20.88)	2.73 (.61)	24.79 (3.86)	33
	2	1.99 (.65)	52.78 (21.03)	3.39 (.76)	23.80 (3.20)	31
	Overall	1.97 (.60)	50.00 (20.95)	3.04 (.76)	24.32 (3.56)	63

Naming latencies

There were virtually no pronunciation errors (1 word out of 882 = 0.11%). Naming latencies greater than 2000 ms were treated as missing data. That constituted 0.68 % of the data. The naming latencies followed a Gaussian distribution. A 2 (presentation mode) x 2 (context) x 2 (word set) ANOVA was performed on the naming latencies. The analysis revealed a highly significant main effect of context on the naming latencies ($F(1,59) = 41.19$, $MSE = 1727.42$, $p < .001$, $\eta^2 p = .41$). As expected, naming latencies in the IR context were shorter than naming latencies in the IU context. This was the same across presentation modes and word sets (all Fs < 1). Figure 9.4 illustrates the naming latencies graphically.

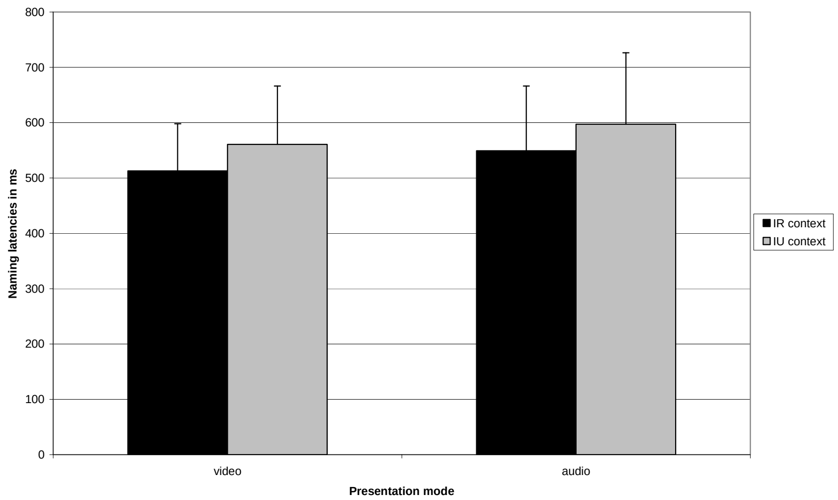


Fig. 9.4: Means and standard deviations (in ms) of the naming latencies as a function of context and presentation mode.

Comprehension scores

Comprehension scores were not analyzed, because comprehension tests were only administered to ensure an elaboration attitude by the participants by administering a task that fosters inference generation (Singer et al., 1997). This is in line with former experiments in the context of inference research that did not further analyze the offline comprehension tests scores (Albrecht & O'Brien, 1993).

Cognitive load and naming task disruption

To analyze the effects of presentation mode and word set on participants' experienced cognitive load measures and participants' experienced overall disruption of the naming task, two 2 (presentation mode) x 2 (word set) ANOVAS were conducted. Due to technical problems, the measures of the first 14 participants were not recorded for these two variables. This affected all conditions comparably. The total sample for these analyses was reduced to 49. Analyses did not reveal any effects for presentation mode nor word set on all three cognitive load measures (all Fs < 1). Refer to Table 9.6 for means and standard deviations. Analyses did not reveal any effects for presentation mode nor word set on the overall disruption of the naming task (all Fs < 1). Refer to Table 9.6 for means and standard deviations. This means that participants experienced the same cognitive load independent from the presentation mode. This is an indication that the pictorial information in films does not demand more cognitive resources than a solely verbal presentation. Overall participants across presentation modes rated the naming task as "not very disrupting", confirming the unobtrusiveness of the naming task.

Table 9.6: Means and (standard deviations) of cognitive load measures and naming task disruption for word set and presentation mode.

Presentation mode	Word set	Effort	Confidence	Stress	Naming task disruption	N
Video	1	2.66 (.32)	3.11 (.35)	1.63 (.41)	1.89 (.71)	14
	2	2.57 (.36)	3.11 (.26)	1.57 (.30)	2.00 (.63)	11
	Overall	2.62 (.33)	3.11 (.31)	1.60 (.36)	1.94 (.67)	25
Audio	1	2.66 (.67)	2.91 (.50)	1.80 (.52)	2.18 (.98)	11
	2	2.67 (.43)	3.08 (.66)	1.81 (.57)	2.19 (.83)	13
	Overall	2.67 (.54)	3.00 (.59)	1.80 (.54)	2.19 (.88)	24
Overall	1	2.66 (.49)	3.02 (.43)	1.70 (.46)	2.02 (.84)	25
	2	2.63 (.39)	3.09 (.50)	1.70 (.47)	2.10 (.74)	24
	Overall	2.64 (.44)	3.06 (.46)	1.70 (.46)	2.06 (.78)	49

9.4 Discussion

Are LCBIs generated in expository film comprehension?

Experiment 2 was conducted to analyze whether LCBIs are generated online in expository film comprehension. A naming task was implemented to unobtrusively assess LCBI generation and to not require the understanders' monitoring of their own inference processes. The results of experiment 2 confirmed the hypothesis that LCBIs are generated in expository film comprehension. Naming latencies are facilitated in an inference-related context compared to an inference-unrelated context. This is in line with most theories derived in narrative text comprehension literature, such as the Constructionist Theory (Graesser et al., 1994, 1997), that forecasts that local and global bridging inferences are generated online, or the Minimalist Hypothesis (McKoon & Ratcliff, 1992), that argues that local bridging inferences are generated online when they are necessary for local coherence.

Considering the special context of expository material, the results are even more substantial, because it is not self-evident that LCBIs are generated online in text comprehension (see Chapter 4). As working memory resources are limited (Baddeley, 1992), the number and variety of inferences that accompany comprehension are significantly restricted (e.g., Graesser et al., 1994, 1997; Singer et al., 1994). Expository text comprehension literature (e.g., Noordman et al., 1992; Singer et al., 1997; Wiley & Myers, 2003, see Section 4.2) suggests that it is crucial for an online LCBI generation that readers have sufficient time and sufficient prior knowledge, that readers encounter a reading task that fosters an inference generation, and that the readers encounter a text that is not too difficult. This experiment successfully accounted for these factors.

It can be concluded from the current results that local causal bridging inferences are generated in expository film comprehension. This is in line with the assumption of higher-order, medium-independent comprehension mechanisms, such as generating a mental representation by drawing bridging inferences online (Graesser et al., 2001; Magliano et al., 1996, 2001; Zwaan & Radvansky, 1998).

Effect of the dynamic pictorial information on online local causal bridging inferences

It has been argued that the dynamic pictorial information in dynamic visualizations, such as films and animations, hinders a deeper elaboration, because, among other things, the processing of dynamic pictorial information demands too many cognitive resources (Lowe, 2003, 2004). However, the current pattern of results did not support this assumption. The present experiment did not reveal any differences in LCBI generation across presentation modes. Participants in the video and the audio condition showed the expected facilitation effect of context on their naming latencies equally, indicating an LCBI generation in expository film and audio tape comprehension. Thus, it can be concluded that the pictorial information in films did not hinder elaboration as claimed by some experts in the field (e.g., DeFleur et al., 1992; Lowe, 2003, 2004; Salomon, 1984; Weidenmann, 1989, 2002). In line with this reasoning, no differences in cognitive load measures across presentation modes were found.

This leads us to the conclusion that bridging inferences are mainly based on the verbal information (this is what both presentation modes had in common) as opposed to the pictorial information, or as opposed to combined verbal and pictorial information. However, the applied method (naming task) only focused on the verbal information. This seems logical, as the applied method and theories are borrowed from text comprehension, but this approach neglects the additional pictorial information in films. It could be that participants draw additional inferences based on the visual pictorial information. This brings us to the idea of developing an analysis scheme for visually presented pictorial information (analogous to propositional analyses as proposed by van Dijk and Kintsch, 1983) and the presentation of pictorial naming stimuli instead of verbal stimuli as applied by Unsöld and Nieding (in press). This approach might also be of interest for the broader field of learning with dynamic visualizations (e.g., animations, digital videos).

Methodological considerations

Naming belongs to the priming methods (or probe techniques) in inference research. In priming methods (see Section 6.1), participants react to a stimulus that represents the inference that needs to be drawn to understand the passage. It is assumed that the generated inference primes the reaction to the stimulus. The latencies in an inference-related context are compared to a base line in an unrelated context that is not supposed to evoke the respective inference.

It has been argued that the latencies are not facilitated by the generated inference, but by the context at the time of the test (Potts et al., 1988). In other words, the found differences in naming latencies across different contexts are not due to the generated inference, but due to the context at the time of the test. Keenan et al. (1990) argued that naming latencies are so short (between average 400 – 600 ms in reading) that the task does not encourage nor require context-checking at the time of the test.

To test empirically whether context or generated inferences led to the found facilitation effect, the location of the naming task could be systematically varied. In the present experiment, the naming task was inserted at positions where the explicit material was omitted and where the LCBI had to be drawn to understand the content. If the context is solely responsible for the facilitation effect, then the facilitation effect should also be found if the naming word is presented at an earlier position than the current LCBI position, but at a position that already belongs to the respective context. This position should be earlier than the current LCBI position, because a later position cannot exclude the possibility that the LCBI had been drawn and caused the facilitation effect. Future experiments could vary the naming task position systematically to clarify the debate.

However, in the present experiment, the facilitation effect is just under 50 ms. It is hard to believe that, if the respective inference had not been generated, solely the activation of the context was strong enough to cause a facilitation effect of 50ms. Concretely, this means that activating the respective context would prime such different concepts, such as “to conduct”, “to reflect”, “lid” and even “comparison” (see appendix for different naming words), that should normally not be activated within the context of a thermos. In this experiment, a variety of naming words were facilitated that do not necessarily have anything to do with a thermos at first glance and without further thinking (and further thinking is an inference).

Even if some of the respective concepts were depicted in the pictorial trace (such as “lid”) and it could be argued that the pictorial information primed the naming words, the present empirical results contradict this hypotheses: first, only some of the naming words were concrete concepts and could, therefore, be presented visually (such as “lid”); most naming words were abstract concepts and thus could not easily be presented visually (such as “conduction”). Second, participants in the audio condition, who did not see the respective visual information, showed the facilitation effect. These arguments lead to the conclusion that it is the respective LCBI that causes the facilitation effect and not the context.

Another aspect that could influence the measurement of inference generation in a naming task is the SOA (stimulus onset asynchrony). Graesser et al. (2007) claim that automatic inferences are generated within approximately half a second of the onset of an inference-triggering word or sentence. Routine inferences are generated within a second or less. LCBI can be considered as one of the two. Consequently, the SOA in this experiment seemed to be sufficient (250 ms black screen + 500 ms asterisk = 750 ms SOA after the inference-triggering sentence) such that an LCBI generation could have taken place. To further analyze the time course of LCBIs in film comprehension, one could vary the SOAs systematically, e.g., 500 ms (in this instance, an LCBI might not have been generated yet if it belongs to the class of routine inferences), and, e.g., 750 ms. The factor SOA could be systematically varied in combination with the factor naming task position in a 2 (SOA, e.g., 500 ms vs. 750 ms) x 2 (early naming task position vs. current naming task position) design.

The occurring interactions could provide further clarification about the nature of LCBI generation. If a facilitation effect occurs at an earlier naming task position and at an SOA = 750 ms, then context checking is responsible for the facilitation effect. If a facilitation effect occurs at the current naming position and at an SOA = 750 ms, but not at the earlier naming task position and at an SOA = 750 ms, then the current findings are replicated and LCBIs are responsible for the facilitation effect. If the facilitation effect occurs at the current naming position and at an SOA = 500 ms, it can be concluded that LCBIs belong to the class of automatic inferences and not to the class of routine inferences (Graesser et al., 2007) and are generated within half a second. Millis and Graesser (1994) also varied systematically different SOAs (see Section 4.2). Further conclusions about the nature of LCBI processes could be drawn by systematically analyzing the interactions of both factors (position x SOA).

A different interesting variation of the present experiment is the position of the naming task in the IU context. In the present experiment, the IU naming task was presented at a position where no LCBI needed to be generated, because the explicit LCBI phrase was left in the text. This was done in preparation for

Experiment 3. However, this might have influenced the naming latencies, because participants possibly invested less mental activity at these positions compared to naming tasks presented at positions that required an LCBI generation. In other words, the latencies in the IU were possibly longer, because less mental activity was invested. It would be interesting to know whether the facilitation effect persisted, if the naming task in the IU context also occurred at positions that required an inference generation, but not the inference that was represented by the respective naming word.

While varying different SOAs, researchers will gain additional insight into the nature of inference processes. Moreover, by varying different methods, researchers can obtain further knowledge about the nature of inference generation (even though it seems a bit tedious). The conclusion from Experiment 1 was that even global bridging inferences are generated automatically and do not have access to a conscious processing. This conclusion could not have been tested when applying a priming method to global bridging inferences, because the latter would not have required a conscious monitoring of the inference processes.

As exemplified above, several ideas for future experiments could be derived from the current experiment. Nevertheless, it should be stated that the current experiment was the first experiment that successfully proved online inference processes in expository film comprehension and can, therefore, be considered as a first step towards a new approach in the area of cognitive film comprehension research.

Generalizations of the present findings

As stated in Section 2.2, it is assumed that the present findings are based on a higher order medium-independent comprehension mechanism. Thus, these findings can probably be transferred to other dynamic visualizations as well, such as animations. The current experiment revealed that the pictorial information did not hinder elaboration in films; therefore, it should not hinder elaboration in other instructional dynamic visualizations either. Nevertheless, this needs to be tested empirically because the findings might depend to some degree on the applied material. The “Sendung mit der Maus” (“Mouse Show”) is made for children and tries to explain complex facts at a relatively easy level. The pictorial information complements the verbal information and visually supports what is being said. The “text-picture gap” (“Text-Bild-Schere”, Wember, 1976), which is often criticized in films, hardly occurs in this material. To analyze whether the present findings can be transferred to other material, empirical studies that vary different sorts of material (i.e., different degrees of complementary verbal and pictorial information, different degrees of complexity and different domains) are necessary.

Can films make you learn?

As stated at the beginning of this paper, the main interest of this dissertation is whether films make you *think* and encourage cognitive processes, such as inferences. This question is associated with the broader question of whether films make you *learn*[7]. To answer the second question, in the past, media comparisons that analyzed whether learning outcomes are better for film versus text (Baggett, 1979; Beentjes & van der Voort, 1993; Neuman, 1989, 1992; Salomon, 1984) or better for film versus audio (Beagles-Roos & Gat, 1983; Furnham, de Siena & Gunter, 2002; Gibbons, Anderson, Smith, Field & Fischer, 1986; Meringoff, 1980) have dominated the field. However, this line of research that asks what medium is best suited to foster learning has been criticized among other things because the answer to this question depends to a highly degree on the type of learning processes addressed (Scheiter, Gerjets, Vollmann, & Catrambone, 2006).

A more interesting question is what type of information can be best conveyed by which medium and which type of cognitive processes can be best supported. In line with this approach, it would seem beneficial for learning from film, if the text-picture gap (Wember, 1976) and an unnecessary redundancy between pictorial and verbal information is avoided (see also Mayer, 2001, 2005). Instead, the verbal and pictorial information should pursue complementary roles in generating a comprehensive mental model. Moreover, it would seem beneficial for learning, if the information is presented at an adequate speed and not too transiently [as dynamic visualizations are often blamed for (Lowe, 1999)], in order to avoid the "missing half-second" ("die fehlende Halbsekunde", Sturm, 1984) and give the recipients the possibility to process the presented content adequately. The impact that the transience of the presented information has on inference generation or on learning outcomes could be analyzed empirically in future experiments. This question could not be addressed in the current experiment, because the audio and video presentation mode had the same transience. In future experiments, it could be analyzed whether the option to reduce the transience of the presented material improves learning outcomes. Possible methods to reduce the transience of dynamic visualizations comprise interactive features, as in interactive videos or simulations, (Schwan & Riempp, 2004), or the integration of multiple static stills depicting crucial information into the film (Arguel & Jamet, 2007) or the learning environment (Rebetez, Bétrancourt, Sangin, & Dillenbourg, 2008).

Summary

Experiment 2 could show that LCBIs are generated in film comprehension. This is rather astonishing, because experiments in the domain of text processing did not always find evidence for an online generation of LCBIs in expository text comprehension. Moreover, it is not self-evident that films are elaborated at all: some researchers argued that films are processed superficially (Salomon, 1984; Weidenmann, 1989, 2001). The findings of this experiment contradict these assumptions. LCBI generation as an elaboration process necessary for comprehension occurs during expository film comprehension. Therefore, the assumption that films are processed superficially is not supportable, at least not if the above-mentioned factors crucial for an inference generation in expository material (e.g., an instruction or a reading goal that encourages the comprehension processes, information necessary to compute the inferences is easily available, and the understander has at least some prior knowledge to relate to the new information) are considered.

[7] This question is raised again in experiment 3.

10 Experiment 3: Local Causal Bridging Inferences: How do they Relate to Offline Comprehension Performance?

Experiment 2 showed that LCBIs are generated online in expository film comprehension. The question that arises is whether the online generated LCBIs that are necessary for comprehension (Kintsch, 1998; van Dijk & Kintsch, 1983) lead to a coherent mental representation of the content, called the *episodic text memory* (Kintsch, 1998). The episodic text memory is a unitary structure, but, for analytical reasons, it is divided into two components: the textbase and the situation model (see Section 3.1). The text by itself usually does not lead to a coherent episodic text memory. To generate a coherent episodic text memory, reader-generated inferences are necessary. The generation of a coherent episodic text memory (also described as a "mental model") is often assessed by offline measures, such as retention and comprehension. If reader-generated inferences are necessary for a coherent mental model, and, if the mental model comprises the textbase and the situation model, and, if bridging inferences are necessary to form a coherent textbase (as they are), bridging inferences are needed for a coherent mental model, as well.

The question that arises is whether online generated bridging inferences are sufficient for generating a coherent mental model? In other words, if understanders generated bridging inferences online, will they automatically generate a coherent model, too? Or are bridging inferences indeed necessary, but not sufficient for a coherent mental model? The current experiment investigates that research question by analyzing whether the online generated LCBIs shown in Experiment 2 will automatically lead to a coherent mental model, as indicated by offline measures, such as retention and comprehension. If it turns out that bridging inferences are not sufficient for a coherent mental model, possible moderators are discussed. Thus, Experiment 3 studies the initial research question concerning the relations between online generated bridging inferences and offline comprehension scores (see Experiment 1). It could be that those understanders, who generate LCBIs online, will automatically generate a coherent mental model. However, on the other hand, it sounds also plausible that even understanders who generate LCBIs online will not automatically generate a coherent mental model, as, in addition to bridging inferences, other cognitive processes, such as elaborative inferences and the integration of new information with prior knowledge structures, are crucial for a coherent mental model building.

There may be questions as to why the relationships between online generated bridging inferences and offline comprehension measures are not analyzed in one comprehensive experiment as we did in Experiment 1. Several factors led up to the decision to separate the online assessment of LCBIs from the examination of their relationships to a coherent mental model. First, it was not yet certain that the method used in Experiment 2 would prove to be valid for measuring online generated LCBIs. It was important to validate the method before designing an experiment that also tests the relationships between online LCBIs and offline comprehension measures. Second, the naming task might influence the natural learning process. Even if the naming task inserted in the film is considered to be unobtrusive and participants rated it as "not very disrupting" in Experiment 2, the fact that a word needs to be named, and then is asked for again in the offline comprehension test, might influence comprehension scores. Moreover, highlighting the naming word and then asking for it again in the offline comprehension test is not a real world setting and, therefore, the obtained results in the offline comprehension test might lack external validity with regard to a more naturalistic film viewing situation as in, e.g., schools. Participants might think about the naming words later on. This might initiate cognitive processes that would not have occurred without the naming task. Thus, two separate experiments were conducted.

It is assumed that the results from Experiment 2 will transfer to Experiment 3. In other words, it is expected that participants in Experiment 3 will also generate the LCBIs online, even if not measured directly. An offline comprehension test administered after the respective film will analyze how the online generated LCBIs relate to offline comprehension scores.

The current experiment aims at answering the following research question:

1. Does the online generation of local causal bridging inferences lead to a coherent mental model as indicated by offline measures of comprehension?
2. Does the presentation of dynamic pictorial information hinder the generation of a coherent mental model?

10.1 Hypotheses

Based on the theoretical assumptions outlined in the previous chapters, the following hypotheses were derived.

H1: Generation of a mental model

If participants generate a mental model or in other words learn from the experimental material (film/audio trace) used in Experiment 2 and 3, participants in the experimental conditions (video/audio) in Experiment 3 will outperform participants in the control condition (no experimental material) in the comprehension test. If participants do not generate a mental model or do not learn from the experimental material, then there should be no differences between the experimental conditions on the one side and the control condition on the other side.

H2: Integration of implicit information into a coherent mental model

If the online generation of LCBIs leads to the comprehension of the content presented, and, if the inferred information is integrated with long-term memory structures into a coherent mental model, then participants' comprehension scores on implicit items should be superior to the comprehension scores of those participants, who had not obtained any experimental material. Additionally, for participants in the experimental conditions, comprehension scores on the items that ask for information *implicitly* stated in the experimental material should be as good as participants' comprehension scores on the items that ask for information *explicitly* stated in the experimental material. If no differences in comprehension scores between implicit and explicit information can be found, then participants inferred the required information and were able to integrate the inferred information with prior knowledge structures from long-term memory into a substantial coherent mental model.

H3: Effect of the dynamic pictorial information

Since the dynamic pictorial information did not have an effect on various dependent variables in Experiment 1 and 2, no differences concerning the comprehension scores between video and audio presentation modes are expected in this experiment either. In other words, it is expected that the films' dynamic pictorial information does not affect the generation of a mental model as indicated by offline comprehension scores.

10.2 Method

The variation of the two factors resulted in a 3 (presentation mode: video, audio, control) x 2 (item type: implicit, explicit) mixed factorial design with presentation mode as a between-subjects factor and item type as a within-subjects factor. Table 10.1 illustrates the design graphically.

As in Experiment 2, the different word sets were varied across participants, but only for methodological reasons. Consequently, word set was not considered a true between-factor, but was included in the respective analyses to control its effect statistically.

Table 10.1: Design.

Presentation mode (between)	Item type (within)	
	Implicit	Explicit
Video	N = 32	N = 32
Audio	N = 32	N = 32
Control (N = 30)	Factor item type not applicable	

Independent variables

Two independent variables were implemented in this experiment: *Presentation mode* and *item type*. Presentation mode was varied as a between-factor and participants were randomly assigned to a video, an audio or a control condition. Participants in the video condition viewed the film, participants in the audio condition listened to the audio trace of the film, but instead of the visual trace, they saw a black screen, and participants in the control condition answered the comprehension questions without receiving any experimental material (i.e., film, audio tape). Additionally, the item type was varied as a within-factor. The items in the comprehension test referred to two types of information included in the experimental material: implicit vs. explicit. *Implicit* information means that the LCBI phrase that is needed to make the text locally coherent at the respective position was omitted in the experimental material. Therefore, to enable an online comprehension, participants had to draw an LCBI at this position. Thus, the implicit item in the offline comprehension test can be considered as measuring the fact that an inference has been generated. *Explicit* information means that the LCBI phrase that was needed to make the text locally coherent at the respective position was explicitly stated in the experimental material. Therefore, the understander did not need to generate online an LCBI at this position. However, to answer the explicit item in the offline comprehension test correctly, the understander needed to have integrated the explicitly given information into a mental model to correctly reject the distractors.

The implicit items referred to positions in Experiment 2, where the naming task was presented in the IR context; whereas the explicit items corresponded to positions in Experiment 2, where the naming task was presented in its IU context.

Dependent variables

The dependent variable was the comprehension scores in a comprehension test. For the analysis of H1, the overall comprehension scores were of interest; whereas, for the analysis of H2, the comprehension scores for implicit vs. explicit items were of interest.

Control variables

To control for the effects of *age*, *prerequisite knowledge* and *interest* on the obtained results, these variables were included as control variables.

Participants

Ninety-four undergraduate students (67 female, 27 male) from the University of Tuebingen, Germany, participated in this study. Their average age was 23.43 years (SD = 3.46) and ranged from 19 to 36 years. 27.7% of the participants were psychology majors. The remaining 72.3% majored in other academic domains, such as business administration, law, etc. Participants in the experimental conditions (video vs. audio, N = 64) received 7€ for participation. Participants in the control condition (N = 30) received 5€ for participation due to the shorter duration of the experiment. Psychology majors also had the possibility to choose course credit instead of the payment. German as the native language was required. All participants had normal vision or visual aids (glasses, contact lenses). The experiment lasted approximately 45 minutes in the experimental conditions and 30 minutes in the control condition.

Apparatus

The experimental procedures were controlled by an HP Compaq laptop and programmed using MediaLab, an experimental software. Films were presented at full screen on a 17” monitor.

Materials

The experimental material (film, audio tape) corresponded to the experimental material used in Experiment 2. However, in the current experiment, the film[8] was presented as a whole and was not interrupted by the naming task. When the LCBI phrase was explicitly stated in the film (cf. IU context in Experiment 2), then the corresponding item in the comprehension test that asked for this information was called *explicit item.* When the LCBI phrase was omitted, participants had to generate the LCBI themselves to make the film coherent (cf. IR context in Experiment 2). The item in the comprehension test that asked for this information was called *implicit item.* The variation of the factor *item type* was presented within a) the experimental material and b) the comprehension test, each time alternating. For instance, first predefined LCBI position: implicit information in the experimental material, implicit item in the comprehension test; second predefined LCBI position: explicit information in the experimental material, explicit item in the comprehension test; third predefined LCBI position: implicit information in the experimental material, implicit item in the comprehension test, and so on. Whether this sequence started with implicit or explicit information was counterbalanced. Only when presented as implicit information did participants have to generate an LCBI. When presented explicitly, the LCBI was explicitly stated in the film and the film was, therefore, coherent at that position. The predefined LCBI positions corresponded to the LCBI positions in Experiment 2. The IR context in Experiment 2 corresponded to the implicit item type in Experiment 3; the IU context in Experiment 2 corresponded to the explicit item type in Experiment 3. Refer to the original material in the appendix.

The *comprehension test* was constructed and discussed with three other people with domain knowledge. The revised version was tested on three pilot people. Their scores revealed a medium item difficulty. Scores ranged from 61.53% correct without watching the films to 84.61% correct with watching the films.

The comprehension test consisted of 13 items. Seven items addressed the “construction of a thermos”, whereas six items addressed the “development of lightning”. The items covered the LCBIs that had to be drawn in Experiment 2. However, in Experiment 2, overall 14 LCBIs in the course of the two films had to be drawn. In Experiment 3, one item was omitted: one LCBI that had to be drawn in Experiment 2 was that of why the protagonist took two glass beakers (in the “thermos” film). The answer was that he wanted *to compare* the effect of air on heat conduction. The inference that the protagonist wanted to compare two experimental settings corresponded to a motivational causal bridging inference (Singer, 1994) that is not suitable for a physics comprehension test. Therefore, the concept “comparison” was omitted as a comprehension test item in Experiment 3. Consequently, there were 13 comprehension test items overall.

[8] As in the other experiments, the term „film“ is used for the film <u>and</u> the audio trace for a better readability.

The items were presented in a multiple choice format: one correct answer plus three distractors (see Section 6.2). The items asked for the same concepts that were covered by the online LCBIs in Experiment 2. Therefore, the correct answer to each comprehension test item reflected the online generated LCBI (measured in Experiment 2) plus its integration with long-term memory structures into a coherent mental model. Moreover, the correct answer contained the naming word from Experiment 2 (the naming word reflected the LCBI that had to be drawn at the respective position). To avoid word-based memory effects, one of the three distractors also contained the naming word from Experiment 2. The items' order of appearance in the comprehension test corresponded to their order of appearance in the films. Refer to the following passage, illustrating two examples from the comprehension test; one item in the domain of the thermos film and one item in the domain of the lightning film.

Thermos
Containers made of different materials (wood, rubber, glass, metal, etc.) vary in their ability to keep hot fluids warm. Why is this?
A hot fluid turns cold, because its heat is conducted to the environment, whereas heat can only be conducted by material…

Answer 1: … and materials vary in their heat conductivity (e.g., metal is a better heat conductor than wood).
Answer 2: … and materials vary in their density (e.g., metal has a higher density than wood).
Answer 3: … and materials vary in their specific weight (e.g., metal has a higher specific weight than wood).
Answer 4: … and materials vary in their electric conductivity (e.g., metal has a higher electric conductivity than wood).

The right answer is answer 1. The distractor that also includes the naming word from Experiment 2 is answer 4. The naming word for the corresponding LCBI position in Experiment 2 was "to conduct".

Lightning
When rubbing on the air, falling raindrops get charged with electricity. Thereby, lightning may be generated. Which one of the four following explanations for the generation of lightning is correct?

Answer 1: The raindrops' storage capacity for electricity is limited and when this limit is reached, they unload in the form of lightning.
Answer 2: Falling raindrops increase their volume. When they have reached a critical size, they divide and lightning is generated.
Answer 3: As soon as falling raindrops go below a critical distance from earth, they unload in the form of lightning.
Answer 4: Falling raindrops heat up when they reach warmer air layers. When a critical temperature and, therewith, a critical intensity of molecule movement is reached, lightning is generated.

The right answer is answer 1. The distractor that also includes the naming word from Experiment 2 is answer 3. The naming word for the corresponding LCBI position in Experiment 2 was "to unload".

To measure participants' *interest in physics and chemistry,* participants answered the same interest items as in Experiment 2.

To measure the *cognitive load* participants experienced while watching the films, participants answered the same cognitive load items as in Experiment 2.

At the end of the experiment, participants answered the same *prerequisite knowledge test* items as in Experiment 2.

The experiment contained three presentation modes: video, audio and control. Participants in the control condition did not watch the films. This fact modified the control condition's procedure from that of the experimental conditions (video, audio). For that reason, the procedure will first be described for the experimental conditions, followed by the description of the procedure for the control condition.

Procedure

Experimental conditions

The experiment was run computer-controlled and quasi-individually. Participants in groups of six were tested together in one room. Participants sat in a room with six desks that were separated by movable walls. Each participant worked at his own laptop and could not see the other participants. The experimenter welcomed each participant and led him to his laptop. The experimenter was present during the experiment. When all participants were present, the experimenter started the experiment individually at each participant's laptop, one after another. Each participant listened to the experimental material with individual headphones. The experiment started by asking five items about the participant's interests in physics and chemistry. Afterwards, participants stated their demographical data. Then, the first film started. Participants were instructed to follow the film attentively and to try to comprehend the content. It was announced that, after the film comprehension test, questions regarding the film content would follow. After the film, participants first answered the comprehension questions and, second, the cognitive load items. Then, the second film started and participants proceeded in the same manner. After answering the comprehension and cognitive load items regarding the second film, participants took the prerequisite knowledge test. When finished, participants waited individually at their desks until everyone had completed the experiment. Participants were not allowed to leave while other participants were still working in order to avoid distracting noises and disruptions. When everyone was finished, participants received their money or course credit and left. Participants did not complete a working memory span test as in Experiment 2, as the test had not yielded any effects in Experiment 2. Thus, it was not expected that it would in this experiment either. Table 10.2 shows the procedure for the experimental conditions at a glance.

Table 10.2: Procedure for the experimental conditions.

	Video (N = 32)	Audio (N = 32)
1	Overall instruction	Overall instruction
2	5 items asking about interest in physics & chemistry	5 items asking about interest in physics & chemistry
3	Demographic data	Demographic data
4	Instruction	Instruction
5	Reception phase 1: Experimental film 1	Reception phase 1: Experimental film 1
6	Test phase 1: 6 or 7 comprehension test items regarding film 1	Test phase 1: 6 or 7 comprehension test items regarding film 1
7	6 cognitive load items regarding film 1	6 cognitive load items regarding film 1
8	Reception phase 2: Experimental film 2	Reception phase 2: Experimental film 2
9	Test phase 2: 6 or 7 comprehension test items regarding film 2	Test phase 2: 6 or 7 comprehension test items regarding film 2
10	6 cognitive load items regarding film 2	6 cognitive load items regarding film 2
11	6 items asking for prerequisite knowledge	6 items asking for prerequisite knowledge

As in Experiment 2, each participant in the experimental conditions saw either version 1 or version 2 of each film (thermos 1 + lightning 1, thermos 2 + lightning 2). As in Experiment 2, the different film versions included different word sets that were varied (implicit vs. explicit). See the appendix for the original material. The films' presentation order and the film versions ("word set" in the following) were counterbalanced. This resulted in eight conditions (see Table 10.3) that corresponded to the conditions in Experiment 2. The eight conditions were administered due to methodological reasons. For content purposes only, the presentation modes (video vs. audio vs. control) and item type (implicit vs. explicit) were of interest. However, the effect of the factor "word set" was controlled statistically in the analyses. When starting with the thermos film in reception phase 1, participants then also started with those comprehension questions referring to the thermos film in test phase 1. When starting with the lightning film in reception phase 1, participants likewise started with those comprehension questions referring to the lightning film in test phase 1. Overall, participants answered the same comprehension questions independent from their experimental condition.

Table 10.3: Experimental conditions.

Condition	Video (N = 32)	N	Condition	Audio (N = 32)	N
1	Thermos 1, Lightning 1	8	5	Thermos 1, Lightning 1	7
2	Lightning 1, Thermos 1	8	6	Lightning 1, Thermos 1	9
3	Thermos 2, Lightning 2	8	7	Thermos 2, Lightning 2	9
4	Lightning 2, Thermos 2	8	8	Lightning 2, Thermos 2	7

Control condition

The material and procedure for the control condition were identical to the experimental conditions, except that participants in the control condition did not receive any experimental material and thus also did not answer cognitive load items. As a consequence, the overall duration of the experiment was reduced to approximately 30 minutes.

Table 10.4: Procedure for the control condition.

	Control (N = 30)
1	Overall Instruction
2	Interest
3	Demographic data
4	Instruction
5	13 comprehension test items
6	6 items asking about prerequisite knowledge

The presentation order of the comprehension test items was counterbalanced. Half of the participants in the control condition answered comprehension test items regarding the construction of a thermos first; the other half answered comprehension test items regarding the formation of lightning first. This resulted in two further methodological conditions regarding the control condition. Refer to Table 10.5.

Table 10.5: Control conditions.

Condition	Control (N=30)	N
9	Thermos items, Lightning items	14
10	Lightning items, Thermos items	16

10.3 Results

Control variables

To ensure that the control variables (*interest, prerequisite knowledge, age*) were equally distributed across conditions, several ANOVAs were conducted. In a first step, three one-factorial ANOVAs were conducted. The analyses revealed that *interest* and *age* were equally distributed across the presentation modes (all *F*s < 1). *Prerequisite knowledge*, however, was not distributed equally across presentation modes ($F(2,91) = 3.65$, $MSE = 342.70$, $p < .05$, $\eta^2 = .07$). Post-hoc analyses with a Bonferroni test revealed that participants in the audio condition had significantly less prerequisite knowledge than participants in the video condition and than participants in the control condition (p< .05). Refer to Table 10.6 for means and standard deviations.

Table 10.6: Means and (standard deviations) for the control variables as a function of presentation mode.

Presentation mode	Interest	Prerequisite knowledge in % correct	Age	N
Video	2.00 (.40)	57.81 (17.95)	24.06 (3.34)	32
Audio	1.93 (.62)	45.83 (19.40)	23.25 (3.76)	32
Control	2.01 (.58)	55.00 (18.13)	22.93 (3.27)	30
Overall	1.98 (.54)	52.83 (19.03)	23.43 (3.46)	94

Furthermore, the equal distribution of the control variables (interest, prerequisite knowledge, age) was analyzed across *word sets.* The factor “word set” was introduced in the design for methodological reasons (see above and Experiment 2). The factor “word set” was not applicable to the control condition, because participants in the control condition did not receive any experimental material. As a consequence, the factor “word set” was excluded from the aforementioned analyses, but was included in three separate 2 (presentation mode) x 2 (word set) ANOVAs that measured the distribution of the control variables only for the experimental conditions (video, audio). The analyses revealed that *interest* was not distributed equally across word sets ($F(1,60) = 5.91$, $MSE = .25$, $p < .05$, $\eta^2p = .09$). Moreover, *prerequisite knowledge* was not distributed equally across presentation modes ($F(1,60) = 7.19$, $MSE = 319.16$, $p < .01$, $\eta^2p = .11$) and word sets ($F(1,60) = 7.19$, $MSE = 319.16$, $p < .01$, $\eta^2p = .11$). Participants in word set 2 showed less prerequisite knowledge and less interest than participants in word set 1. Additionally, as already mentioned above, participants in the audio condition demonstrated less prerequisite knowledge than participants in the video condition. Age, on the other hand, was equally distributed across word sets ($F(1,60) = 1.82$, $MSE = 12.39$, $p > .10$, $\eta^2p = .02$). See Table 10.7 for means and standard deviations.

Table 10.7: Means and (standard deviations) for the control variables as a function of presentation mode and word set.

Presentation mode	Word set	Interest	Prerequisite knowledge in % correct	Age	N
Video	1	2.11 (.50)	65.62 (15.48)	25.19 (3.60)	16
	2	1.89 (.23)	50.00 (17.21)	22.94 (2.72)	16
	Overall	2.00 (.40)	57.81 (17.95)	24.06 (3.34)	32
Audio	1	2.13 (.70)	50.00 (16.10)	23.31 (3.65)	16
	2	1.74 (.48)	41.67 (21.94)	23.19 (3.99)	16
	Overall	1.93 (.62)	45.83 (19.40)	23.25 (3.76)	32
Overall	1	2.12 (.59)	57.81 (17.45)	24.25 (3.69)	32
	2	1.81 (.38)	45.83 (19.86)	23.06 (3.36)	32
	Overall	1.97 (.52)	51.82 (19.50)	23.66 (3.55)	64

Comprehension scores

To investigate whether participants learned from the experimental material, a one-factorial ANCOVA was conducted. *Presentation mode* was included as the independent variable, the *overall comprehension performance* as the dependent variable. *Prerequisite knowledge* was included as a covariate, because it was not equally distributed across presentation modes. *Word set* was not analyzed here, because word set was not applicable for the control condition. The analysis yielded a significant main effect for presentation mode ($F(2,90) = 20.43$, $MSE = 166.00$, $p < .001$, $\eta^2 p = .31$) and prerequisite knowledge ($F(1,90) = 6.46$, $MSE = 166.00$, $p \leq .05$, $\eta^2 p = .07$). Post-hoc Bonferroni analyses revealed that video and audio did not differ significantly ($p > .10$), but both experimental conditions (video, audio) differed from the control condition ($p < .001$). Overall comprehension test performance was better in the video and audio condition than in the control condition. Refer to Figure 10.1 for a graphical illustration of the data.

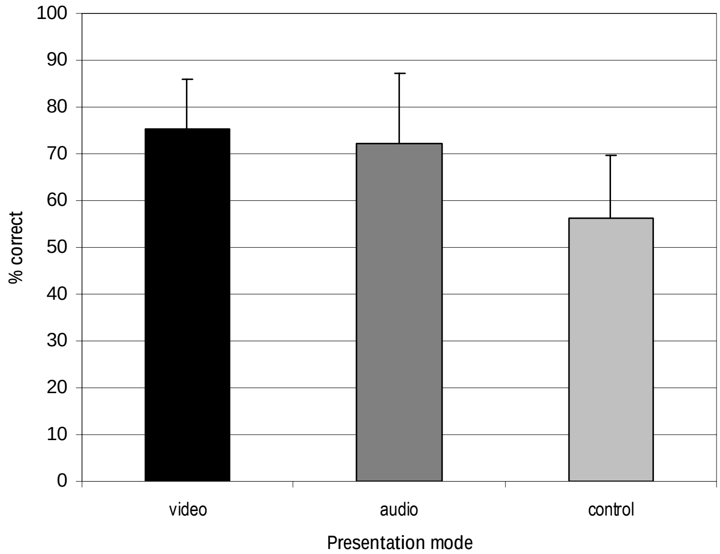


Fig. 10.1: Means and standard deviations (in % correct) of the comprehension test performances as a function of presentation mode.

Item type

To test H2, the comprehension test items were divided into implicit and explicit items across both films. *Implicit items* were those items whose answers were only stated *implicitly* in the experimental material and thus are supposed to be inferred online by the understander. *Explicit items* were those items whose answers were stated *explicitly* in the experimental material and thus did not have to be inferred online by the understander. However, not only the implicit but also the explicit items required a deeper level of offline understanding (beyond mere retention) to correctly identify the right answer from the distractors in the multiple choice comprehension test.

To analyze whether participants in the experimental conditions were able to integrate the online inferred information into a coherent mental model, it was analyzed whether comprehension performance for implicit items was better for participants in the experimental conditions than for participants in the control condition, who did not obtain any experimental material. A one-factorial ANCOVA for these implicit items was conducted with presentation mode as independent variable. Prerequisite knowledge was included as a covariate, because it was not equally distributed across presentation modes. The analysis revealed a significant effect for presentation mode ($F(2,90) = 9.26$, $MSE = 250.43$, $p < .001$, $\eta^2p = .17$) and prerequisite knowledge ($F(2,90) = 7.12$, $MSE = 250.43$, $p < .01$, $\eta^2p = .07$). Post-hoc Bonferroni analyses revealed that participants' implicit comprehension performance in video and audio did not differ ($p > .10$), but participants in both experimental conditions outperformed participants in the control condition ($p < .01$).

This finding is a first indication that participants in the experimental conditions were able to integrate the inferred information into a coherent mental model. However, it could be that other factors, such as "time on task" during the reception phase, led to the superiority of implicit comprehension scores for participants in the experimental condition over participants in the control condition. To rule out this possibility, a second analysis was conducted to investigate whether participants in the experimental conditions integrated the implicit information as coherently into a mental model as they did the explicit information.

To test this assumption, a 2 (item type) x 2 (presentation mode) x 2 (word set) ANCOVA with interest and prerequisite knowledge as covariates was conducted. The covariates were included, as interest and prerequisite knowledge were not equally distributed across word sets. The methodological factor "word set" was included to control its effect statistically. If participants answered implicit items as correctly as they did explicit items and, therefore, no differences in comprehension scores across item types occurred, then participants integrated the inferred information as coherently into a mental model as they had the explicit information. However, if a difference in comprehension scores across item types occurred and if this difference is in favor of the explicit items, then participants were indeed able to infer the implicit information online, but they were not able to integrate the new information into a coherent mental model as they had been able to integrate explicitly stated information. A mental model was thus only generated for the explicitly stated information, but not for the implicitly stated information. The analysis revealed a significant interaction of item type x word set ($F(1,57) = 8.27$, $MSE = 230.22$, $p < .01$, $\eta^2p = .13$). No other main effects or interactions were found (all Fs < 1). See Table 10.8 for means and standard deviations.

Table 10.8: Means and (standard deviations) of comprehension test performances (in % correct) as a function of item type, presentation mode and word set.

Item type	Presentation mode	Word set	Comprehension test performances in % correct	N
Implicit	Video	Word set 1	78.13 (15.77)	16
		Word set 2	65.89 (14.21)	16
		Overall	72.01 (16.02)	32
	Audio	Word set 1	70.83 (19.72)	16
		Word set 2	67.19 (18.19)	16
		Overall	69.01 (18.75)	32
	Overall	Word set 1	74.48 (17.95)	32
		Word set 2	66.54 (16.07)	32
		Overall	70.51 (17.37)	64
Explicit	Video	Word set 1	75.52 (15.13)	16
		Word set 2	81.25 (13.44)	16
		Overall	78.39 (14.37)	32
	Audio	Word set 1	67.45 (22.16)	16
		Word set 2	82.29 (16.63)	16
		Overall	74.87 (20.70)	32
	Overall	Word set 1	71.48 (19.11)	32
		Word set 2	81.77 (14.88)	32
		Overall	76.63 (17.77)	64

The interaction between item type and word set means that participants in word set 2 showed a difference in comprehension scores for implicit vs. explicit items in favor of the explicit items ($F(1,31) = 15.76$, $MSE = 235.56$, $p < .001$, $\eta^2 = .34$), but participants in word set 1 did not ($F < 1$). See Table 10.9 for means and standard deviations. This means that only participants in word set 1 were able to integrate the inferred information into a mental model as coherently as the explicit information. This seems awkward at first glance, because the factor “word set” was only introduced for methodological reasons and it was not expected that it would have an effect at all.

Upon closer examination, it might be reasonable to infer that the relation between word set and comprehension performance might have been mediated by the fact that participants in word set 1 had more prerequisite knowledge than participants in word set 2 (see descriptive analyses). Therefore, participants in word set 1 might have been able to integrate the online generated LCBIs into long-term memory structures, such as prerequisite knowledge, and thus into a coherent mental representation as measured by comprehension scores for implicit items; but participants with low prerequisite knowledge, as in word set 2, were not able to integrate the inferred information with long-term memory structures, because they were lacking those structures. To test this assumption, two mediator analyses were conducted.

Furthermore, interest was not equally distributed across word sets either and it is theoretically possible that interest also mediated the relation between word set and comprehension scores. Even though it is less likely that interest detached from prerequisite knowledge has an effect on the integration of online inferred information into a coherent mental representation, as interest as a motivational variable is not represented in cognitive long-term memory structures as prerequisite knowledge is, its mediating effect should be tested statistically.

Four separate mediator analyses (Baron & Kenny, 1986; James & Brett, 1984; Kenny, 2007) were conducted that examined the mediating effect of prerequisite knowledge and of interest on the relation between word set and comprehension scores for implicit and explicit items. A mediator analysis examines whether the relation between the predictor and the outcome variable is mediated by a third variable. A mediator analysis comprehends four steps that will briefly be explained in the following passage. However, the interested reader is referred to the aforementioned references for further details. The four steps that belong to a mediator analysis comprehend three regression analyses that analyze the relations between predictor, mediator and outcome variable and one statistical test (Sobel test; MacKinnon, Lockwood, Hoffman, West, & Sheets., 2002; MacKinnon, Fairchild, & Fritz, 2007; Preacher & Leonardelli, 2007; Sobel, 1982) that compares whether the addition of the mediator to the model reduces the effect of the predictor on the outcome variable. Each of the four steps relies on the significance of the previous step. In other words, if one regression fails to reveal a significant relation between the respective predictor and outcome variable, the possible mediator does not mediate the effect. In this analysis, word set was included as predictor variable, whereas comprehension scores for implicit as well as explicit items were included as outcome variables and either prerequisite knowledge or interest were included as possible mediator variables. In the following paragraphs, the four mediator analyses are described.

A first regression analysis revealed that word set marginally predicted the comprehension scores for implicit items (β = -.23, p < .10; see path c in Figure 10.2). The beta weight is negative, signifying that participants in word set 1 had better comprehension scores for implicit items than participants in word set 2. A second regression analysis revealed that word set successfully predicted the possible mediator variable prerequisite knowledge (β = -.31, p < .05, path a). The beta weight is negative, signifying that participants in word set 1 had more prerequisite knowledge than participants in word set 2. A third regression analysis revealed that prerequisite knowledge successfully predicted the outcome variable comprehension scores for implicit items (β = .28, p < .05, path b). The beta weight is positive, signifying that participants with more prerequisite knowledge scored higher on the implicit items than participants with low prerequisite knowledge. To test whether the effect of word set (predictor) on comprehension scores for implicit items (outcome) shrinks upon adding prerequisite knowledge (mediator) to the model, a Sobel test (MacKinnon et al., 2002; 2007; Preacher & Leonardelli, 2007; Sobel, 1982) was conducted. The Sobel test is significant when the mediator mediates the effect from predictor on outcome variable. For the current analysis, the Sobel test reached marginal significance (Sobel's $t = -1.70$, $p < .10$).

In other words, prerequisite knowledge mediated the effect of word set on comprehension scores for implicit items. A variable is called a complete mediator, if the relation between the predictor and the outcome is reduced to zero when the mediator is included in the model. A variable is called a partial mediator, if the relation between the predictor and the outcome is greater than zero, but gets significantly smaller when the mediator is included in the model. In the current experiment, prerequisite knowledge can be regarded as a complete mediator, because the effect of word set on comprehension scores for implicit items was reduced to zero ($\beta = -.16$, $p > .10$) when the mediator "prerequisite knowledge" was included in the model ($\beta = .23$, $p < .10$). Figure 10.2 illustrates the relations between word set and comprehension scores for implicit items when the mediator "prerequisite knowledge" is included in the model.

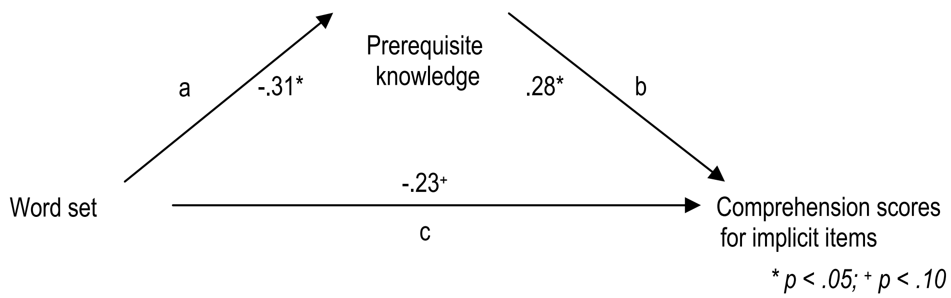


Fig. 10.2: Model when prerequisite knowledge is included as a mediator variable in the relation between word set and comprehension scores for implicit items.

A second mediator analyses examined the mediating effect of prerequisite knowledge on comprehension scores for explicit items. The first regression analysis revealed a significant effect of word set on comprehension scores for explicit items ($\beta = .29$, $p < .05$). The beta weight is positive, signifying that participants in word set 2 scored higher on explicit items. The second regression analysis revealed a significant effect of word set on prerequisite knowledge ($\beta = -.31$, $p < .05$). The beta weight is negative; signifying that participants in word set 1 had more prerequisite knowledge than participants in word set 2. The third regression analysis failed to reveal a significant effect of prerequisite knowledge on comprehension scores for explicit items ($\beta = .13$, $p > .10$). Thus, prerequisite knowledge cannot be considered to be mediating the effect of word set on comprehension scores for explicit items.

In a third mediator analysis, it was examined whether interest mediated the effect from word set on comprehension scores for implicit items. The first regression analyses showed a significant effect of word set on comprehension scores for implicit items ($\beta = -.23$, $p < .10$). The beta weight is negative, signifying that participants in word set 1 scored higher on implicit items than participants in word set 2. The second regression analyses revealed a significant effect for word set on interest ($\beta = -.30$, $p < .05$). The beta weight is negative; signifying that participants in word set 1 had more interest than participants in word set 2. The third regression analysis revealed a marginal significant effect of interest on comprehension scores for implicit items (β = .22, p < .10). The beta weight is positive, signifying that participants with high interest scored higher on implicit items than participants with low interest. However, the Sobel test did not reach statistical significance (Sobel's t = -1.42, p > .10). When adding interest to the model, neither word set (β = -.18, p > .10) nor interest (β = .16, p > .10) yielded significant effects on the comprehension scores for implicit items. Thus, interest cannot be regarded as a mediator variable between the relations of word set on comprehension scores for implicit items.

In a fourth mediator analysis, the mediating effect of interest on the relation between word set and comprehension scores for explicit items was analyzed. The first regression analysis revealed a significant effect of word set on comprehension scores of explicit items (β = .29, p < .05). The beta weight is positive, signifying that participants in word set 2 scored higher on explicit items. The second regression analysis revealed a significant effect of word set on interest (β = -.30, p < .05). The beta weight was negative, signifying that participants in word 1 had more interest than participants in word set 2. However, the third regression analysis failed to reveal a significant effect for interest on comprehension scores for explicit items (β = .15, p > .10). Hence, interest does not mediate the relation between word set and comprehension scores for explicit items.

To summarize, prerequisite knowledge explains the interaction between item type and word set, because it completely mediates the relation between word set and comprehension scores for implicit items. The model is depicted in Figure 10.2. As the other possible mediators did not yield significant effects, they are not illustrated graphically. The following tables show the summarized results for all four mediator analyses, divided by the criterion comprehension scores for implicit items (Table 10.9) and comprehension scores for explicit items (Table 10.10).

Table 10.9: Summary of regression analyses for mediator analyses for the relation between word set and comprehension scores for implicit items.

	B	SE B	β
Step 1 (Path c)			
Comprehension scores for implicit items (C)			
Word set (P_1)	- 7.94	4.26	-.23$^+$
Word set (P_2)	- 7.94	4.26	-.23$^+$
Step 2 (Path a)			
Prerequisite knowledge (C_1)			
Interest (C_2)			
Word set (P_1)	-11.98	4.67	-.31*
Word set (P_2)	-.31	.13	-.30*
Step 3 (Path b)			
Comprehension scores for implicit items (C)			
Word set (P_1)	-5.50	4.40	-.16
Word set (P_2)	-6.28	4.44	-.18
Prerequisite knowledge (M_1)	.25	.11	.28*
Interest (M_2)	7.25	4.16	.22$^+$

Note. C = Criterion; P = Predictor; M = Mediator; $^+$ $p<.10$; * $p<.05$
P_1 = P if possible mediator = prerequisite knowledge, P_2 = P if possible mediator = interest

Table 10.10: Summary of regression analyses for mediator analyses for the relation between word set and comprehension scores for explicit items.

	B	SE B	β
Step 1 (Path c)			
Comprehension scores for explicit items (C)			
Word set (P_1)	10.29	4.28	.29*
Word set (P_2)	10.29	4.28	.29*
Step 2 (Path a)			
Prerequisite knowledge (C_1)			
Interest (C_2)			
Word set (P_1)	-11.98	4.67	-.31*
Word set (P_2)	-.31	.13	-.30*
Step 3 (Path b)			
Comprehension scores for explicit items (C)			
Word set (P_1)	12.96	4.40	.37**
Word set (P_2)	12.96	4.37	.37**
Prerequisite Knowledge (M_1)	.12	.12	.13
Interest (M_2)	4.99	4.31	.15

Note. C = Criterion; P = Predictor; M = Mediator; $^+$ $p < .10$; * $p <. 05$; ** $p < .01$
P_1 = P if possible mediator = prerequisite knowledge, P_2 = P if possible mediator = interest

Cognitive load

To analyze the effects of presentation mode and word set on participants' experienced *cognitive load measures* (effort, confidence, stress) during the reception phases, three 2 (presentation mode) x 2 (word set) ANOVAs were conducted. Analyses did not reveal any effects for presentation mode nor word set on all three cognitive load measures (all *F*s < 1). Refer to Table 10.11 for means and standard deviations.

Table 10.11: Means and (standard deviations) of three cognitive load measures as a function of presentation mode and word set.

Presentation mode	Word set	Effort	Confidence	Stress	N
Video	1	2.56 (.43)	3.11 (.49)	1.66 (.60)	16
	2	2.61 (.35)	2.98 (.35)	1.56 (.31)	16
	Overall	2.59 (.39)	3.05 (.42)	1.61 (.47)	32
Audio	1	2.63 (.52)	3.02 (.32)	1.77 (.40)	16
	2	2.66 (.30)	3.00 (.38)	1.67 (.46)	16
	Overall	2.64 (.42)	3.01 (.34)	1.72 (.43)	32
Overall	1	2.59 (.47)	3.06 (.41)	1.71 (.51)	32
	2	2.63 (.32)	2.99 (.36)	1.62 (.39)	32
	Overall	2.61 (.40)	3.03 (.38)	1.66 (.45)	64

10.4 Discussion

Did participants learn from film?

Participants in the experimental conditions (video, audio) outperformed participants in the control condition in their overall comprehension scores. In other words, participants were able to comprehend the film and generated a mental model. These findings contradict those assumptions that presume that film is not an effective learning tool (Salomon, 1984; Weidenmann, 1989, 2002). Average comprehension scores for participants in the experimental conditions lay at 73.80%. Those for participants in the control condition lay at only 56.15%. This is a pretty substantial comprehension gain of almost 18%. Moreover, the current finding that participants were able to comprehend the experimental material and to generate a mental model from it can be interpreted as a manipulation check for Experiment 2, where the generation of a mental model was not directly tested. It can be concluded that the experimental material applied in Experiments 2 and 3 was an effective instructional tool, because participants generated online local causal bridging inferences (Experiment 2) and demonstrated better offline comprehension scores than a control group (Experiment 3). This is good news for all students and teachers in school, who are eager to implement films in class (Hobbs, 2006).

Effect of the pictorial information on offline comprehension scores

It has been argued that the pictorial information in dynamic visualizations, such as films, hinders a deeper elaboration, because, among other things, the simultaneous processing of verbal and pictorial information demands too many cognitive resources (Lowe, 2003, 2004). However, as in Experiment 1 and 2, no differences in dependent measures across presentation modes were found in the current experiment. This stable finding throughout three experiments with different on- and offline measures leads one to the conclusion that the pictorial information in films did not hinder elaboration as claimed by e.g., Lowe, Salomon, Weidenmann or foster elaboration as claimed by Mayer and Paivio. In line with this reasoning, no differences in cognitive load measures across presentation modes were found in the current experiment or in Experiment 1 and 2.

Implicit versus explicit information

Experiment 2 showed that recipients generate LCBIs online in expository film comprehension. To examine whether the online generated LCBIs are integrated into a coherent mental model, offline comprehension test items were divided into implicit and explicit items. Implicit items asked questions that were only implicitly stated in the experimental material and had thus to be inferred by the understander, whereas explicit items asked about information that was explicitly stated in the experimental material.

The question that arises is whether understanders were able to integrate the online inferred LCBIs into a coherent mental model by integrating the new information into long-term memory structures, as measured by implicit comprehension scores, or were understanders only able to integrate explicitly stated information into long-term memory structures, as assessed by explicit comprehension scores? To answer these questions, three analyses were conducted: First, comprehension scores on implicit items (only applicable to participants in the experimental condition) were compared to the comprehension scores of participants in the control condition, who did not obtain any experimental material. Results indicated that participants in the experimental condition outperformed participants in the control condition. This finding also held true for comprehension scores on explicit items.

In other words, participants were able to integrate the explicitly stated information into a coherent mental representation. Additionally, participants were able to infer the necessary information to answer even implicit items correctly and outperformed those participants, who did not obtain any experimental material. This led to the assumption that participants are able to integrate the online inferred information into a coherent mental model and that, in terms of learning from the material, it does not matter whether the information is stated implicitly or explicitly in the experimental material, because understanders generate inferences to integrate the implicitly stated information into a mental model. However, this assumption seems very optimistic when compared to findings from expository text comprehension research, where often no online causal inference generation was found at all.

To test this assumption, a within comparison between the comprehension scores on implicit and explicit items was conducted. If implicit items were answered as correctly as explicit items and, if no difference between the comprehension scores across item types occurred, then understanders integrated both types of information into a coherent mental model. However, if a difference in comprehension scores across item types occurs and, if this difference is in favor of the explicit items, then understanders were indeed able to infer the implicit information online, but they were not able to integrate the new information into a coherent mental model by integrating the new information with long-term memory structures.

The results showed no differences in comprehension scores across item types, but an interaction between item type and word set was found. Concerning the interaction word set x item type, it was found that prerequisite knowledge mediated this relation. In other words, the factor "word set", which had been introduced due to methodological reasons, only had an influence on the item type, because prerequisite knowledge mediated the relation. Only participants with high(er) prerequisite knowledge were able to integrate the inferred information into a mental model as coherently as they had the explicit information. Participants with low(er) prerequisite knowledge were indeed able to infer the implicit information online to enable an online understanding as shown in Experiment 2, where no effect of prerequisite knowledge was found, but they were not able to integrate the new information with existing long-term memory structures, as they had with explicit information. It seems that existing long-term memory structures, such as prerequisite knowledge structures, are necessary to integrate the inferred information into a coherent mental model that lasts longer than the presentation of the material.

However, prerequisite knowledge had no influence on the explicit comprehension scores. In other words, prerequisite knowledge structures are necessary to integrate inferred information into a coherent mental model, but not to build a mental model from explicit information. Thus, understanders are able to integrate explicitly stated information into long-term memory structures, unlike implicitly stated information. For implicitly stated information, prerequisite knowledge stored in long-term memory is essential. This is in line with findings from inference research in narrative (Graesser et al., 1994; 1997) and expository texts (e.g., Wiley & Myers, 2003) that always emphasize the dominant role of sufficient prior or prerequisite knowledge for inference generation.

Summary and conclusions

Films are suitable instructional tools. Participants in two experimental conditions, who either watched expository films or listened to expository audio tapes, demonstrated superior offline comprehension scores than participants in a control condition, who did not receive any experimental material. If the information was explicitly stated in the experimental material, participants were able to integrate the new information into a coherent mental model. However, if the information was implicitly stated in the experimental material and, therefore, had to be inferred by the understanders, only understanders with high prerequisite knowledge were able to integrate the online inferred information into a coherent mental model. Regarding the role of online local causal bridging inferences (LCBIs) in expository film comprehension, it can be concluded from the current experiment that LCBIs are necessary for online comprehension, but are not sufficient for a coherent mental model generation. For the generation of a coherent mental model, existing long-term memory structures, such as prerequisite knowledge, seem crucial for the integration of online inferred causal information.

The current set of experiments is one of the first to systematically analyze the interplay of online bridging inferences and offline comprehension scores. As the current experiments demonstrated, this seems to be a promising approach to analyze the complex nature of comprehension processes.

11 General Discussion

Three experiments analyzed whether expository films are elaborated deeply and can, therefore, be considered to be effective instructional tools. A crucial form of elaboration that is necessary for comprehension is that of online generated bridging inferences, because bridging inferences establish coherence on a semantic level. The current experiments tested whether bridging inferences are generated online in expository film comprehension and how bridging inferences relate to offline comprehension scores.

The central finding was that recipients are able to generate local causal bridging inferences in expository film and audio comprehension (Experiment 2), as indicated by shorter naming latencies in an inference-related context compared to an inference-unrelated context. Compared to a control group, which did not receive any experimental material, participants in both experimental groups (video and audio) also showed superior performances in an offline comprehension test (Experiment 3). Experiment 1 revealed that it is extremely important in online inference generation research to implement the right method and to consider the automatic generation of bridging inferences, even for global bridging inferences. As the key press applied in this experiment was not a valid indicator for online inferences, the more unobtrusive naming task was implemented in Experiment 2.

Stable throughout all three experiments, recipients of expository films, as well as of expository audio tapes, did not differ in online measures of inference generation as well as offline measures of comprehension. Therefore, it is concluded that the dynamic pictorial information in films did neither hinder nor foster inference and comprehension processes, as claimed by several researchers. Films are elaborated as deeply as single media like audio tapes and thus seem to be equally suitable as instructional tools in schools or elsewhere.

This chapter discusses the current findings with respect to the broader subject realm of cognitive, educational and media psychology. Based on this discussion, promising directions for further research will be pointed out whenever possible. The chapter is divided into a) *theoretical considerations* that discuss the findings from a theoretical perspective, b) *methodological considerations* that discuss the findings from a methodological perspective, and lastly c) *practical educational considerations* that discuss the practical educational implications for the current findings.

11.1 Theoretical considerations

Are expository films elaborated deeply or does the dynamic pictorial information hinder elaboration?

Film viewers and audio tape listeners generated online local causal bridging inferences, a type of inference that has been found crucial for understanding. Moreover, both film viewers and audio tape listeners generated a mental model of the learning content. From these considerations, it can be concluded that expository films are elaborated deeply.

As shown by all three experiments, no differences in online and offline comprehension processes across film viewers and audio tape listeners were found. Therefore, the assumption that the pictorial information hinders elaboration cannot be supported by the current set of experiments. It seems that understanders are able to concentrate on the verbal information and not let themselves be distracted by seductive details (Harp & Mayer, 1998). Moreover, the results suggest that film viewers do not adopt a passive or ineffective reception attitude in response to the entertaining character of the pictorial information or the overwhelming character of the additional pictorial information presented. Thus, the common assumption that films are elaborated superficially (DeFleur et al., 1992; Salomon, 1984; Weidenmann, 1989, 2002) cannot be supported by the current findings.

However, the results obtained may be at least partially artifacts of the codalities the performance measures were based on. In particular, this thesis applied on- and offline measures based on verbal propositional analyses, verbal naming stimuli and verbal comprehension test items. However, film is a medium that combines verbal *and* pictorial information. As audio tapes just rely on verbal codes to transport the information, it is reasonable to assume that participants in the video group would outperform participants in the audio group on pictorial stimuli in naming tasks or pictorial comprehension test items, at least for items that address spatial relations as found by Stone and Glock (1981) for static pictures. Examples of pictorial comprehension test items could be drawing sketches or detecting errors in drawings. Examples of online naming stimuli could be simply naming pictorial objects as applied by Unsöld and Nieding (in press). However, pictorial online stimuli seem hard to implement for automatic bridging inferences in expository films, because the inferred information is often abstract and not easy to depict.

To further consider the impact of the films' pictorial trace, the pictorial trace could be analyzed analogous to the propositional analysis for verbal information in texts. However, one important distinction needs to be made: While language is often assumed to be mentally represented in propositions (Kintsch, 1988, 1998; Lorch Jr. & van den Broek, 1997) some researchers assume that pictorial information is mentally represented analogous to the external representation, in other words, *not* in propositions (e.g., Kosslyn, 1980, 1994). In line with this reasoning, some authors claim that even situation models based on language are not mentally represented in propositions, but analogous to the external representations (Barsalou, 1999; Zwaan, 1999, 2004). Concerning film, this first implies that it is unclear whether the verbal and pictorial information are represented in a similar propositional format or in different representational systems. Second, it might be argued that for films even the verbal information conveyed in the audio trace would be augmented by perceptual visual elements in the sense of perceptual symbols. It is beyond the scope of this dissertation to go into any further detail regarding the perceptual symbol approach, particularly as it is not the currently underlying theoretical fundament. However, an integrative approach in future studies should consider a non-propositional approach, especially for pictorial information. The integrative model for text-picture processing (Schnotz, 2005) might be a good first theoretical approach, albeit for static text-picture combinations.

Are global bridging inferences generated online in expository film comprehension?

To measure global bridging inferences in expository film comprehension, recipients were instructed to press a key whenever they were spontaneously thinking back to a previously viewed position or scene. The key press was interpreted as an indicator for a global bridging inference. However, the question as to whether *global bridging inferences* are generated online in expository film comprehension could not be answered satisfactorily, because no systematic correlations between the key press and a) a priori defined bridging inference positions and b) an offline measure of comprehension scores were found. Therewith, it was concluded that either no global bridging inferences are generated online or that the Newtson paradigm failed as a valid method for inference detection. The latter assumption seemed to be plausible, because bridging inferences are assumed to be generated automatically without the understander's awareness (e.g., Graesser et al., 2007), whereas the key press required some monitoring of one's own cognitive processes.

Thus, to answer the question of whether global bridging inferences are generated online in expository film comprehension, empirical studies are necessary that implement an unobtrusive method, not requiring the understander's awareness of his own cognitive processes. The naming task that proved to be a valid indicator for local causal bridging inference generation in expository film comprehension (Experiment 2) seems to be a promising approach to test whether global bridging inferences are generated online in expository film comprehension. However, it remains difficult to measure global bridging inferences, because for the same inference-triggering position there are often several semantic backward connections conceivable, making it hard to assess the precise global bridging inference that the understander is generating. One option of overcoming this drawback would be to implement material that only allows one specific global connection at a time.

Differential cognitive and motivational aspects

Even though working memory did not reveal a significant relation to online inference generation in Experiment 2, several findings in the context of text comprehension suggest that understanders with greater working memory capacity are better able to generate online inferences (Calvo, 2004; Linderholm, 2002), because they are better able to integrate information from several sources (generate bridging inferences) than understanders with a low working memory capacity. Nevertheless, in Experiment 2, no such correlations were found. One reason could be that working memory capacity only affects global bridging inference generation, because the information that has to be connected is spatially and temporally further apart in the discourse itself and thus needs to be reactivated in working memory during the understander's processing of the discourse.

Another reason for the lack of correlations found could be that a *reading span task* was applied in Experiment 2. Even if the reading span serves as a general measure for working memory span (Hacker et al. 1998; Daneman & Carpenter, 1980; Just & Carpenter, 1992), it could be that the aspects of working memory measured with the reading span task were not predictive for the understanding of the filmic and/or auditory content presented. A variation for future experiments could be to implement other working memory measures more relevant to film and audio understanding like an auditory or listening span test (e.g., Daneman & Carpenter, 1980).

In addition to working memory capacity, another cognitive variable plays an important role for inference generation and comprehension processes, namely, prior knowledge. McNamara et al. (1996) found that low coherent expository texts were only beneficial for high prior knowledge learners, whereas high coherent expository texts were only beneficial for low prior knowledge learners. It seems to be important to adjust the experimental material according to the learners' cognitive prerequisites to enable an active processing. In line with this reasoning, McNamara et al. found that learning outcomes for high prior knowledge learners decreased if the text was too coherent, because their active text processing was impaired. As Experiment 3 revealed, only learners with sufficient prerequisite knowledge were able to integrate the implicit information as coherently as the explicit information into a mental model. Thus, to enable a truly coherent mental model for low prior knowledge learners, more information should have been stated explicitly in the experimental material. This conclusion is in line with findings from previous research who found that directly increasing the coherence of a text increased the learners' recall and comprehension performances (Beck et al., 1991; Beyer, 1990; Britton & Gülgoz, 1991).

The beneficial effect of addressing the learners' cognitive prerequisites has been shown by Schwan and Riempp (2004). Recipients of interactive videos that directly address the recipients' cognitive prerequisites, e.g., different degrees of prior knowledge, with such interactive features as stop, rewind, and fast forward, demonstrated better learning outcomes than recipients of regular non-interactive videos. An explanation for this finding is that interactive videos allow for a better adaptation to the learners' prerequisites and hence enable active processing like inference generation. Thus, the implementation of interactive features in expository films seems to foster inferences and, thereby, might enhance comprehension and learning processes.

Beyond cognitive variables, such as prior knowledge or working memory span, it is assumed that motivational variables, such as interest, motivation and enjoyment, play a significant role in film comprehension (Weidenmann, 1989). It seems reasonable to consider whether motivational variables have an impact on inference generation and comprehension processes in expository and narrative film comprehension. Thus far, these variables have rarely been analyzed with regards to inferences in text comprehension, but it seems plausible that these variables might have an effect on inference generation, particularly in (narrative) film comprehension. For instance, if a recipient is very interested and involved (Vorderer, 1992; Wirth, 2006) in a filmic presentation, his elaboration and his desire to understand the content might be enhanced such that he invests more mental effort, elaborates the content more deeply, and generates more inferences. The concept of involvement, however, refers mainly to narrative media.

Relations between online and offline measures

This dissertation is the first set of experiments to systematically analyze the relations between online inference processes and offline comprehension scores within the same setting and material. The results revealed that all participants had sufficient prerequisite knowledge to generate online local causal bridging inferences (Experiment 2). Compared to a control group that did not receive any experimental material, participants in both experimental groups (video and audio) also exhibited superior performances in an offline comprehension test (Experiment 3). When comparing comprehension scores on implicit versus explicit items among the experimental groups, it was demonstrated that only participants with high prior knowledge were able to integrate the implicit information as coherently as the explicit information into a mental model.

These findings clarify the role that online inferences play in the learning process. It seems as if local causal bridging inferences are generated online to construct the textbase and to enable online comprehension. At this level, prior knowledge plays a less significant role compared to the generation of the situation model (Kintsch, 1988, 1998), which requires integrating explicitly stated and inferred information with long-term memory structures, such as prior knowledge, into a coherent mental model.

To further analyze the interplay between online generated inferences and offline comprehension scores and to test the sustainability of the mental model, the current design could be extended by the delay of the comprehension test for, e.g., a couple of days or several weeks, resulting in a 2 (presentation mode: video vs. audio) x 2 (delay of comprehension test: immediately vs. delayed) x 2 (item type: implicit vs. explicit) design. A possible hypothesis regarding the delay would be that only participants with high prior knowledge are able to sustain their mental model over time, based on long-term memory structures, such as prior knowledge, explicit *and* implicit information, because their prior knowledge structures allow a comprehensive, sustainable integration of the explicit and implicit information with existing prior knowledge structures. Participants with low prior knowledge, by contrast, might not be able to sustain either implicit only or implicit *and* explicit information over time, as a lack of prior knowledge structures prevents a coherent integration of the new information with long-term memory structures.

This thesis revealed that systematically analyzing on- and offline measures in combination seems to be a promising approach that should be pursued further in order to scrutinize the complexity of comprehension and learning processes. The goal of future research should be to provide an innovative, comprehensive theoretical framework that combines on- and offline processes in the broader subject realm of cognitive, educational, and media psychology. Moreover, existing theories in the context of cognitive media psychology should be extended by explicitly addressing process variables like inferences. For example, the Cognitive Theory of Multimedia Learning (Mayer, 2001, 2005) just names selection, organization and integration as processing variables that could be further specified by the inclusion of specific cognitive processes, such as inferences.

Generalizations of the present findings

It seems as if the current findings are of interest beyond the scientific community addressing discourse comprehension. For instance, the current findings are of broader interest for multimedia learning, especially for learning with dynamic visualizations, such as animations. It seems as if the aspect of analyzing online inference processes is of interest for basically every instructional medium. Most learning theories and instructional experiments so far focus on learning outcomes variables and neglect process variables like inference generation. Additionally integrating measures addressing these online processes might substantially enhance the repertoire of methods to account for the underlying cognitive processes.

As previously discussed in this dissertation, the current findings might be material-dependent. Therefore, it seems important to try to replicate the current findings with a different material and perhaps different content domains to test how stable the current findings are.

11.2 Methodological considerations

As stated by several authors (e.g., Graesser et al., 1997; Potts et al., 1988) and revealed throughout this thesis, the methods that measure online inferences are an important factor in the research design. For instance, inferences can be revealed by one method, but remain undetected by the other. Depending on the method, different conclusions are drawn. This became evident in Experiment 1, but also in earlier research. Potts et al. (1988) drew two different conclusions regarding the online generation of elaborative inferences by implementing two different methods. By implementing a lexical decision task, they found that elaborative inferences are generated online. However, by implementing a naming task, the authors found that elaborative inferences are *not* generated online. Since the naming task was considered to be the more unobtrusive, but still sensitive enough method, the authors finally concluded that elaborative inferences are not generated online.

This example illustrates how the method impacts the obtained results and thus the conclusions drawn. Regarding Experiment 1 of this dissertation, the results could have either indicated a) that global bridging inferences are not generated online or b) that the Newtson paradigm is not a valid indicator for their detection. The latter assumption seemed to be more plausible, because empirical results in text comprehension have shown that global bridging inferences are generated online (e.g., Albrecht & O'Brien, 1993; Myers et al., 1994; Tapiero & Otero, 2002). Furthermore, bridging inferences are assumed to be generated automatically, without the understander's awareness (e.g., Graesser et al., 2007), whereas the key press requires some monitoring of one's own cognitive processes.

These two examples illustrate how crucial the applied method in inference research truly is, but also how difficult to develop (Graesser et al., 1994; Keenan et al., 1990). In text comprehension research, inference methods have been considered "stumbling blocks" (Keenan et al., 1990; p. 295), but, in film comprehension, it seems even harder to implement the most appropriate online method due to specific filmic characteristics, such as the simultaneous presentation of verbal and pictorial information, the dynamic presentation of the content, and the lack of interactive possibilities. The naming task applied in Experiment 2 proved to be a successful method for detecting online bridging inferences. Fortunately, traditional text comprehension research has shown where assets and drawbacks of the naming task lie.

A different methodological possibility for film comprehension, which has not yet been discussed, is that of making experimental films oneself to implement the manipulations as precisely as possible. As texts are easier to create, texts have always been created by the researchers themselves to allow for a precise manipulation. The difficulty of creating a precise manipulation in film comprehension research was exemplified by Experiment 2. In this experiment, a true cross-mapping of the naming words was hardly possible, because some words in proximity to the LCBI position were either already mentioned or too similar to present the naming words in their unrelated context. Although this was not a drawback for the current experiment, this potentially obtrusive factor can be avoided by self-made films.

Further methodological considerations include thinking about even more unobtrusive methods to avoid interrupting the film. One possibility could be dual task measures to measure mental load during film comprehension (Della Sala, Baddeley, Papagno, & Spinnler, 2004; Logie, Maylor, Della Sala, & Smith, 2004). If understanders demonstrate high mental load measures at predefined inference positions, this could be an indicator for an inference generation. However, a suitable dual task is hard to find. Since the film's audio trace might be inconsistent regarding loudness, tone, etc., a pure auditory tone is not a valid indicator, as the film's auditory inconsistencies might influence the understanders' detection of the auditory tone. Thus, an unequivocal interpretation of the responses obtained to an auditory tone implemented at different positions may be exacerbated. The same applies to a visual dual task, because the film's visual trace is probably even less consistent than the audio trace. Lightness, brightness, etc., might change over the course of the film and influence the dual task performances. Therefore, a suitable dual task that considers these aspects and that is indicative for inference processes must be developed first. If that occurred, the dual task paradigm would appear to be a promising approach for measuring online inferences in film comprehension.

Another unobtrusive method is eye tracking. Even if stated in Chapter 6 that eye tracking is not a suitable method to indicate inference processes in film comprehension, because a) analyses of dynamic material are difficult and time-consuming and b) interpretational problems regarding the eye-tracking parameters exist, the implementation of eye-tracking measures as a *very* unobtrusive method should be raised again. The analyses of eye-tracking parameters in dynamic visualizations, such as films, will be improved in the future. Automatic computational possibilities for an online definition of dynamic areas of interest (Rötting, 2001) are under development (Fischer, Papenmeier, & Huff, 2007). This would reduce the effort and possible inaccuracy regarding the analyses of eye-tracking parameters in dynamic stimulus material, such as films. Recently, different authors scrutinized how visual attention relates to auditorily presented information by studying eye movements during listening while simultaneously being presented with either static pictorial stimuli (Boland, 2005; Yee & Sedivy, 2006) or a blank screen (Altmann, 2004; Spivey, Richardson, Tyler, & Young, 2000), Recording eye movements to draw conclusions about the online cognitive processes that connect verbal and pictorial information seems to be an innovative approach that could possibly be transferred to measure online inference generation in film comprehension. However, the interpretational problems associated with eye-tracking measures might remain.

Other methodological trends include neurological images and fMRI (Beeman, Bowden, & Gernsbacher; 2000; Ferstl, 2007; Griesel, Friese, & Schmalhofer, 2003), EEG (van Berkum, Brown, Zwitserlood, Kooijman, Hagoort, 2005) and computational models (Goldman, Golden, & van den Broek, 2007). These methods have been successfully applied to measure and model inference processes in text comprehension and could be adapted to measure inference processes in film comprehension, as well.

11.3 Practical educational implications

Practical educational considerations refer to the questions of how an online inference generation can be encouraged and thus comprehension outcomes improved.

An interesting idea derived from Experiment 2 refers to the online encouragement of inferences. It sounds plausible that presenting the inference-presenting naming word at the text position where learners should draw this inference could help learners (e.g., with low prior knowledge) to draw the required inference.

This approach is in line with an area of research in text comprehension, where learning from texts is supposed to be facilitated by computer-based auto-tutor systems that foster online reading strategies and thus online inferences (e.g., iSTART developed by McNamara, Levinstein, & Boonthum, 2004). In iSTART, an animated conversational agent gives prompts, e.g., to link two sentences in a text causally (causal bridging inference). These systems have shown promising results in tests of learning gains and learning strategies in text comprehension (Graesser, McNamara, & VanLehn, 2005).

An inference training for children was developed by McGee and Johnson (2003). The authors examined whether inference training affected skilled and less skilled comprehenders by instructing six- to nine-year-old children in how to make inferences from -- and generate questions about -- a text. The authors report that the less skilled group improved more than the skilled group of children. Hence, inference trainings could be systematically implemented to encourage inference and learning processes, especially for less-skilled students. It seems conceivable that these findings also transfer to expository film comprehension. Thus, inference trainings and auto-tutor systems might improve inference generation and thus comprehension and learning processes in expository film comprehension, as well.

Are inferences generated voluntarily?

An important role regarding the efficiency of films as instructional tools could be played by the instruction to learn. As found in the context of expository text comprehension, the instruction to learn from the text improved inference generation (e.g., Noordman et al., 1992). In all three experiments outlined in this dissertation, participants were instructed to learn from the film and to follow the film attentively. Nevertheless, in Experiment 1, participants were instructed to “memorize” the presented content. This was probably not an especially encouraging instruction for inference generation. Therefore, in Experiments 2 and 3, participants were instructed to “understand” the content. Even though it cannot be concluded from the current set of experiments that the particular instruction in Experiment 1 was responsible for the unsystematic results regarding inference generation, it seems reasonable that the instruction to “understand” in Experiments 2 and 3 encouraged more online bridging inferences than the “memorize” instruction in Experiment 1. Bridging inferences are necessary to comprehend the content and not to memorize the content (Kintsch, 1998) and are thus probably more activated by an instruction that stresses the role of comprehension compared to an instruction that stresses the role of retention.

Conclusions

As stated in the introduction, films are more frequently used as instructional tools in schools than a variety of other media (Feierabend & Klingler, 2003). However, the effectiveness of films as effective learning tools has often been questioned, because it was assumed that films are not elaborated deeply (DeFLeur et al., 1992; Salomon, 1984; Weidenmann, 1989, 2002). This dissertation demonstrated that this assumption cannot be supported when directly looking at elaboration processes like causal bridging inference generation in expository film comprehension. Consequently, expository films can be implemented in schools and on TV to successfully communicate knowledge and information, as long as the factors crucial for an inference generation in expository material (e.g., an instruction or a reading goal that encourages the comprehension processes, the information necessary to compute the inferences is easily available, and the understander has at least some prior knowledge to relate to the new information) are considered.

12 Summary

This dissertation is motivated by the widespread assumption that films are processed in a superficial manner and that they are, therefore, not particularly suitable for instructional purposes. For instance, it has been claimed that the dynamic-pictorial information, a core characteristic of film, rather hinders than supports deep cognitive processing. However, it has to be noted that the supposed superficial elaboration of film yet needs to be tested in a direct way by means of *process measures*. Existing research findings mostly rely on analyses of *outcome measures* that are obtained subsequently to the film reception. Typically, these outcome measures are then used in a second step, to draw post-hoc conclusions regarding cognitive processes that took place during film reception. This indirect approach may, however, easily suffer from misinterpretations and is, moreover, not suited to analyze the specific cognitive processes during film comprehension in greater detail. Thus, the strong focus of the existing research on analyzing outcome measures can be seen as a pivotal reason for the yet ambiguous pattern of results with regard to cognitive processes involved in film reception.

Based on this state of research, the current dissertation aims at directly investigating elaboration processes in expository film (i.e., films that target at conveying knowledge) reception instead of only analyzing outcome measures that are obtained post-hoc. As prototypical elaborations inferences are addressed that can be broadly characterized as information that is not conveyed explicitly but needs to be generated by the recipient himself in order to understand the information presented. A specific type of inference that is crucial for comprehension are bridging inferences because bridging inferences construct semantic relations between different sections of the information presented. In analogy to the distinction between local and global coherence, a distinction is made between local and global bridging inferences depending on the spatial distance between the sections that are related.

The generation of local and global bridging inferences during expository film reception can be seen as an important indicator for a deep cognitive processing of film. Therefore, this dissertation scrutinizes to what extent local and global bridging inferences are generated during expository film reception. Moreover, the effect of the dynamic-pictorial information on the generation of bridging inferences will be investigated by means of a film/audio comparison. This comparison is of particular interest as often the dynamic-pictorial information inherent to film is made responsible for a possibly superficial cognitive processing. Answering these two research questions based on suitable process measures can provide important hints as to what extent expository films are elaborated deeply. Thus, in the three experiments reported in this dissertation bridging inferences are analyzed online, i.e., during film reception, (Experiment 1 and 2) as well as offline, i.e., subsequently to film reception, (Experiment 1 and 3).

Experiment 1 uses an expository film on German post-war history to analyze the role of global bridging inferences during film comprehension. The generation of global bridging inferences was registered by means of process measures (online) as well as by means of outcome measures (offline). To derive

indicators for the online generation of bridging inferences the Newtson paradigm was applied. Up to now, this paradigm was predominantly used in fields like cognitive film psychology or event cognition to reveal specific cognitive processes during film reception.

The participants in Experiment 1 had the task to watch a film and to indicate with a key press whenever they spontaneously thought back to a previously seen film section (i.e., when they generated a global bridging inference). It was hypothesized that a systematic relationship would show between the key presses (and the global bridging inferences indicated by them) and a priori defined global bridging inference positions. Additionally, it was expected that the degree of fit between key presses and a priori defined bridging inference positions would allow predicting offline comprehension performance. This expectation was derived from the assumption that generating global bridging inferences would increase the global coherence of the presented information, finally resulting in a more coherent mental model.

However, these predictions could not be confirmed in Experiment 1. Different explanations could account for the fact that the results contradicted the hypotheses. First, it might have been that the recipients did not generate global bridging inferences online. However, this seems rather improbable in the light of the offline comprehension data obtained in Experiment 1. Second, it might have been that the assumed relationship involving the degree of fit between key presses and a priori defined bridging inference positions on the one hand and comprehension performance on the other hand does not exist. However, this would be contradictory to established models of text comprehension. A third explanatory approach questions the assumption that the Newtson paradigm is a suitable method to measure online generated global bridging inferences. To support this explanation one could point to the finding that local bridging inferences are generated automatically. Thus, their conscious accessibility is limited. Up to now, it has been unclear whether this holds also true for global bridging inferences in film comprehension. If it does then an approach like the Newtson paradigm that presupposes a substantial amount of conscious accessibility of cognitive processes might not be a valid measure. Based on the results of Experiment 1, however, it can be concluded that global bridging inferences, too, are generated rather automatically and unconsciously so that methods that depending on conscious accessibility might be limited regarding the detection of these inferences.

Based on these results, a different methodological approach was chosen in Experiment 2 to answer the question whether expository films are elaborated deeply. First, *local* bridging inferences were addressed instead of *global* bridging inferences. Local bridging inferences allow identifying more precisely which pieces of information are connected by the bridging inference, thus enabling are more unambiguous analysis. Second, the Newtson paradigm was replaced by the so-called naming paradigm, a method used in text comprehension research, which is less intrusive and depends less on conscious accessibility. In text comprehension research the naming paradigm has been established as a valid method to measure automatic inference processes. Thus, it can be assumed that this method can also be applied to measure automatic bridging inferences in film comprehension. Third, Experiment 2 focuses on *causal* bridging inferences, which are a specific type of local bridging inferences. Causal bridging inferences are pivotal for a substantial knowledge acquisition. Moreover, they are characterized by the fact that they can be defined very precisely. The understanding of causal relationships is particularly important for scientific domains.

In Experiment 2, the methodological modifications compared to Experiment 1 were implemented by studying expository film comprehension in scientific domains (physics/chemistry) by means of the naming paradigm. The basic idea of the naming paradigm is that the generation of inferences can be detected by measuring naming latencies for words that represent the generated inferences. These latencies will be shorter in a context that requires the generation of the inference compared to a control context that does not require the generation of the respective inference.

The current findings support the hypothesis that local causal bridging inferences are generated online during expository film comprehension. Naming latencies for words presented in an inference-related context were shorter than naming latencies for the same words presented in an inference-unrelated context. This was true for the experimental condition watching films as well as for the experimental condition that merely received the respective audio information to test for the effect of dynamic-pictorial information on the generation of local causal bridging inferences. Based on this pattern of results, both the assumption that films are not elaborated deeply as well as the assumption that their dynamic-pictorial information hinders elaboration can be refuted.

To furthermore explore the relation between online generated local causal bridging inferences on the one hand and the construction of a coherent mental model on the other hand, a third experiment was conducted. In Experiment 3 the same materials as in Experiment 2 were used to investigate whether the online comprehension of causal relationships demonstrated in Experiment 2 results in the generation of a coherent mental model.

To answer this question, an offline comprehension test consisting of multiple-choice questions referring to the local causal bridging inferences that, according to Experiment 2, are inferred online, were administered in Experiment 3 subsequently to the reception phase. As first independent variable it was manipulated whether the material was presented as film or as an audio tape. Additionally, it was manipulated within-subjects whether the materials presented to the participants contained the critical causal information only implicitly as in Experiment 2 or whether this information was explicitly stated. Performance on answering the multiple-choice items addressing the critical causal information was compared to a control group that did not receive any experimental materials. The comparison revealed that both participants of the film and the audio condition yielded better comprehension performances than participants of the control condition. Moreover, it could be demonstrated that participants in the experimental conditions outperformed participants in the control group not only with regards to explicitly stated causal information but also with regards to implicitly stated causal information. Thus, it can be concluded that participants in the experimental condition generated a rich mental model by integrating both implicit and explicit causal information with long-term memory structures. However, a within comparison of participants' implicit versus explicit comprehension performances for the experimental conditions revealed that only participants with high domain specific prior knowledge were able to integrate the implicit information as coherently into a mental model as the explicit information.

Thus, it can be concluded that domain specific prior knowledge is necessary to enable understanders to integrate implicit information - that is verifiably inferred online (Experiment 2) - successfully into a coherent mental model (Experiment 3). This holds true for both film and audio comprehension processes.

In sum, the assumption that films are only superficially elaborated can be refuted based on the pattern of results obtained in the three current experiments. It could not only be demonstrated that local causal bridging inferences are generated online, but also that the cognitive processing of expository films results in a deeper offline comprehension. Furthermore, it could be shown in all three experiments that the dynamic-pictorial information that characterizes film has no effect on both online and offline comprehension processes.

This dissertation provides a first set of experiments that investigate elaboration processes in expository film comprehension by means of direct online methods and by means of a systematic relation of online comprehension processes to offline comprehension processes. Thus, this work presents an innovative and promising approach for the analysis of inference and learning processes at the interface of cognitive, educational, and media psychology that deserves further refinement in future research.

Zusammenfassung

Ausgangspunkt dieser Dissertation ist die weit verbreitete Annahme, dass Filme vergleichsweise oberflächlich verarbeitet werden und somit zur Wissensvermittlung ungeeignet sind. So wird z.B. angenommen, dass die für Filme charakteristische dynamisch-piktoriale Information eher hinderlich als förderlich für eine tiefe kognitive Verarbeitung ist. Allerdings steht die Annahme einer oberflächlichen Verarbeitung von Filmen in einem gewissen Widerspruch zu der Tatsache, dass eine direkte empirische Überprüfung dieser Annahme anhand von *Prozessvariablen* bislang noch aussteht. Bisherige Forschungsergebnisse beruhen überwiegend auf *Ergebnisvariablen*, die im Anschluss an die Filmrezeption erhoben wurden. Aus diesen Ergebnisvariablen werden üblicherweise in einem Folgeschritt Rückschlüsse auf zugrunde liegende Rezeptionsprozesse während der Filmverarbeitung gezogen. Diese indirekte Vorgehensweise ist jedoch anfällig für Fehlinterpretationen und erscheint überdies nicht geeignet, um spezifische Verarbeitungsprozesse während des Filmverstehens detaillierter zu untersuchen. Die Beschränkung der bisherigen Forschung auf eine ergebnisorientierte Vorgehensweise kann als ein zentraler Grund dafür angesehen werden, dass bislang noch kein eindeutiges Ergebnismuster zur Verarbeitung von Filmen gewonnen werden konnte.

Aufbauend auf diesem Forschungsstand zielt die vorliegende Dissertation darauf ab, erstmals nicht nur Ergebnisvariablen zu erfassen, sondern auch auf direkte Weise Elaborationsprozesse bei der Verarbeitung von Filmen zu untersuchen. Als prototypische Elaborationen werden dabei Inferenzen adressiert, die im weitesten Sinne als Information charakterisiert werden können, die nicht explizit vermittelt wird, sondern die durch den Rezipienten selber generiert werden muss, um die dargebotene Information zu verstehen. Die vorliegende Arbeit beschäftigt sich mit expositorischen Filmen, d.h. mit Filmen, die dem Ziel der Wissensvermittlung dienen. Eine Inferenzform, die von entscheidender Wichtigkeit für den Aufbau einer semantisch kohärenten mentalen Repräsentation während der Verarbeitung ist und die damit in hohem Ausmaß das Verständnis bestimmt, wird als Brückeninferenz bezeichnet. Brückeninferenzen stellen Bezüge zwischen einzelnen Abschnitten der dargebotenen Informationen her. Je nach räumlicher Distanz dieser Abschnitte wird - analog zu der Unterscheidung von lokaler und globaler Kohärenz - zwischen lokalen und globalen Brückeninferenzen unterschieden.

Die Bildung von lokalen und globalen Brückeninferenzen bei der Rezeption expositorischer Filme kann als wichtiger Indikator für eine tiefe kognitive Filmverarbeitung betrachtet werden. Aus diesem Grund bezieht sich eine spezifische Forschungsfrage dieser Dissertation darauf, in welchem Ausmaß bei der Rezeption expositorischer Filme lokale und globale Brückeninferenzen gebildet werden. Als zweite spezifische Frage adressiert diese Dissertation auf der Basis eines Film/Audio Vergleichs den Effekt der dynamisch-piktorialen Information auf die Bildung von Brückeninferenzen, da oft vor allem die dynamisch-piktoriale Information, die dem Medium Film inhärent ist, als Ursache einer möglichen oberflächlichen Verarbeitung angesehen wird. Die Beantwortung dieser beiden Fragen auf der Basis geeigneter Prozess-und Ergebnisvariablen kann wichtige Hinweise darauf geben, ob und in welchem Umfang expositorische Filme elaboriert verarbeitet werden. In den in dieser Dissertation berichteten Experimenten werden Brückeninferenzen daher einerseits online, d.h. während der Rezeption erfasst (Experiment 1 und 2) und andererseits offline, d.h. im Anschluss an die Rezeption analysiert (Experiment 1 und 3).

In Experiment 1 wurde anhand eines expositorischen Films zur deutschen Nachkriegsgeschichte untersucht, welche Rolle globale Brückeninferenzen bei der Filmrezeption spielen. Die Bildung globaler Brückeninferenzen wurde sowohl durch Prozessvariablen (online-Messung) als auch durch Ergebnisvariablen (offline-Messung) erfasst. Um Indikatoren für die online-Generierung von Brückeninferenzen zu gewinnen, wurde das Newtson Paradigma verwendet. Dieses Paradigma wurde bislang vor allem in Forschungsarbeiten zur kognitiven Filmpsychologie bzw. zu Ereigniskognitionen erfolgreich eingesetzt, um spezifische kognitive Prozesse während der Filmverarbeitung zu untersuchen.

Die Aufgabe der Teilnehmer in Experiment 1 bestand darin, bei der Betrachtung des Films immer dann mit einem Tastendruck zu reagieren, wenn sie spontan an eine vorangegangene Filmpassage zurückdenken mussten (d.h., wenn sie eine globale Brückeninferenz generierten). Dabei wurde von der Hypothese ausgegangen, dass systematische Beziehungen zwischen den Tastendrücken (und den durch sie indizierten globalen Brückeninferenzen) und a priori definierten globalen Brückeninferenzpositionen bestehen. Des weiteren wurde angenommen, dass sich aus dem Ausmaß der Passung zwischen Tastendrücken und a priori definierten Brückeninferenzpositionen offline-erfasste Verstehensleistungen vorhersagen lassen, da die Generierung globaler Brückeninferenzen die globale Kohärenz der präsentierten Information erhöht und damit letztlich in einem kohärenteren mentalen Modell resultieren sollte.

Diese Erwartungen konnten jedoch in Experiment 1 nicht bestätigt werden. Für diesen hypothesenkonträren Befund sind verschiedene Erklärungsansätze denkbar. Erstens könnten die Rezipienten keine globalen Brückeninferenzen online generiert haben, was jedoch aufgrund der offline-Verständnisdaten eher unwahrscheinlich erscheint. Zweitens besteht die Möglichkeit, dass der vermutete Zusammenhang zwischen der Passung von Tastendruck und a priori definierten Brückeninferenzpositionen einerseits und Verstehensleistung andererseits nicht besteht, was jedoch den etablierten Textverarbeitungstheorien widerspricht. Ein dritter Erklärungsansatz stellt die Annahme in Frage, dass das Newtson Paradigma eine geeignete Methode ist, um online generierte globale Brückeninferenzen zu erfassen. Für diesen Erklärungsansatz spricht der Befund, dass zumindest lokale Brückeninferenzen automatisch generiert werden. Sie sind damit nur eingeschränkt bewusst zugänglich und dementsprechend mit einem Ansatz wie dem Newtson Paradigma, das eine gewisse bewusste Zugänglichkeit der kognitiven Prozesse voraussetzt, nicht valide erfassbar. Unklar war bisher, ob dies auch für globale Brückeninferenzen beim Filmverstehen gilt. Den Ergebnissen von Experiment 1 zufolge kann allerdings davon ausgegangen werden, dass auch die Generierung globaler Brückeninferenzen weitgehend automatisiert und unbewusst erfolgt, so dass diese Inferenzen nur sehr bedingt mit bewusstseinsabhängigen Methoden zugänglich sein dürften.

Aufbauend auf diesen Ergebnissen wurde in Experiment 2 ein anderer methodischer Ansatz gewählt, um die Frage nach der elaborierten Verarbeitung expositorischer Filme zu beantworten. Erstens wurden lokale anstelle von globalen Brückeninferenzen untersucht, da erstere eine genauere Zuordnung der verknüpften Information erlauben und sich damit eindeutiger analysieren lassen. Zweitens wurde anstelle des Newtson Paradigmas eine nicht intrusive und weniger bewusstseinsabhängige Methode aus der Textverstehensforschung verwendet, und zwar das sogenannte Naming Paradigma. In der Textverstehensforschung wird dieses Paradigma als valide Methode zur Messung von automatisierten Inferenzprozessen verwendet und es kann daher angenommen werden, dass es auch in der Filmverstehensforschung zur Messung automatisierter Brückeninferenzen angewendet werden kann. Drittens wurden in Experiment 2 kausale Brückeninferenzen untersucht, die eine spezifische Form lokaler Brückeninferenzen darstellen. Kausale Brückeninferenzen sind für den Wissenserwerb von zentraler Bedeutung und zeichnen sich darüber hinaus dadurch aus, dass sie sich sehr präzise definieren lassen. Diese methodischen Veränderungen gegenüber Experiment 1 wurden umgesetzt, indem in Experiment 2 die Verarbeitung naturwissenschaftlicher expositorischer Filme (Physik/Chemie) mit Hilfe des Naming Paradigmas untersucht wird. Die Grundidee dieser Methode ist dabei, dass die Bildung kausaler Brückeninferenzen sich daran feststellen lässt, dass die Benennungslatenzen von Wörtern verkürzt sind, wenn diese Wörter in einem Kontext präsentiert werden, in dem sie eine semantische Nähe zu aktuell generierten Brückeninferenzen aufweisen als wenn sie in semantisch nicht assoziierten Kontexten präsentiert werden.

Die Ergebnisse bestätigen die Hypothese, dass lokale kausale Brückeninferenzen während des Filmverstehens generiert werden. Naminglatenzen für Wörter im Inferenzkontext sind kürzer als Naminglatenzen für dieselben Wörter in einem Nichtinferenzkontext. Dies gilt sowohl für die Filmbedingung als auch für eine Experimentalbedingung, in der nur die entsprechende Audiospur dargeboten wurde, um den Effekt der dynamisch-piktorialen Information auf die Bildung von Brückeninferenzen zu untersuchen. Auf der Grundlage dieses Ergebnismusters kann sowohl die Annahme, dass Filme nicht elaboriert verarbeitet werden, als auch die Annahme dass die dynamische piktoriale Information eine Elaboration behindert, widerlegt werden.

Um über diese Ergebnisse hinaus den Zusammenhang zwischen online generierten lokalen kausalen Brückeninferenzen und dem Erwerb eines kohärenten mentalen Modells zu untersuchen, wurde ein drittes Experiment durchgeführt. In Experiment 3 wurde anhand des gleichen Materials wie in Experiment 2 mittels eines offline-Verständnistests untersucht, ob das in Experiment 2 nachgewiesene online Verständnis der kausalen Zusammenhänge zum Aufbau eines kohärenten mentalen Modells führt. Dazu wurden im Anschluss an die Rezeptionsphase multiple-choice Items vorgegeben, die gezielt das Verständnis der in Experiment 2 adressierten kausalen Brückeninferenzen abfragen. Als unabhängige Variable wurde einerseits interindividuell manipuliert, ob das Material als Film dargeboten wurde oder nur als Audiospur. Andererseits wurde intraindividuell manipuliert, ob die kritischen Kausalinformationen wie in Experiment 2 lediglich implizit in den präsentierten Informationen enthalten waren oder ob sie explizit dargeboten wurden. Die Bearbeitungsleistungen für die multiple-choice Items für diese kritischen Kausalinformationen wurden mit einer Kontrollgruppe verglichen, die keine Experimentalmaterialien erhalten hatte. Ein Vergleich mit dieser Kontrollgruppe ergab, dass sowohl Teilnehmer der Film- als auch der Audiobedingung bessere Verständnisleistungen zeigten als Teilnehmer der Kontrollgruppe. Darüber hinaus wurde gefunden, dass die Experimentalgruppen auch hinsichtlich der Leistungen für nur implizit enthaltene Kausalinfomationen den Leistungen der Kontrollgruppe überlegen waren. Es wurde deutlich, dass die Rezipienten in den Experimentalgruppen ein durch Inferenzen angereichertes mentales Modell auch bezüglich der nur implizit vorhandenen Informationen erworben haben. Jedoch zeigte ein Vergleich der impliziten Lernleistung mit der expliziten Lernleistung innerhalb der Experimentalgruppen, dass nur Teilnehmer mit hohem domänenspezifischem Vorwissen in der Lage sind, die implizite Information ebenso gut wie die explizite Information in ein gesamtes kohärentes mentales Modell zu integrieren. Es scheint also, als ob domänenspezifisches Vorwissen notwendig ist, um die implizite Information, die nachweisbar online inferiert wird (Experiment 2), auch erfolgreich in ein kohärentes mentales Modell zu integrieren (Experiment 3). Dies gilt sowohl für Film- als auch für Audiorezepienten.

Insgesamt widerlegt das gefundene Ergebnismuster der drei präsentierten Experimente die Annahme, dass Filme nicht oder nur in geringem Ausmaß elaboriert verarbeitet werden. Es konnte sowohl nachgewiesen werden, dass während der Filmrezeption online-Brückeninferenzen gebildet werden als auch, dass die Filmrezeption in einem vertieften offline-Verständnis resultiert. Dabei konnte in allen drei Experimenten gezeigt werden, dass die dynamisch-piktoriale Informationen, die charakteristisch für Film ist, keinen Einfluss auf online- und offline-Verarbeitungsprozesse ausübt. Die in dieser Dissertation dargestellten Experimente untersuchen erstmalig anhand direkter online Methoden Elaborationsprozesse bei der Filmrezeption und bringen diese systematisch mit offline-Verstehensprozessen in Zusammenhang. Dieser Ansatz stellt damit einen innovativen und viel versprechenden Zugang zur Analyse von Inferenz- und Lernprozessen an der Schnittstelle von kognitiver, pädagogischer und Medienpsychologie dar, dessen weitere Ausarbeitung lohnenswert erscheint.

List of Figures

Fig. 3.1: Relationships between narrative texts, expository texts, narrative films, and expository films. 10
Fig. 3.2: The comprehension process. 14
Fig. 3.3: Text, textbase, and situation model and their interrelations 23
Fig. 5.1: Shots from the James Bond film applied by Magliano et al. (1996). 34
Fig. 5.2: Shots from a film clip applied by Unsöld and Nieding (in press). 36
Fig. 6.1: Offline measures of learning outcomes in inference research. 52
Fig. 8.1: Stills from the experimental film “Post-War Germany under Allied Occupation”. 62
Fig. 8.2: Adjusted number of key pressing for each OII in the video sample. 68
Fig. 8.3: Adjusted number of key pressing for each OII in the audio sample. 68
Fig. 8.4: Means and standard deviations (in % correct) of the comprehension scores as a function of presentation mode. 71
Fig. 9.1: Stills from the expository film “Construction of a thermos”. 82
Fig. 9.2: Stills from the expository film “Formation of lightning”. 82
Fig. 9.3: Naming task procedure. 90
Fig. 9.4: Means and standard deviations (in ms) of the naming latencies as a function of context and presentation mode. 92
Fig. 10.1: Means and standard deviations (in % correct) of the comprehension test performances as a function of presentation mode. 109
Fig. 10.2: Model when prerequisite knowledge is included as a mediator variable in the relation between word set and comprehension scores for implicit items. 113

List of Tables

Table 3.1: Inferences generated in scientific text comprehension ... 16
Table 4.1: Inference-relevant differences between narrative and expository texts. ... 32
Table 6.1: Contrasting naming tasks and Newtson paradigm at a glance. ... 50
Table 7.1: Overview of Experiments 1 through 3. ... 57
Table 8.1: Design. ... 60
Table 8.2: Definition of question types. ... 63
Table 8.3: Number and format of different comprehension test questions. ... 63
Table 8.4: Procedure. ... 65
Table 8.5: Means and standard deviations for control variables as a function of presentation mode. ... 66
Table 8.6: Number of OII hits per minute and NII hits per minute. ... 67
Table 8.7: EIIs across conditions. ... 69
Table 8.8: Means and (standard deviations) for cognitive load measures as a function of presentation mode. ... 72
Table 9.1: Design. ... 81
Table 9.2: Examples from the experimental material. ... 85
Table 9.3: Procedure. ... 90
Table 9.4: Condition set-up. ... 91
Table 9.5: Means and (standard deviations) for control variables as a function of presentation mode and word set. ... 92
Table 9.6: Means and (standard deviations) of cognitive load measures and naming task disruption for word set and presentation mode. ... 93
Table 10.1: Design. ... 101
Table 10.2: Procedure for the experimental conditions. ... 106
Table 10.3: Experimental conditions. ... 106
Table 10.4: Procedure for the control condition. ... 107
Table 10.5: Control conditions. ... 107
Table 10.6: Means and (standard deviations) for the control variables as a function of presentation mode. ... 108
Table 10.7: Means and (standard deviations) for the control variables as a function of presentation mode and word set. ... 108
Table 10.8: Means and (standard deviations) of comprehension test performances (in % correct) as a function of item type, presentation mode and word set. ... 111

Table 10.9: Summary of regression analyses for mediator analyses for the relation between word set and comprehension scores for implicit items. 115

Table 10.10: Summary of regression analyses for mediator analyses for the relation between word set and comprehension scores for explicit items. 115

Table 10.11: Means and (standard deviations) of three cognitive load measures as a function of presentation mode and word set. 116

Bibliographical References

Adam, K. M., & Butler, B. E. (1999). Comprehending cognition: An integrative paradigm for psychology. *The American Journal of Psychology, 112*, 489-495.

Albrecht, E. J., & Myers, L. J. (1998). Assessing distant text information during reading: Effects of contextual cues. *Discourse Processes, 26,* 87-107.

Albrecht, E. J., & O'Brien, J. E. (1993). Updating a mental model: Maintaining local and global coherence. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 19*, 1061-1070.

Arguel, A., & Jamet, E. (2007). Using static pictures associated with dynamic visualizations to improve the learning of procedural documents. In B. Csapó & C. Csíkos (Eds.), *Proceedings of the 12th Biennial Conference for Research on Learning and Instruction (EARLI)* (p. 512). Budapest, Hungary: University of Szeged.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*, 199-209.

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556-559.

Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior, 18,* 333-356.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22*, 577-660.

Baumann, H. (1995). Probleme historischer Unterrichtsfilme [Problems of educational films in history classes]. In I. Wilharm (Ed.), *Geschichte in Bildern* (pp. 156-198). Pfaffenweiler, Germany: Centaurus-Verlagsgesellschaft.

Beagles-Roos, J., & Gat, I. (1983). Specific impact of radio and television on children's story comprehension. *Journal of Educational Psychology, 75*, 128-137.

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies texts from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly, 27*, 251-276.

Beeman, M. J., Bowden E. M., & Gernsbacher, M. A. (2000). Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language, 71,* 310–336.

Beentjes, J. W. J., & Van der Voort, T. H. A. (1993). Television viewing versus reading: Mental effort, retention, and inferential learning. *Communication Education, 42*, 191-205.

Beyer, R. (1990). *Psychologische Analyse kognitiver Prozesse bei der Textverarbeitung* [Psychological analysis of cognitive processes during text processing]. Berlin, Germany: Habilitationsschrift, Technische Universität Berlin.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, MA: Cambridge University Press.

Black, J. B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior, 20,* 267-275.

Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An on-line assessment of causal reasoning during comprehension. *Memory & Cognition, 18*, 65-71.

Boland, J. E. (2005). Visual arguments. *Cognition, 95*, 237-274.

Bordwell, D. (1985). *Narration in the fiction film.* London: Methuen.

Bordwell, D., & Thompson, K. (1993). *Film art: An introduction* (4th ed.). New York: McGraw-Hill.

Bovair, S., & Kieras, D. E. (1985). A guide to propositional analysis for research on technical prose. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 315-362). Hillsdale, NJ: Erlbaum.

Brewer, W. F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 221-239). Hillsdale, NJ: Erlbaum.

Britton, B. K., & Gülgöz, S. (1991). Using Kintsch`s computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structure. *Journal of Educational Psychology, 83*, 329-345.

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory, 9,* 365-381.

Calvo, M. G. (2004). Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading. *Learning and Individual Differences, 15,* 53-65.

Carpenter, P. A., & Daneman M. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8,* 293-332.

Cirillo, R. K., & Foss, D. J. (1980). Text structure and reading time for sentences. *Journal of Verbal Learning and Verbal Behavior, 19,* 96-109.

Clark, R. E., & Salomon, G. (1986). Media in teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 464-478). New York: Macmillan.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

DeFleur, M. L., Davenport, L., Cronin, M., & DeFleur, M. (1992). Audience recall of news stories presented by newspaper, computer, television, and radio. *Journalism Quarterly, 69,* 1010-1022.

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 121-132.

Della Sala, S., Baddeley, A., Papagno, C., & Spinnler, H. (2004). Dual-task paradigm: A means to examine the central executive. *Brain, 124,* 1479-1481.

Die Sendung mit der Maus (2007). In Wikipedia, the free encyclopedia. Retrieved October 2, 2007, from http://de.wikipedia.org/wiki/Die_Sendung_mit_der_Maus.

Durgunoğlu, A. Y., & Jehng, J.-C. J. (1991). Elaborative inferences on an expository text. *Contemporary Educational Psychology, 16,* 314-330.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Eysenck, M. W., & Keane, M. T. (2005). *Cognitive Psychology, A Student's Handbook* (5th Edition). Hove, England: Psychology Press.

Feierabend, S., & Klinger, W. (2003). Kinder und Medien 2002. Ergebnisse der Studie KIM 2002 zum Medienumgang sechs- bis 13-Jähriger in Deutschland [Results of the KIM 2002 study regarding media exposure of six to 13-year-old children]. Baden-Baden, Germany: Medienpädagogischer Forschungsverbund.

Ferstl, E. (2007). The functional neuroanatomy of text comprehension: What's the story so far? In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 53-102). Mahwah, NJ: Erlbaum.

Fincher-Kiefer, R. (1993). The role of predictive inferences in situation model construction. *Discourse Processes, 16,* 99-124.

Fincher-Kiefer, R. (2001). Perceptual components of situational models. *Memory & Cognition, 29,* 336-343.

Fincher-Kiefer, R., & D'Agostino, P. R. (2004). The role of visuospatial resources in generating predictive and bridging inferences. *Discourse Processes, 37*, 205-224.

Fischer, G., Henninger, S., & Redmiles, D. (1991). Cognitive tools for locating and comprehending software objects for reuse. In *the 13th International Conference on Software Engineering* (pp. 318-328). Austin, TX: ACM, IEEE.

Fischer, S., Lowe, R. K., & Schwan, S. (in press). Effects of presentation speed of a dynamic visualization on the understanding of a mechanical system. *Applied Cognitive Psychology.*

Fischer, S., Papenmeier, F., & Huff, M. (2007, November). *Dynamic AOIs by the use of Blenders 3D-engine*. Paper presented at the Eye Tracking and Multimedia Workshop. University of Burgundy, Dijon, France.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27,* 235-244.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33,* 465-495.

Furnham, A. (2001). Remembering stories as a function of the medium of presentation. *Psychological Reports, 89,* 483-486.

Furnham, A., Benson, I., & Gunter, B. (1987). Memory for television commercials as a function of the channel of communication. *Social Behavior, 3,* 105-112.

Furnham, A., de Siena, S., & Gunter, B. (2002). Children's and adults recall of children's news stories in both print and audio-visual presentation modalities. *Applied Cognitive Psychology, 16,* 191-210.

Furnham, A., & Gunter, B. (1985). Sex, presentation mode and memory for violent and non violent news. *Journal of Educational Television, 11,* 99-105.

Furnham, A., & Gunter, B. (1989). The primacy of print: Immediate cued recall of news as a function of the channel of communication. *Journal of General Psychology, 116,* 305-310.

Furnham, A., Gunter, B., & Green, A. (1990). Remembering science: the recall of factual information as a function of the presentation mode. *Applied Cognitive Psychology, 4,* 203-212.

Furnham, A., & Williams, C. (1987). Remembering commercials presented in different media. *Journal of Educational Television, 13,* 115-124.

Furnham, A., Proctor, E., & Gunter, B. (1988). Memory for material presented in the media: The superiority of written communication. *Psychological Reports, 63,* 935-938.

Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 250-257.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 430-445.

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

Gilabert, R., Martinez, G., & Vidal-Abarca, E. (2005). Some good texts are always better: Text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction, 15,* 45-68.

Goetz, E. T., Anderson, R. C., & Schallert, D. L. (1981). The representation of sentences in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 369-385.

Goldman, S. R., Golden, R. M., & Van den Broek, P. (2007). Why are computational models of text comprehension useful? In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension Processes* (pp. 27-52). Mahwah, NJ: Erlbaum.

Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading, 2*, 247-269.

Graesser, A. C., & Bower, G. H. (1990). *Inferences in text comprehension.* San Diego, CA: Academic Press, Inc.

Graesser, A. C., Haberlandt, K., & Koizumi, D. (1987). How is reading time influenced by knowledge-based inferences and world knowledge? In B. Britton (Ed.), *Executive control processes in reading* (pp. 217-251). Hillsdale, NJ: Erlbaum.

Graesser, A. C., León, J. A., & Otero, J. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 1-15). Mahwah, NJ: Erlbaum.

Graesser, A., Louwerse, M., McNamara, D., Olney, A., Cai, Z., & Mitchell, H. (2007). Inference generation and cohesion in the construction of situation models: Some connections with computational linguistics. In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 289-310). Mahwah, NJ: Erlbaum.

Graesser, A. C., & Magliano, J. P. (1991). Cognition and context. In R. Cohen & A. W. Siegel (Eds.), *Context and development* (pp. 57-76). Hillsdale, NJ: Erlbaum.

Graesser, A. C., McNamara, D. A., Louwerse, M. M., & Cai, Z. (2004). Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36,* 193-202.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist, 40,* 225-234.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 163-189.

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

Graesser, A. C., Wiemer-Hastings, P., & Wiemer-Hastings, K. (2001). Constructing inferences and relations during text comprehension. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 249-271). Amsterdam: John Benjamins Publishing Company.

Green, R. (1981). Remembering ideas from text: The effect of modality of presentation. *British Journal of Educational Psychology, 51,* 83-89.

Greenfield, P. M. (1984). *Mind and media: The effects of television, computers and video games*. London: Fontana.

Griesel, C., Friese, U., & Schmalhofer, F. (2003). What are the differences in the cognitive representations of predictive and bridging inferences? In F. Schmalhofer, R. Young, & G. Katz (Eds.), *Proceedings of EuroCogSci 2003: The European Cognitive Science Conference 2003* (pp. 145–50). Mahwah, NJ: Erlbaum.

Gunter, B. (1987). *Poor reception: Misunderstanding and forgetting broadcast news.* Hillsdale, NJ: Erlbaum.

Gunter, B., & Furnham, A. (1986). Sex and personality differences in recall of violent and non violent news from three presentational modalities. *Personality and Individual Differences, 7,* 829-837.

Gunter, B., Furnham, A., & Gietson, G. (1984). Memory for the news as a function of the channel of communication. *Human Learning, 3,* 265-271.

Gunter, B., Furnham, A., & Griffiths, S. (2000). Children`s memory for news: A comparison of three presentation media. *Media Psychology, 2,* 93-118.

Gunter, B., Furnham, A., & Leese, J. (1986). Memory for information from a party political broadcast as a function of the channel of communication. *Social Behavior, 1,* 135-142.

Haberlandt, K. (1994). Methods in reading research. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1-31). San Diego, CA: Academic Press.

Hacker, W., Handrick, S., & Veres, T. (1998). Lesespannen-Test (LSPT). *Verfahren zur Bestimmung der mentalen Verarbeitungskapazität (Arbeitsgedächtnis) für Text* [Procedures for determining the mental processing capacity (working memory)]. (Forschungsbericht, Bd 54). TU Dresden, Institut für Allgemeine Psychologie und Methoden der Psychologie. Dresden, Germany: TU-Eigenverlag.

Haladyna, T. (2004). *Developing and validating multiple-choice test items.* Mahwah, NJ: Erlbaum.

Halldorson, M., & Singer, M. (2002). Inference processes: Integrating relevant knowledge and text information. *Discourse Processes, 34*, 145-161.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414-434.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Haviland, E. S., & Clark, H. H. (1974). What´s new? Acquiring information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512-521.

Hobbs, R. (2006). Non-optimal uses of video in the classroom. *Learning, Media and Technology, 31*, 35-50.

James, L. R., & Brett, J. M. (1984). Mediators, moderators and tests for mediation. *Journal of Applied Psychology, 69,* 307-321.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness.* Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Just, M. A., & Carpenter, P. M. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychology Review, 99,* 122-149.

Keenan, J. M., Golding, J. M., Potts, G. R., Jennings, T. M., & Aman, C. J. (1990). Methodological issues in evaluating the occurrence of inferences. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation* (pp. 295-312). New York: Academic Press.

Kenny, D. (2007). *Mediation.* Retrieved October 23, 2007, from http://davidakenny.net/cm/mediate.htm

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A Construction-Integration Model. *Psychological Review, 95,* 163-182.

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

Kintsch, W. (2004). The Construction-Integration Model of text comprehension and its implications for instruction. In R. Ruddell & N. Unrau (Eds.) *Theoretical models and processes of reading.* (pp. 1270-1328). Newark, NJ: International Reading Association.

Kintsch, W., & van Dijk, A. T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Klin, C. M., & Myers, J. L. (1993). Reinstatement of causal information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 554-560.

Kluwe, R. H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen [Psychological methods for gaining data about human knowledge]. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp 359-385). München, Germany: PVU.

Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1994). *Image and brain - The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Kozma, R. B. (1991). Learning with media. *Review of Educational Research, 61*, 179-211.

Kürschner, C., Schnotz, W., & Eid, M. (2006). Konstruktion mentaler Repräsentationen beim Hör- und Leseverstehen [Construction of mental representations during listening and reading comprehension]. *Zeitschrift für Medienpsychologie 18*, 48-59.

Lehmkuhl, M. (2007). *Viel Sendezeit, viele Programme - Wissenschaft im deutschen Fernsehen* [Much airtime, many programs – Science on German TV]. Retrieved December 22, 2007, from http://www.goethe.de/ ins/th/prj/wif/vor/2007/art07/de2756264.htm

Linderholm, T. (2002). Predictive inference generation as a function of working memory capacity and causal text comprehension. *Discourse Processes, 34,* 259-280.

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94*, 778-784.

Linderholm, T., Gaddy Everson, M., Van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers` comprehension of easy and difficult texts. *Cognition and Instruction, 16,* 525-556.

Logie, R. H. (1995). *Visuo-spatial working memory.* Hove, England: Erlbaum.

Logie, R. H., Maylor, E. A., Della Sala, S., & Smith, G. (2004). Working memory in event- and time-based prospective memory tasks: Effects of secondary demand and age. *The European Journal of Cognitive Psychology, 16,* 441-456.

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language, 31*, 634-647.

Lorch Jr., R. F., & van den Broek, P. (1997). Research for the future - Understanding reading comprehension: Current and future contributions of cognitive science. *Contemporary Educational Psychology*, 22, 213-246.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225-244.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157-176.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14*, 257–274.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58,* 593-614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test the significance of the mediated effect. *Psychological Methods, 7,* 83-104.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes, 22*, 199-224.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics, 20*, 193-232.

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology, 15,* 533-545.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*, 615-629.

Magliano, J. P., Zwaan, R. A., & Graesser, A. (1999). The role of situational continuity in narrative understanding. In S. R. Goldman & H. van Oostendorp (Eds.), *The construction of mental representations during reading* (pp. 219-245). Mahwah, NJ: Erlbaum.

Marx, H., & Jungmann, T. (2000). Abhängigkeit der Entwicklung des Leseverstehens von Hörverstehen und grundlegenden Lesefertigkeiten im Grundschulalter: Eine Prüfung des Simple View of Reading-Ansatzes [Dependence of development of reading comprehension and basic reading skills in elementary school age: An examination of the Simple View of Reading Approach]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 32*, 81-93.

Maury, P., Pérez, O., & León, J. (2002). Predictive inferences in scientific and technological contexts. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 199-221). Mahwah, NJ: Erlbaum.

Mayer, R. E. (2001). *Multimedia learning.* Cambridge, MA: Cambridge University Press.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.). *The Cambridge handbook of multimedia learning* (pp. 31-48). Cambridge, MA: Cambridge University Press.

McGee, A., & Johnson, H. (2003). The effect of inference training on skilled and less skilled comprehenders. *Journal of Educational Psychology, 23*, 49-59.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 82-91.

McKoon, G., & Ratcliff, R. (1989). Assessing the occurrence of elaborative inference with recognition: Compatibility checking vs compound cue. *Journal of Memory and Language, 28,* 547-563.

McKoon, G., & Ratcliff, R. (1990). Dimensions and inferences. In A. Graesser & G. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension* (pp 313-328). New York: Academic Press.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*, 440-466.

McKoon, G., & Ratcliff, R. (1995). The minimalist hypothesis: Directions for research. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension - Essays in honor of Walter Kintsch* (pp. 97-116). Hillsdale, NJ: Erlbaum.

McKoon, G., Ratcliff., R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the online lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1219-1228.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes, 22,* 247-288.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1-43.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iStart: Interactive strategy trainer for active reading and thinking. *Behavioral Research Methods, Instruments, and Computers, 36*, 222-233.

Meringoff, L. K. (1980). Influence of the medium on children's story apprehension. *Journal of Educational Psychology, 72*, 240-249.

Miller, J. R., & Kintsch, W. (1980). Readability and recall for short passages: A theoretical analysis. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 335-354.

Millis, K. K., & Graesser, A. C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language, 33,* 583-599.

Murray, J. D., & Burke, K. A. (2003). Activation and encoding of predictive inferences: The role of reading skill. *Discourse Processes, 35*, 81-102.

Murray, J. D., Klin, C. M., & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory and Language, 32,* 464-473.

Myers, G. (1991). Lexical cohesion and specialized knowledge in science and popular science texts. *Discourse Processes, 14*, 1-26.

Myers, J. L., O'Brien, E. J., Albrecht, J. E., & Mason, R. A. (1994). Maintaining global coherence during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 876-886.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of learning environments. *Cognition and Instruction, 9,* 329-389.

Neuman, S. B. (1989). The impact of different media on children's story comprehension. *Reading Research and Instruction, 28*, 38-47.

Neuman, S. B. (1992). Is learning from media distinctive? Examining children's inferencing strategies. *American Educational Research Journal, 29*, 119-140.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology, 28,* 28-38.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Noordman, L. G. M., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language, 31,* 573-590.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 410-420.

Ohler, P. (1994). *Kognitive Filmpsychologie. Verarbeitung und mentale Repräsentation narrativer Filme* [Cognitive film psychology. Processing and mental representations of narrative films]. Münster, Germany: MAkS-Publikationen.

Paas, F., Tuovinen, J., Tabbers, H., & van Gerven, P. (2003). Cognitive load measurement as means to advance cognitive load theory. *Educational Psychologist, 38*, 63-74.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*, 241-263.

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

Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based tasks. *Human-Computer Interaction, 8,* 193-216.

Potts, G. R., Keenan, J. M, & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language, 27*, 399-415.

Poynor, D. V., & Morris, R. K. (2003). Inferred goals in narratives: Evidence from self-paced reading, recall, and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 3-9.

Preacher, K. J., & Leonardelli, G. J. (2007). *Calculation for the Sobel Test. An interactive calculation tool for mediation tests.* Retrieved November 23, 2007, from http://www.psych.ku.edu/preacher/sobel/sobel.htm

Rayner, K. (1997). Understanding eye-movements in reading. *Scientific Studies of Reading 1*, 317-339.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

Rayner, K., Sereno, S., Morris, R. K., Schmauder, R., & Clifton, C. J. (1989). Eye movements and on-line comprehension processes. *Language and Cognitive Processes, 4,* 121-150.

Rebetez, C., Bétrancourt, M., Sangin, M., & Dillenbourg, P. (2008). *Learning from animation enabled by collaboration.* Manuscript submitted for publication.

Reed, S. (1999). *Word problems: Research and curriculum reform.* Mahwah, NJ: Erlbaum.

Rötting, M. (2001). *Parametersystematik der Augen und Blickbewegungen für arbeitswissenschaftliche Untersuchungen* [Parameter taxonomy of eye and gaze movements for ergonomic studies]. Aachen, Germany: Shaker.

Rummer, R., Schweppe, J., Scheiter, K., & Gerjets, P. (2008). Lernen mit Multimedia: Die kognitiven Grundlagen des Modalitätseffekts [Learning with multimedia: The cognitive basics of the modality effect]. *Psychologische Rundschau, 59*, 98-107.

Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attribution. *Journal of Educational Psychology, 76*, 647-658.

Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2006). A methodological alternative to media comparison studies: Linking information utilization strategies and instructional approach in hypermedia learning. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 2117-2122). Mahwah, NJ: Erlbaum.

Schelling, G., Aarnoutse, C., & van Leuwe, J. (2006). Third graders' think aloud protocols: Types of reading activities in reading an expository text. *Learning and Instruction*, *16,* 549-568.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: Detection, search, and attention. Psychological Re*view, 84*, 1-66.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review, 14,* 101-120.

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49-69). Cambridge, MA: Cambridge University Press.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*, 47-58.

Schwan, S. (2007). Lernen mit Filmen [Learning with films]. In H. Mandl & J. Zumbach (Eds.), *Pädagogische Psychologie in Theorie und Praxis. Ein fallbasiertes Lehrbuch* (pp. 99-104). Göttingen, Germany: Hogrefe.

Schwan, S., & Garsoffky, B. (2008). The role of segmentation for perception and understanding of events. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding Events* (pp. 391-414). New York: Oxford University Press

Schwan, S., Garsoffky, B., & Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Memory & Cognition, 28*, 214-223.

Schwan, S., Hesse, F. W., & Garsoffky, B. (1998). The relationship between formal filmic means and the segmentation behavior of film viewers. *Journal of Broadcasting and Electronic Media, 42*, 237-249.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14,* 293-305.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489-537.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition, 12*, 315-328.

Seifert, C. M., Robertson, S. P., & Black, (1985). Types of inferences generated during reading. *Journal of Memory and Language, 24,* 405-422.

Sharkey, A. J. C., & Sharkey, N. E. (1992). *Weak contextual constraints in text and word priming. Journal of Memory and Language, 31*, 543-572.

Singer, M. (1994). Discourse inference processes. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 479-515). San Diego, CA: Academic Press.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 22,* 437–448.

Singer, M., & Gagnon, N. (1999). Detecting causal inconsistencies in scientific text. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 179–194). Mahwah, NJ: Erlbaum.

Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language, 33,* 421-441.

Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology, 30,* 1–38.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31*, 507-524.

Singer, M., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes, 24,* 199-228.

Singer, M., Revlin, R., & Halldorson, M. (1990). Bridging-inferences and enthymeme. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension* (pp. 35-51). San Diego, CA: Academic Press.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodolo*gy 1982 (pp. 290-312). Washington DC: American Sociological Association.

Spivey, M. J., Richardson, D. C., Tyler, M. J., & Young, E. E. (2000). Eye movements during comprehension of spoken scene descriptions. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society* (pp. 487-492). Mahwah, NJ: Erlbaum.

Sturm, H. (1984). Wahrnehmung und Fernsehen - Die fehlende Halbsekunde. Plädoyer für eine zuschauerfreundliche Mediendramaturgie [Perception and TV – The missing half second. Pleadings for a spectator friendly media dramaturgy]. *Media Perspektiven, 1*, 58-64.

Suh, S., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of Memory and Language, 32*, 279-300.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257-285.

Sweller, J. (1999). *Instructional design in technical areas.* Camberwell, Australia: ACER.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19-30). Cambridge, MA: Cambridge University Press.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Swinney, D. A. (1979). Lexical decision during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-659.

Tapiero, I., & Otero, J. (2002). Situation models as retrieval structures: Effects on the global coherence of science texts. In J. Otero, J. A. León & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 179-198). Mahwah, NJ: Erlbaum.

Tompkins, C. A., Lehman-Blake, M. T., Baumgaertner, A., & Fassbinder, W. (2001). Mechanisms of discourse comprehension impairment after right hemisphere brain damage: Suppression in inferential ambiguity resolution. *Journal of Speech, Language, and Hearing Research, 44,* 400-415.

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16,* 3-34.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24,* 612-630.

Trabasso, T., van den Broek, P, & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12,* 1-26.

Tversky, B., Bauer-Morrison, J., & Bétrancourt, M. (2002) Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.

Unsöld, I. H., & Nieding, G. (in press). Die Bildung von Prädiktiven Inferenzen von Kindern und Erwachsenen während der kognitiven Verarbeitung von audiovisuellen Texten [The construction of predictive inferences by children and adults during the cognitive processing of audiovisual texts]. *Zeitschrift für Entwicklungs- und Pädagogische Psychologie.*

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467.

Van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423-445). Hillsdale, NJ: Erlbaum.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies in Discourse Comprehension.* New York: Academic Press.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*, 167-174.

Veenman, M. V. J. (1993). *Intellectual ability and metacognitive skill: Determinants of discovery learning in computerized learning environments.* Amsterdam: University of Amsterdam.

Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 77-99). Münster, Germany: Waxmann.

Vorderer, P. (1992). *Fernsehen als Handlung* [TV watching as action]. Berlin, Germany: Sigma.

Walma van der Molen, J. H., & van der Voort, T. H. A. (1997). Children's recall of television and print news: A media comparison study. *Journal of Educational Psychology, 89,* 82-91.

Walma van der Molen, J. H., & van der Voort, T. H. A. (2000). Children's and adults' recall of television and print news in children's and adult news formats. *Communication Research, 27,* 132-160.

Wanner, E. (1975). *On remembering, forgetting, and understanding sentences.* The Hague, Netherlands: Mouton.

Weidenmann, B. (1989). Der mentale Aufwand beim Fernsehen [The mental effort during watching TV]. In J. Groebel & P. Winterhoff-Spurk (Eds.), *Empirische Medienpsychologie* (pp. 134-149). München, Germany: Psychologie-Verlags-Union

Weidenmann, B. (2002). Abbilder in Multimedia-Anwendungen. In L. J. Issing & P. Klimsa (Eds.), *Information und Lernen mit Multimedia* (pp. 83-96). Weinheim, Germany: Psychologie Verlags Union

Wember, B. (1976). *Wie informiert das Fernsehen? Ein Indizienbeweis*. (How informs TV? An evidence proof]. München, Germany: List.

Wetzel, C. D., Radtke, P. H., & Stern, H. W. (1994). *Instructional effectiveness of video media.* Hillsday, NJ: Erlbaum.

Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from scientific text. *Discourse Processes, 26*, 109-129.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Mahwah, NJ: Erlbaum.

Wirth, W. (2006). Involvement. In J. Bryant & P. Vorderer (Eds.), *Psychology of Entertainment* (pp. 199-213). Mahwah, NJ: Erlbaum.

Wolfe, M. W. (2005). Memory for narrative and expository text: Independent influences of semantic associations and text organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 359-364.

Wolfe, M. W., & Mienko, J. A. (2007). Learning and memory of factual content from narrative and expository text. *British Journal of Educational Psychology, 77,* 541-564.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1-14.

Zacks., J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*, 3-21.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 920-933.

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situational models. *Discourse Processes, 28,* 81-88.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 35-62). New York: Academic Press.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 386-397.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

Filmographical References

FWU (2003). *Sieger und Besiegte im Nachkriegsdeutschland 1945-1950* [Winners and defeated in post-war Germany] [Motion picture]. Grünwald, Germany.

FWU (1999). *Der See* [The lake] [Motion picture]. Grünwald, Germany.

Gilbert, L. (Producer), & Broccoli, A. R. (Director). (1979). *Moonraker* [Motion picture]. United States: CBS/Vox Video.

Maiwald, A. (Writer), & Biermann, C. (Director). (2006). Abblendspiegel [Rear view mirror] [Television series episode]. In J. Lachmuth (Producer), Sachgeschichten aus der Sendung mit der Maus. Köln, Germany: WDR.

Maiwald, A. (Writer), & Biermann, C. (Director). (2006). Blitzentstehung [Development of a lightning] [Television series episode]. In J. Lachmuth (Producer), Sachgeschichten aus der Sendung mit der Maus. Köln, Germany: WDR.

Maiwald, A. (Writer), & Biermann, C. (Director). (2006). Thermoskanne [Thermos] [Television series episode]. In J. Lachmuth (Producer), Sachgeschichten aus der Sendung mit der Maus. Köln, Germany: WDR.

Appendixes

A Original Film Transcripts

A.1 German post-war history (Experiment 1)

A.2 Thermos (Experiment 2 and 3)

A.3 Lightning (Experiment 2 and 3)

B Propositional Analyses

B.1 German post-war history (Experiment 1)

B.2 Thermos (Experiment 2 and 3)

B.3 Lightning (Experiment 2 and 3)

C Experimental Film Transcripts

C.1 Thermos (Experiment 2 and 3)

C.2 Lightning (Experiment 2 and 3)

C.3 Naming words

D Offline Comprehension Tests

D.1 Offline comprehension test (Experiment 1)

D.2 Offline comprehension test (Experiment 3)

A

Original Film Transcripts

A.1 German post-war history (Experiment 1)

„Was soll aus Deutschland werden?“

„Dieser Krieg ist nicht wie in der Vergangenheit. Wer immer ein Gebiet besetzt, erlegt ihm auch sein eigenes gesellschaftliches System auf. Jeder führt sein eigenes System ein, so weit seine Armee vordringen kann.“ – Stalin, 1945

„Was machen wir am Tage nach unserem Sieg? Unser Einsatz hat das Ziel, Frankreich auf Dauer hier zu etablieren.“ – de Gaulle, 1945

„Es empfiehlt sich, Deutschland in ein Land mit vorwiegend land- und weidewirtschaftlichen Charakter umzuwandeln.“ – Morgenthau, 1944

Was soll aus Deutschland werden? Die meisten Deutschen in den zerstörten Städten dachten nicht an die Zukunft. Sie sorgten sich ums Überleben am nächsten Tag. Hunger herrschte. Hamsterfahrten für ein paar Kartoffeln. Tauschgeschäfte. „Biete Markenfüllhalter - neu, suche Briketts.“ Bitter die Kälte in den unbeheizbaren Behausungen. Bis zu 80 % der Häuser und Wohnungen waren zerstört. In den Ruinenlandschaften Flüchtlinge, Vertriebene, Kriegsheimkehrer, ehemalige Kriegsgefangene. Menschen, die ihre Familien suchten. Ein Schweizer an seinen amerikanischen Freund: „der Deutsche unserer Tage glaubt nichts, hofft nichts und duldet alles. Wer hier noch Demokratie und Menschenwürde predigen will, gibt Steine statt Brot. Deutschland muss untergehen oder neu erstehen. Weiter vegetieren darf es um Europawillen nicht.“

Schon vor der bedingungslosen Kapitulation Deutschlands auf der Konferenz von Jalta, im Februar 1945, hatten die großen Drei, Großbritannien, USA und die Sowjetunion beschlossen, Teile des deutschen Reiches im Osten abzutrennen. So fiel der nordöstliche Teil Ostpreußens an die Sowjetunion. Der Rest wurde unter polnische Verwaltung gestellt. Ebenso Pommern und Schlesien. Ein Ausgleich dafür, dass Polen seine östlichen Gebiete an die Sowjetunion abgeben musste. Diese Westverschiebung Polens bis zu Oder und Neiße sollte in einem späteren Friedensvertrag endgültig festgeschrieben werden. Diese Grenze wurde erst 1990 von der Bundesrepublik anerkannt. Trotz unterschiedlicher Auffassungen, ein wirkliches Konzept, was aus Deutschland werden sollte, hatte keiner der Siegermächte. Aber sie waren sich in vier Punkten einig, die als die vier D’s in die Geschichte eingegangen sind: Denazifizierung, Demilitarisierung, Demontage, Demokratisierung.

„Es ist nicht unsere Absicht, das deutsche Volk zu vernichten, aber erst nach der Auslöschung des Nazitums und des Militarismus wird für die Deutschen Hoffnung auf ein bescheidenes Leben und auf einen Platz in der Gemeinschaft der Nationen bestehen.“

Potsdam 1945: Nach dem Sieg über den gemeinsamen Gegner Deutschland traten die gegensätzlichen machtpolitischen Interessen der Sieger offen zu Tage. Dennoch verständigte man sich auf der Konferenz von Potsdam darauf, Deutschland nicht in mehrere Teilstaaten zu zerstückeln. So wurde es nun in vier Besatzungszonen und Berlin in vier Sektoren aufgeteilt. Die vierte Besatzungsmacht war Frankreich, das jedoch in Potsdam nicht vertreten war. Wesentlich für den Erhalt der staatlichen Einheit war der Beschluss, Deutschland als ein wirtschaftliches Ganzes zu betrachten und zu verwalten. Oberstes Regierungsorgan war der alliierte Kontrollrat in Berlin. Hier sollten alle, Deutschland als Ganzes betreffende Angelegenheiten geregelt werden. Wohl die wichtigste von den wenigen gemeinsamen Aktionen der Siegermächte war der Prozess gegen die Hauptverantwortlichen des NS-Staates in Nürnberg. Sie wurden der Kriegsverbrechen und der Verbrechen gegen die Menschlichkeit angeklagt und verurteilt.

Die Demontage und der Abtransport deutscher Industrieanlagen begannen im Sommer 1945. Die Alliierten waren in der Frage der Reparationen uneins. Amerikaner und Briten verfuhren schon bald zurückhaltend, doch die Sowjetunion und Frankreich, die unter der deutschen Besatzung und Kriegshandlungen am schwersten gelitten hatten, bestanden auf Wiedergutmachungen im großen Umfang. Um jedoch den Sowjets den Abtransport von Maschinen und Geräten im Ruhrgebiet zu verwehren, und damit das größte deutsche Industriegebiet am Leben zu erhalten, setzten die Amerikaner durch, dass jede Besatzungsmacht nur Zugriff auf ihre eigene Zone haben dürfe. Mit dieser folgenschweren Reparationsregelung war eine Trennungslinie durch Deutschland gezogen. Es war der Anfang der Teilung Deutschland in zwei Wirtschaftsgebiete, und schließlich in zwei deutsche Staaten.

„Das russische Fernziel heißt Weltherrschaft. Macht ist das einzige, was die Russen verstehen.“ Diese Bemerkung des amerikanischen Präsidenten Truman signalisierte eine Wende in der amerikanischen Politik. Keine Zusammenarbeit mit den Russen, sondern weltweite Eindämmung des kommunistischen Einflusses. Westdeutschland, zunächst die amerikanische Zone, war ein wichtiges Feld der amerikanischen Außenpolitik.

Die Deutschen in der amerikanischen Zone spürten die Besatzer weniger als die Menschen in den anderen Zonen. Ihre Politik verstanden die Amerikaner als Demokratisierungsprozess, vor allem der Jugend, aber auch durch Presse und Rundfunk. An den Schulen wurden die nationalsozialistisch-belasteten Lehrer entfernt und die Schülermitverwaltung eingeführt. In den Medien achteten die amerikanischen Presseoffiziere auf eine saubere Trennung von Nachricht und Kommentar und sorgten für eine positive Darstellung der freiheitlichen Staats- und Gesellschaftsform. Kritik war willkommen, allerdings nicht an der Besatzungsmacht. Die Amerikaner strebten für ihre Zone eine föderale Struktur nach dem Vorbild der USA an, wo starke Bundessstaaten die Politik bestimmen. So gliederten sie ihre Zone in vier Verwaltungseinheiten, aus denen schon bald die Bundesländer Bayern, Württemberg-Baden, Hessen und Bremen werden sollten. Schon im Januar 1946 ließen sie die ersten Gemeinderats- und Bürgermeisterwahlen zu. Zu dieser stufenweisen Demokratisierung von unten gehörte es, dass schon sehr bald Aufgaben von der Militärregierung auf deutsche Verwaltungen übertragen wurden. So war Amerika auch zur Wahrung eigener Interessen von Anfang an bemüht, ein demokratisches Deutschland zu etablieren.

„Das amerikanische Volk will dem deutschen Volk helfen, seinen Weg zurückzufinden zu einem ehrenvollen Platz unter den freien und friedliebenden Völkern der Welt." – Byrnes, Stuttgart-Rede, 1945

Großbritannien sah das dringlichste Ziel darin, den preußischen Militarismus zu zerschlagen. Doch bereits ein Jahr nach Kriegsende erschien es wichtiger, Deutschland als Bollwerk gegen den Kommunismus zu stabilisieren. „Die russische Gefahr ist mit Sicherheit genauso groß, möglicherweise noch größer als die Gefahr eines wieder erstarkten Deutschlands." –Bevin. So begannen die Engländer in ihrer Zone, ähnlich wie die Amerikaner, sich nachhaltiger für den Wiederaufbau einzusetzen. Die Lebensmittelrationen wurden erhöht und die Reparationen gedrosselt. Aus demokratischer Tradition tolerierten die Briten die neu gegründeten deutschen Parteien. Aber im Gegensatz zu den Amerikanern, behielt die britische Militärregierung alle Fäden in der Hand. Wie in der bewährten Verwaltungspraxis britischer Kolonien gab sie den Ländern nur geringen Spielraum zur Eigenverantwortlichkeit.

Die Deutschland-Politik Frankreichs war von einem starken Sicherheitsbedürfnis bestimmt. Ursache waren die mehrfachen leidvollen Erfahrungen mit dem aggressiven Nachbar Deutschland. Deshalb wünschte Frankreich ein zerstückeltes Deutschland aus wirtschaftlich schwachen Einzelstaaten und verhinderte jede zentrale, zonenübergreifende Einrichtung. Zudem fordert es eine internationale Kontrolle des Ruhrgebietes, der einstigen Waffenschmiede Deutschlands. Frankreich kapselte seine Besatzungszone im Widerspruch zu dem Potsdamer Abkommen von den anderen ab und gliederte das Saargebiet in das Wirtschafts- und Währungsgebiets Frankreichs ein. Churchill sprach von einem „seidenen Vorhang", der die französische Zone von den anderen Besatzungszonen trennte. Die Bevölkerung bekam die harte Besatzungspolitik der Franzosen im Alltag der Nachkriegszeit zu spüren. Gleichzeitig aber versuchte auch Frankreich, ähnlich wie die USA durch eine gezielte Schul- und Kulturpolitik gerade die Jugend zu gewinnen und einen langfristigen Demokratisierungsprozess einzuleiten. Diese Politik ebnete schließlich den Weg für die spätere deutsch-französische Verständigung.

Die Sowjetunion war, wie zunächst Frankreich, an einer alle Zonen umfassenden Zusammenarbeit im Interesse Gesamtdeutschlands nicht interessiert, wenn auch zum Teil aufgrund anderer politischen Überlegungen als Frankreich. Auf eigene Faust hatte die Sowjetunion begonnen, in ihrer Besatzungszone tief greifende Veränderungen durchzusetzen. Mit 60 000 sowjetischen Kadern und einer Gruppe deutscher Exilkommunisten wurden nach und nach die Schlüsselpositionen in den rasch gegründeten Ländern Mecklenburg, Brandenburg, Sachsen-Anhalt, Thüringen und Sachsen besetzt. Anders als in den anderen drei westlichen Zonen waren Entnazifizierung und Bildungsreform Instrumente gesellschaftlich ideologischer Umwälzungen, die alle Lebensbereiche durchdringen sollten. Enteignungen und Verstaatlichungen von Banken und Industrien waren Anfänge einer zentralgelenkten staatsmonopolitstischen Wirtschaft. Die Bodenreform eine Vorstufe zur Kollektivierung der Landwirtschaft. Die Vereinigung von SPD und KPD zur SED, von den kommunistischen gefeiert als Bündelung aller proletarischen Kräfte, geschah nicht freiwillig, sondern war von der KPD durch Pression erzwungen. So äußerte der deutsche KPD-Führer Walter Ulbricht: „ Es muss demokratisch aussehen, aber wir müssen alles in der Hand haben."

Die unterschiedlichen Interessen der vier Siegermächte verhinderten ihre ursprüngliche Absicht, Deutschland als Ganzes zu regieren. Der Graben zwischen der kommunistisch geführten Ostzone und denen im Sinne westlicher Demokratie geführten Westzonen wurde tiefer. Vor dem Hintergrund des sich verschärfenden Ost- Westkonfliktes und angesichts wachsender wirtschaftlicher Not war es nur folgerichtig, dass die amerikanische und britische Zone näher zusammenrückten. Die Amerikaner bezogen die Westzonen in das CARE-Programm mit ein. Eine Überlebenshilfe für die hungernde Bevölkerung. Ebenso wie die Schulspeisung für 3,5 Millionen Kinder in der amerikanischen und britischen Zone. Anfang 1947 schlossen sich die britische und amerikanische Zone zur Bi-Zone zusammen. Erst zwei Jahre später sollte auch Frankreich auf diese gemeinsame Linie einschwenken. „Setzt die Deutschen in die Lage zu arbeiten“; forderten amerikanische Journalisten und Kongressabgeordnete. Dafür erwies sich der Marshall-Plan, der auf alle drei Westzonen ausdehnt wurde als geeignetes Instrument. Aus dieser Aufbauhilfe für ganz Europa flossen 1,5 Milliarden US-Dollar allein in die westdeutsche Wirtschaft. Der Marshall-Plan wurde zur Grundlage des westdeutschen Nachkriegswohlstandes. Die Sowjetunion hatte für ihre Zone Marshall-Plan-Hilfe abgelehnt.

Im März 1948 verließ die Sowjetunion den alliierten Kontrollrat, der danach nie wieder zusammentrat. Die Vier-Mächte-Verwaltung Gesamtdeutschlands war damit endgültig beendet. Längs hatten die Sowjetunion und die Westalliierten zusammen mit ihren Verbündeten erste Schritte zur Gründung deutscher Staaten in ihrem Besatzungsgebiet eingeleitet. In Deutschland war im Zuge des kalten Krieges eine Grenze zwischen zwei verfeindeten Machtblöcken gezogen worden. „Ein eiserner Vorhang“, sagte Churchill, „habe sich gesenkt.“ Die Gründung zweier deutscher Staaten 1949 war die Konsequenz aus vier Jahren Besatzungsherrschaft in Deutschland. So sind beide Staaten in ihren Grundzügen geprägt von den Siegermächten, deren politische Systeme gegensätzlicher kaum vorstellbar sind.

Schon 1946 hatte Churchill prophezeit: „wir müssen der Tatsache ins Auge sehen, dass so wie die Dinge gegenwärtig stehen, zwei Deutschlands im Entstehen sind. Das eine, mehr oder weniger organisiert im russischen Interesse, das andere nach dem der westlichen Demokratie.“

Die Grenze durch Deutschland wurde für 40 Jahre zur Todeszone zwischen zwei verfeindeten Machtblöcken. Erst die Veränderungen im Ostblock und das Aufbegehren der Menschen in der DDR schufen die Voraussetzungen für eine Vereinigung Deutschlands unter demokratischen Vorzeichen.

A.2 Thermos (Experiment 2 und 3)

Ein schönes Glas Tee. Kann Christoph in aller Ruhe trinken. Uhh, heiß. Kann er kaum anfassen. Wenn das Telefon nicht wäre. Ist wieder einer dran, der erzählt und erzählt und erzählt. Und wenn das Gespräch dann beendet ist, dann ist der Tee kalt geworden. Mhmm. Wie kann man das verhindern? Das muss doch raus zu kriegen sein, denkt Christoph. Mit einem Topf voll mit heißem Wasser? Da ist die Wärme jetzt noch drin, klar. Dann hängt er da noch was rein. Und ne Stoppuhr. Er will testen, welches Material am wenigsten die Wärme wegnimmt. Nach 30 Sekunden der Fingertest. Der Löffel aus Metall … heiß. Glas, mhm, geht so. Ins Gummi ist fast gar keine Wärme reingewandert. Und das Holz, ist kalt geblieben. Das is es. Das leitet die Wärme wohl am wenigsten weg.

Tja, Christoph ist ein Mann der Tat. Gleich greift er zur Bohrmaschine und bohrt ein Riesenloch in einen Holzklotz. Ja, sieht irgendwie praktisch aus. Und was soll es sein? Eine Teetasse – aus Holz. Da müsste die Wärme doch länger drin bleiben. Ja, heiß! Aber, nich lange. Mhm. Das muss noch besser gehn, findet Christoph, mal überlegen. Mhm. Von den vier Gegenständen war der aus Holz schon der beste. Aber vielleicht gibt’s etwas, das noch weniger von der Wärme wegnimmt.

Neuer Versuch: Wieder heißes Wasser. Zwei Gefäße. Christoph will wohl was vergleichen. Auf ein Gefäß legt er ein Glas und dann noch eins. Und auf das andere Gefäß auch ein Glas, dann Korkstückchen und dann erst das zweite Glas. Dazwischen ist also Luft. Und jetzt, der Buttertest. Wo schmilzt die Butter zuerst? Natürlich da, wo die Wärme aus dem heißen Wasser am besten zur Butter geleitet wird. Links ist sie geschmolzen und da wo die Luft dazwischen ist, ist die Wärme nicht durchgekommen.

Ein Glas, ein paar Glasstücken als Abstandshalter zum Boden und darein jetzt das Teeglas. Jetzt nur von Luft umgeben. Mal gucken, ob das funktioniert. Test mit Vergleichsteeglas und der Anfangstemperatur. Mhm, 80, 81, gute 82 Grad. Und da, 80, 81, auch da, 82 Grad. Also, die Zeit läuft. Nach einer Viertelstunde, stopp und nachmessen. Ohne Luftpolster 46 Grad. Bei dem Glas im Glas mit Luft dazwischen 60 Grad. Ganz schöner Unterschied. Nach unten und zur Seite kann die Wärme also nicht mehr wegwandern. Aber, noch nach oben. Da geht noch Wärme weg. Und zwar ganz schön viel, denn Wärme will ja immer nach oben. Wie kann man die aufhalten?

Zweiter Test mit dem Vergleichsteeglas. Diesmal mit Deckel. Aber nicht nur ein Deckel, einer für jedes Glas – beim Glas im Glas. Wieder nach ner Viertelstunde messen. Im Vergleichsteeglas, das überrascht jetzt nicht, wieder so 45, 46 Grad. Und jetzt mal gucken, was der Deckel gebracht hat. 70, 71, 72 Grad! Also, 12 Grad mehr als ohne Deckel. Da ist der Tee noch heiß, 72 Grad, Christoph, 72, pffff, heiß! Aber Christoph ist trotzdem stolz auf seine neue Tasse. Vielleicht gibt’s aber noch was zu verbessern. Er hat ne Idee.

Alufolie, silbrig-spiegelnd! Gibt’s in fast jedem Haushalt. Damit wird ja auch schon mal Essen warm gehalten. Natürlich kann man damit auch ein Teeglas drumwickeln. Einpacken! Aber, welchen Sinn soll das haben? Um das zu verstehen, nimmt Christoph eine Wärmelampe, ein Glas ohne Alu und eins mit. Und jetzt, wieder der Buttertest. Klares Ergebnis. Mit Alufolie bleibt die Butter ganz, weil die Wärme von der Alufolie wieder zurückgespiegelt wird. Und nichts ins Glas reinkommt. Und was Wärme nicht rein lässt, lässt sie auch nicht raus.

Dann müsste man doch jetzt das Glas mit dem heißen Tee gut anfassen können. Wenn die Alufolie drum ist. Ohne Alufolie… vorsicht… ohne… Aber mit Alufolie kann man es ganz bequem in die Hand nehmen. Also ist Alufolie für Christoph's neue Tassenerfindung noch eine Bereicherung. Rein damit in das andere Glas mit den Korkstücken als Abstandshalter für das Luftpolster. Tee rein. Und dann noch den praktischen Doppeldeckel mit dem Strohhalm drauf. Super! Das alles hält den Tee jetzt schön heiß. Ja, Christoph. Heiß! Aber trotzdem ist er stolz auf seine neue Teewarmhaltetasse.

Aber Christoph, weißt du? Das gibt's schon. Schon mal gehört? Thermoskanne? Ja, die hält den Tee auch schön warm. Funktioniert ganz ähnlich wie deine Erfindung. Wir ham mal eine aufgeschnitten. Auch eine Flasche in der Flasche. Innen verspiegelt und ringsrum Luft, die die Wärme nicht wegleitet.

Aber, Christoph trinkt natürlich am liebsten aus seiner eigenen Teewarmhaltetasse. Darin bleibt der Tee auch wunderbar heiß. Ja, heiß. Er wird's nicht lernen. Trotzdem, köstlich!

A.3 Lightning (Experiment 2 und 3)

Ihr habt alle schon mal erlebt, dass es blitzt. Was hat dieser Blitz mit einem Luftballon zu tun, den man mit einem Wollpullover reibt und der danach Papierkonfetti anzieht? Es ist beides ungefähr das Gleiche. Wie bitte? Was ist ein Blitz, und warum blitzt es? Die Zutaten dazu hat jeder zuhause. Wenn ich einen Luftballon nehme und reibe den mit einem Wollpullover, stelle ich Elektrizität her. Ich bin also Armin das Elektrizitätswerk. Kann man nicht sehen. Aber man kann sehen, dass die Elektrizität da ist. Sie ist auf der Ballonhülle gespeichert. Und sie zieht Franziskas lange Haare an. Die Elektrizität kann nicht von der Ballonhülle weg. Und Elektrizität ist die Hauptzutat für den Blitz, aber noch nicht der eigentliche Blitz selbst. Wenn ich den Ballon jetzt nehme, lege den ins Wasser, dann kann die Elektrizität den Ballon verlassen. Sie fließt ins Wasser ab. Wenn ich den Ballon jetzt wieder neben Franziskas Haare tue, dann kommen die Haare nicht hoch. Sie werden nicht mehr angezogen. Was sich da bewegt, macht nur der Wind. Wenn ich jetzt neue Elektrizität auf dem Luftballon draufhaben will, dann muss ich den erst wieder trocken machen und mit dem Wollpullover neue Elektrizität da drauf reiben, erzeugen. Das kann allerdings jeder von euch zuhause selber ausprobieren. Tut nicht weh, und es passiert nichts bei. Wie viel Elektrizität krieg ich denn auf so nen Luftballon drauf? Hängt ein bisschen von der Größe ab, und wie viel ich reibe. Hier ist die auf dem Ballon gespeicherte Elektrizität wieder. Jetzt kann ich euch zeigen was passiert, wenn dieser Speicher zu voll ist.

Was eben mein Pullover war, das sind jetzt diese beiden Bänder, die ich einschalte. Die reiben aneinander und stellen Elektrizität her. Und die Kugel ist das, was eben mein Luftballon war. Und wenn dieser Speicher auf der Kugel jetzt zu voll ist, und ich komm mit ner Metallspitze in die Nähe, dann springt irgendwann die Elektrizität von der Kugel weg. Und das ist ein kleiner Blitz. Das Band hat also durch Reibung so viel Elektrizität erzeugt, dass ein kleiner Blitz entsteht.

Aber was reibt denn am Himmel? Da reibt doch niemand dran. Nein, es reibt kein Mensch dran. Im Zeitraffer sieht man, wie die Wolken über den Himmel rasen. Das sind die Luftschichten, die sich aneinander vorbeibewegen. Und winzig kleine Eiskristalle reiben sich an der Luft und stellen dadurch die Elektrizität her. Die Elektrizität für den Blitz entsteht also durch die Reibung der Luftschichten. Und wann blitzt´s? Da nehmen wir wieder unsern Luftballon. Reiben aber jetzt da nicht außen Elektrizität drauf, sondern wir füllen den mit Wasser. Das geht ne Weile gut, aber irgendwann ist der Speicher voll, und dann platzt der Ballon. Man kann auch sagen, der Ballon entlädt sich blitzartig. Und wenn der Himmel keine Elektrizität mehr speichern kann, dann blitzt´s.

Und was passiert genau wenn´s blitzt? Um euch das zeigen zu können mussten wir nach München reisen, zur technischen Universität. Die können Blitze künstlich erzeugen. Und nur solche künstlichen Blitze kann man sich genau angucken. Wollte man darauf in der Natur warten, müsste man Geduld haben. Vielleicht ein Jahr oder noch länger. Und um so einen Blitz zu filmen, braucht man auch so ein Monstrum an Kamera. Das ist so schwer, dass es zwei Leute tragen müssen. Von innen sieht diese Kamera so aus. Die kann einen Blitz 40 000 mal langsamer zeigen, als er in Wirklichkeit passiert. Und diese Geschwindigkeit würde kein Filmmaterial der Welt aushalten. Deswegen bedient man sich eines Tricks. Da ist ein Spiegel, der sich dreht. Und der lenkt das Bild von dem Blitz auf das Filmmaterial um, das außen am Rand eingespannt ist und sich überhaupt nicht bewegt. Da

kommt´s hin. Und damit man mit der Kamera auch den richtigen Punkt trifft, wo´s blitzt, ist da eine Blitzanlage. Von da oben kommt der Blitz. Und der wird dann nachher unten in dieses Haus einschlagen. Sonst in der Natur würd man so nen Blitz ja nie treffen. Man wüßt ja gar nicht wo es blitzt. Vorher noch die Optik rein. Dann kommt ein Signal. Das ist Elektrizität, alle müssen raus. Das ist gefährlich. Und jetzt wird´s spannend. Jetzt muss zuerst mal so viel Elektrizität gespeichert werden, dass es blitzen kann. Das kann man in der Universität München machen. Und wenn dann die Kamera eingeschaltet ist und hochgefahren ist mit ihrer Geschwindigkeit, dann kann man in dem Blitz so tolle Dinge sehen, die man noch nie im Leben gesehen hat. So ein Blitz, der ja so schnell geht, hat eigentlich drei Geheimnisse. Erstes Geheimnis: bevor man den Blitz sieht, ist vorher eine Vorentladung. Die bildet so nen Schlauch. Und dann rutscht der eigentliche Blitz erst nach. Und jetzt beim nächsten mal das zweite Geheimnis: Man denkt ja immer der Blitz saust von oben nach unten. Das ist gar nicht so. Ein Teil des Blitzes kommt von oben. Und wenn er drei Viertel des Weges gegangen ist, kommt von unten einer entgegen. Und dann leuchtet´s erst. Und weil´s so schön war jetzt das dritte Geheimnis: Achtung gleich! Der Blitz rutscht ruckweise nach, nicht in einem schnellen Zug, wie man das immer sieht. Und dann leuchtet´s. Das ist die Hauptentladung. Und genau dasselbe passiert blitzschnell, bei jedem Blitz, immer dann, wenn ein Gewitter ist. Hättet ihr das gedacht?

B

Propositional Analyses

B. 1 German post-war history (Experiment 1)

P1 (werden-sollen $)
P2 (aus P1 Deutschland)
P3 (was P2)

P4 (sein Krieg)
P5 (mod Krieg dieser)
P6 (mod P4 nicht)
P7 (wie P4 P9)
P8 (sein Krieg)
P9 (in P8 Vergangenheit)

P10 (besetzen wer Gebiet)
P11 (mod wer immer)
P12 (auferlegen P10 System)
P13 (mod P12 auch)
P14 (mod auferlegen ihm)
P15 (mod System eigenes)
P16 (mod System gesellschaftliches)
P17 (mod System sein)

P18 (einführen jeder System)
P19 (mod System eigenes)
P20 (mod System sein)
P21 (vordringen-können Armee)
P22 (mod Armee seine)
P23 (mod P21 so-weit)

P24 ($ Stalin)
P25 (mod Stalin 1945)

P26 (machen wir)
P27 (am P26 Tag)
P28 (nach P27 Sieg)
P29 (mod Sieg unser)
P30 (was P27)

P31 (haben-das Einsatz Ziel)
P32 (mod Einsatz unser)
P33 (etablieren Frankreich)
P34 (mod P33 hier)
P35 (auf P33 Dauer)

P36 ($ de-Gaulle)
P37 (mod de-Gaulle 1945)

P38 (empfehlen es P42)
P39 (mod empfehlen sich)
P40 (umwandeln Deutschland)
P41 (in P40 Land)
P42 (mit P41 Charakter)
P43 (mod Charakter land-und-weidewirtschaftlich)
P44 (mod land-und-weidewirtschaftlich vorwiegend)
P44.1 ($ Morgenthau)
P44.2 (mod Morgenthau 1944)

P45 ($ Deutschland)
P46 (nach P45 Kapitulation)

P47 (werden-sollen $)
P48 (aus P47 Deutschland)
P49 (was P48)

P50 (denken-an Deutsche Zukunft)
P51 (mod Deutsche meiste)
P52 (mod P50 nicht)
P53 (in P50 Städte)
P54 (mod Städte zerstörte)

P55 (sorgen-um sie Überleben)
P56 (mod sorgen sich)
P57 (am P55 Tag)
P58 (mod Tag nächster)

P59 (herrschen Hunger)

P60 ($ Hamsterfahrten)
P61 (für P60 Kartoffeln)
P62 (mod Kartoffeln paar)

P63 ($ Tauschgeschäfte)

P64 (bieten ich Markenfüllhalter)
P65 (mod Markenfüllhalter neu)
P66 (suchen ich Briketts)

P67 (bitter-sein Kälte)
P68 (in P67 Behausungen)
P69 (mod Behausungen unbeheizbaren)

P70 (zerstören man Prozent)
P71 (mod Prozent 80)
P72 (mod Prozent bis-zu)
P73 (mod Prozent Häuser-und-Wohnungen)

P74 ($Flüchtlinge-Vertriebene-Kriegsheimkehrer-Kriegsgefangene)
P75 (mod Kriegsgefangene ehemalige)
P76 (in P74 Ruinenlandschaften)

P77 (suchen Menschen Familien)
P79 (mod Familie ihre)

P80 ($ Schweizer)

P81 (an P80 Freund)
P82 (mod Freund sein)
P83 (mod Freund amerikanisch)
P84 (glauben Deutscher nichts)
P85 (mod Deutscher Tage)
P86 (mod Tage unsere)
P87 (hoffen Deutscher nichts)
P88 (und P87 P89)
P89 (dulden Deutscher alles)

P90 (noch-predigen-wollen wer P91)
P91 (und Demokratie Menschenwürde)
P92 (geben der Steine)
P93 (statt P92 Brot)

P94 (untergehen-müssen Deutschland)
P95 (oder P94 P96)
P96 (entstehen Deutschland)
P97 (mod entstehen neu)

P98 (vegetieren-dürfen es)
P99 (mod P98 nicht)
P100 (mod P98 weiter)
P101 (um P98 Europawillen)

P102 ($ Ziele)
P103 (mod Ziele Alliierte)

P104 (beschließen Drei P109)
P105 (mod Drei groß)
P106 (mod Drei Großbritannien-USA-Sowjetunion)
P107 (abtrennen Drei Teile)
P108 (mod Teile deutsches-Reich)
P109 (im P107 Osten)
P110 (vor P104 Kapitulation)
P111 (mod Kapitulation bedingungslos)
P112 (mod Kapitulation Deutschland)
P113 (auf P110 Konferenz-von-Jalta)
P114 (im P113 Februar)
P115 (mod Februar 1945)
P116 (mod P110 schon)

P117 (so-fallen Teil)
P118 (mod Teil Ostpreußen)
P119 (mod Ostpreußen nordöstlich)
P120 (an P117 Sowjetunion)

P121 (stellen $ Rest)
P122 (unter P121 Verwaltung)
P123 (mod Verwaltung polnisch)

P124 ($ $ P126)
P125 (mod P124 ebenso)
P126 (und Pommern Schlesien)
P127 ($-ein Ausgleich)
P128 (dafür P127 P131)
P129 (abgeben-müssen Polen Gebiete)
P130 (mod Gebiete östlich)
P131 (an P129 Sowjetunion)

P132 (festschreiben-sollen $ Westverschiebung)
P133 (mod Westverschiebung diese)
P134 (mod Westverschiebung Polen)
P135 (mod P132 endgültig)
P136 (bis P132 P137)
P137 (und Oder Neiße)
P138 (in P136 Friedensvertrag)
P139 (mod Friedensvertrag später)
P140 (mod Friedensvertrag ein)

P141 (anerkennen Bundesrepublik Grenze)
P142 (mod Grenze diese)
P143 (mod P141 1990)
P144 (mod 1990 erst)

P145 (haben Siegermächte Konzept)
P146 (mod Siegermächte keine)
P147 (mod Konzept wirklich)
P148 (werden-sollen $)
P149 (aus P148 Deutschland)
P150 (was P149)
P151 (trotz P145 Auffassungen)
P152 (mod Auffassungen unterschiedlich)

P153 (einig-sein sie sich)
P154 (mod P153 aber)
P155 (in P153 Punkt)
P156 (mod Punkt vier)
P157 (eingehen Punkt)
P158 (als P157 D)
P159 (in P158 Geschichte)
P160 (mod D vier)
P161 ($ Denazifizierung-Demilitarisierung-Demontage-Demokratisierung)

P162 (sein es Absicht)
P163 (mod P162 nicht)
P164 (vernichten Volk)
P165 (mod Volk deutsch)
P166 (aber P162 P176)
P167 (bestehen Hoffnung)
P168 (mod P167 erst)
P169 (nach P167 P172)
P170 (mod Auslöschung Nazitum)
P171 (mod Auslöschung Militarismus)
P172 (und P170 P171)
P173 (für P169 Deutsche)
P174 (auf P173 Leben)
P175 (mod Leben bescheiden)
P176 (und P174 P178)

P177 (auf P173 Platz)
P178 (in P177 Gemeinschaft)
P179 (mod Gemeinschaft Nationen)
P180 ($ Potsdamer-Konferenz)

P181 ($ Potsdam)
P182 (mod Potsdam 1945)

P183 (zu-Tage-treten Interessen)
P184 (mod Interessen Sieger)
P185 (mod Interessen machtpolitisch)
P186 (mod Interessen gegensätzlich)
P187 (mod zu-Tage-treten offen)
P188 (nach P183 Sieg)
P189 (über P188 Gegner)
P190 (mod Gegner gemeinsam)
P191 (mod Gegner Deutschland)

P192 (verständigen man sich)
P193 (mod P192 dennoch)
P194 (auf P192 Konferenz)
P195 (von P194 Potsdam)
P196 (auf P195 das)
P197 (zerstückeln man Deutschland)
P198 (mod P197 nicht)
P199 (in P197 Teilstaaten)
P200 (mod Teilstaaten mehrere)

P201 (aufteilen man es)
P202 (mod aufteilen nun)
P203 (mod aufteilen so)
P204 (in P201 Besatzungszonen)
P205 (mod Besatzungszonen vier)
P206 (und P204 P208)
P207 (aufteilen man Berlin)
P208 (in P207 Sektoren)
P209 (mod Sektoren vier)

P210 (sein Frankreich Besatzungsmacht)
P211 (mod Besatzungsmacht vierte)
P212 (sein das)
P213 (mod sein vertreten)
P214 (mod P212 nicht)
P215 (in P212 Potsdam)
P216 (mod P213 jedoch)

P217 (sein Beschluss)
P218 (mod P217 wesentlich)
P219 (für P217 Erhalt)
P220 (mod Erhalt Einheit)
P221 (mod Einheit staatlich)
P222 (betrachten Deutschland)
P223 (als P222 Ganzes)
P224 (mod Ganzes wirtschaftlich)
P225 (mod Ganzes ein)
P226 (und P223 P228)
P227 (verwalten Deutschland)
P228 (als P227 Ganzes)

P229 (sein Kontrollrat Regierungsorgan)
P230 (mod Kontrollrat alliierter)
P231 (mod Regierungsorgan oberstes)
P232 (in P229 Berlin)

P233 (regeln-sollen man Angelegenheiten)
P234 (mod P233 hier)
P235 (mod Angelegenheiten alle)
P236 (betreffend P233 Deutschland)
P237 (als P236 Ganzes)

P238 (sein Prozess wichtigste)
P239 (mod P238 wohl)
P240 (von P238 Aktionen)
P241 (mod Aktionen gemeinsam)
P242 (mod Aktionen wenige)
P243 (mod Aktionen Siegermächte)
P244 (gegen P240 Hauptverantwortliche)
P245 (mod Hauptverantwortliche NS-Staat)
P246 (in P244 Nürnberg)

P247 (anklagen man sie)
P248 (mod anklagen Kriegsverbrechen)
P249 (mod anklagen Verbrechen)
P250 (gegen Verbrechen Menschlichkeit)
P251 (und P247 P252)
P252 (verurteilen man sie)

P253 ($ Reparationen)

P254 (beginnen Demontage)
P255 (mod Demontage Industrieanlagen)
P256 (mod Industrieanlagen deutsch)
P257 (und P254 P258)
P258 (beginnen Abtransport)
P259 (im P257 Sommer)
P260 (mod Sommer 1945)

P261 (sein Alliierte)
P262 (mod sein uneins)
P263 (in P261 Frage)
P264 (mod Frage Reparationen)

P265 (verfahren Amerikaner-und-Briten)
P266 (mod verfahren zurückhaltend)
P267 (mod P265 bald)
P268 (mod bald schon)
P269 (bestehen Sowjetunion-und-Frankreich)
P270 (auf P269 Wiedergutmachungen)
P271 (in P270 Umfang)
P272 (mod Umfang groß)
P273 (leiden S-u-F)
P274 (unter P273 Besatzung)

P275 (mod Besatzung deutsch)
P276 (und P274 P277)
P277 (unter P273 Kriegshandlungen)
P278 (mod leiden am-schwersten)
P279 (mod P269 doch)

P280 (durchsetzen Amerikaner P284)
P281 (haben-dürfen Besatzungsmacht Zugriff)
P282 (mod Besatzungsmacht jede)
P283 (mod P281 nur)
P284 (auf P281 Zone)
P285 (mod Zone eigene)
P286 (verwehren $ Sowjets Abtransport)
P287 (von P286 Maschinen)
P288 (und P287 P289)
P289 (von P286 Geräten)
P290 (im P287 Ruhrgebiet)
P291 (und P290 P296)
P292 (erhalten $ Industriegebiet)
P293 (mod Industriegebiet größtes)
P294 (mod Industriegebiet deutsches)
P295 (am P292 Leben)
P296 (mit P295 dem)
P297 (mod Zone ihre)

P298 (ziehen $ Trennungslinie)
P299 (mod Trennungslinie eine)
P300 (mit P298 Entscheidung)
P301 (mod Entscheidung diese)
P302 (mod Entscheidung folgenschwer)
P303 (durch P300 Deutschland)

P304 (sein-der es Anfang)
P305 (mod Anfang Teilung)
P306 (mod Teilung Deutschland)
P307 (in P304 Wirtschaftsgebiete)
P308 (mod Wirtschaftsgebiete zwei)
P309 (und P307 P310)
P310 (in P304 Staaten)
P311 (mod Staaten deutsch)
P312 (mod Staaten zwei)
P313 (mod P310 schließlich)

P314 ($ Wende)
P315 (mod Wende Politik)
P316 (mod Politik amerikanisch)

P317 (heißen Fernziel Weltherrschaft)
P318 (mod Fernziel russisch)

P319 (sein-das Macht Einziges)
P320 (verstehen Russen)
P321 (was P320)

P322 (signalisieren Bemerkung Wende)
P323 (mod Bemerkung diese)
P324 (mod Bemerkung Präsident-Truman)
P325 (mod Präsident-Truman amerikanisch)
P326 (in P322 Politik)
P327 (mod Politik amerikanisch)

P328 ($ Zusammenarbeit)
P329 (mod Zusammenarbeit keine)
P330 (mit P338 Russen)
P331 (sondern P330 P332)
P332 ($ Eindämmung)
P333 (mod Eindämmung weltweit)
P334 (mod Eindämmung Einfluss)
P335 (mod Einfluss kommunistisch)

P336 (sein-ein Westdeutschland Feld)
P337 (mod Feld wichtig)
P338 (mod Feld Außenpolitik)
P339 (mod Außenpolitik amerikanisch)
P340 (mod Westdeutschland Zone)
P341 (mod Zone amerikanisch)
P342 (mod P340 zunächst)

P343 ($ Besatzungszone)
P344 (mod Besatzungszone amerikanisch)

P345 (spüren Deutsche Besatzer)
P346 (in P345 Zone)
P347 (mod Zone amerikanisch)
P348 (mod P345 weniger)
P349 (als P346 P351)
P350 (spüren Menschen Besatzer)
P351 (in P350 Zonen)
P352 (mod Zonen andere)

P353 (verstehen Amerikaner Politik)
P354 (mod Politik ihre)
P355 (als P353 Demokratisierungsprozess)
P356 (vor P355 allem)
P357 (mod Demokratisierungsprozess Jugend)
P358 (verstehen Amerikaner Politik)
P359 (mod P358 auch)
P360 (mod P358 aber)
P361 (durch P358 Presse)
P362 (durch P358 Rundfunk)
P363 (und P361 P362)

P364 (entfernen man Lehrer)
P365 (mod Lehrer belastet)
P366 (mod belastet nationalsozialistisch)
P367 (an P364 Schulen)
P368 (und P367 P369)
P369 (einführen man Schülermitverwaltung)

P370 (achten Presseoffiziere)
P371 (in P370 Medien)
P372 (mod Presseoffiziere amerikanisch)

P373 (auf P371 Trennung)
P374 (mod Trennung sauber)
P375 (von P373 P376)
P376 (und Nachricht Kommentar)
P377 (und P375 P378)
P378 (sorgen-für Presseoffiziere Darstellung)
P379 (mod Darstellung positiv)
P380 (mod Darstellung Staats- und Gesellschaftsform)
P381 (mod Staats- und Gesellschaftsform freiheitlich)

P382 (sein Kritik willkommen)
P383 (sein Kritik willkommen)
P384 (mod P383 nicht)
P385 (mod P384 allerdings)
P386 (an P383 Besatzungsmacht)

P387 (anstreben Amerikaner Struktur)
P388 (mod Struktur föderal)
P389 (für P387 Zone)
P390 (mod Zone ihre)
P391 (nach P389 Vorbild)
P392 (mod Vorbild USA)
P393 (bestimmen Bundesstaaten Politik)
P394 (mod Bundesstaaten stark)
P395 (wo P393)

P396 (gliedern sie Zone)
P397 (mod P396 so)
P398 (mod Zone ihre)
P399 (in P396 Verwaltungseinheiten)
P400 (mod Verwaltungseinheiten vier)
P401 (werden-sollen Bundesländer)
P402 (aus P401 denen)
P403 (mod Bundesländer Bayern-Baden-Württemberg-Baden-Hessen-Bremen)
P404 (mod P401 bald)
P405 (mod bald schon)

P406 (zulassen sie Gemeinderats-und-Bürgermeisterwahlen)
P407 (mod P406 schon)
P408 (im P406 Januar)
P409 (mod Januar 1946)
P410 (mod Gemeinderats-und-Bürgermeisterwahlen erste)

P411 (gehören es P420)
P412 (zu P411 Demokratisierung)
P413 (mod Demokratisierung diese)
P414 (mod Demokratisierung stufenweise)
P415 (von P412 unten)
P416 (übertragen man Aufgaben)
P417 (mod P416 bald)
P418 (mod bald sehr)
P419 (von P416 Militärregierung)
P420 (auf P419 Verwaltungen)
P421 (mod Verwaltungen deutsch)
P422 (mod P416 schon)

P423 (so-sein Amerika)
P424 (mod sein bemüht)
P425 (mod P423 auch)
P426 (zu P423 Wahrung)
P427 (mod Wahrung Interessen)
P428 (mod Interessen eigene)
P429 (von P426 Anfang)
P430 (mod P429 an)
P431 (etablieren Amerika Deutschland)
P432 (mod Deutschland demokratisch)

P433 (helfen-wollen Volk1 Volk2)
P434 (mod Volk1 amerikanisch)
P435 (mod Volk2 deutsch)
P436 (zurückfinden Volk2 Weg)
P437 (mod Weg sein)
P438 (zu P436 Platz)
P439 (mod Platz ein)
P440 (mod Platz ehrenvoll)
P441 (unter P438 Völker)
P442 (mod Völker frei)
P443 (mod Völker friedliebend)
P444 (mod Völker Welt)

P445 ($ Brynes)
P446 (mod Brynes Stuttgart-Rede)
P447 (mod Stuttgart-Rede 1945)

P448 ($ Besatzungszone)
P449 (mod Besatzungszone britisch)

P450 (sehen Großbritannien Ziel)
P451 (mod Ziel dringlichstes)
P452 (in P450 dem)
P453 (zerschlagen $ Militarismus)
P454 (mod Militarismus preußisch)

P455 (erscheinen es)
P456 (mod P455 wichtiger)
P457 (mod P455 bereits)
P458 (mod P455 doch)
P459 (nach P455 Kriegsende)
P460 (mod nach Jahr)
P461 (mod Jahr ein)
P462 (stabilisieren $ Deutschland)
P463 (als P462 Bollwerk)
P464 (gegen P463 Kommunismus)

P465 (sein-die Gefahr)
P466 (mod sein groß)
P467 (mod groß genauso)

P468 (mod Gefahr russisch)
P469 (mit P465 Sicherheit)
P470 (sein-die Gefahr)
P471 (mod sein noch-größer)
P472 (mod P470 möglicherweise)
P473 (als P470 P474)
P474 (sein-die Gefahr)
P475 (mod Gefahr Deutschland)
P476 (mod Deutschland wiedererstarkt)

P477 ($ Bevin)

P478 (so-beginnen Engländer P484)
P479 (in P478 Zone)
P480 (mod Zone ihre)
P481 (wie P479 Amerikaner)
P482 (einsetzen Engländer sich)
P483 (mod P482 nachhaltiger)
P484 (für P482 Wiederaufbau)

P485 (erhöhen man Lebensmittelrationen)
P486 (und P485 P487)
P487 (drosseln man Reparationen)

P488 (tolerieren Briten Parteien)
P489 (mod Parteien deutsch)
P490 (mod Parteien gegründet)
P491 (mod gegründet neu)
P492 (aus P488 Tradition)
P493 (mod Tradition demokratisch)

P494 (behalten Militärregierung Fäden)
P495 (mod Militärregierung britisch)
P496 (mod Fäden alle)
P497 (mod P494 aber)
P498 (im P494 Gegensatz)
P499 (zu P498 Amerikaner)
P500 (in P499 Hand)

P501 (geben sie Spielraum)
P502 (mod Spielraum gering)
P503 (wie P501 P505)
P504 (sein es)
P505 (in P504 Verwaltungspraxis)
P506 (mod Verwaltungspraxis bewährt)
P507 (mod Verwaltungspraxis Kolonien)
P508 (mod Kolonien britisch)
P509 (mod gering nur)
P510 (zur P501 Eigenverantwortlichkeit)

P511 ($ Besatzungszone)
P512 (mod Besatzungszone französisch)

P513 (sein Deutschlandpolitik)
P514 (mod sein bestimmt)
P515 (mod Deutschlandpolitik Frankreich)
P516 (von P513 Sicherheitsbedürfnis)
P517 (mod Sicherheitsbedürfnis stark)

P518 (sein Ursache Erfahrungen)
P519 (mod Erfahrungen leidvoll)
P520 (mod Erfahrungen mehrfach)
P521 (mit P518 Nachbar)
P522 (mod Nachbar aggressiv)
P523 (mod Nachbar Deutschland)

P524 (wünschen Frankreich Deutschland)
P525 (mod Deutschland zerstückelt)
P526 (aus P524 Einzelstaaten)
P527 (mod Einzelstaaten schwach)
P528 (mod schwach wirtschaftlich)
P529 (und P526 P530)
P530 (verhindern Frankreich Einrichtung)
P531 (mod Einrichtung jede)
P532 (mod Einrichtung zentral)
P533 (mod Einrichtung zonenübergreifend)

P534 (fordern es Kontrolle)
P535 (mod P534 zudem)
P536 (mod Kontrolle international)
P537 (mod Kontrolle Ruhrgebiet)
P538 (mod Kontrolle Waffenschmiede)
P539 (mod Waffenschmiede einstig)
P540 (mod Waffenschmiede Deutschland)

P541 (abkapseln Frankreich Besatzungszone)
P542 (mod Besatzungszone seine)
P543 (im P541 Widerspruch)
P544 (zu P543 Abkommen)
P545 (mod Abkommen Potsdamer)
P546 (von P544 anderen)
P547 (und P546 P549)
P548 (eingliedern Frankreich Saargebiet)
P549 (in P548 Wirtschafts- und Währungsgebiet)
P550 (mod Wirtschafts- und Währungsgebiet Frankreich)

P551 (sprechen Churchill)
P552 (von P551 Vorhang)
P553 (mod Vorhang seiden)
P554 (mod Vorhang ein)
P555 (trennen der Zone)
P556 (mod Zone französisch)
P557 (von P555 Besatzungszonen)
P558 (mod Besatzungszonen andere)

P559 (spüren-bekommen Bevölkerung Besatzungspolitik)
P560 (mod Besatzungspolitik hart)
P561 (mod Besatzungspolitik Franzosen)
P562 (im P559 Alltag)
P563 (mod Alltag Nachkriegszeit)

P564 (versuchen Frankreich)
P565 (mod P564 gleichzeitig)
P566 (mod Frankreich auch)
P567 (wie P564 USA)
P568 (mod wie ähnlich)
P569 (durch P567 Schul-und-Kulturpolitik)
P570 (mod Schul-und-Kulturpolitik gezielt)
P571 (gewinnen Frankreich Jugend)
P572 (mod P571 gerade)
P573 (und P571 P574)
P574 (einleiten Frankreich Demokratisierungsprozess)
P575 (mod Demokratisierungsprozess langfristig)

P576 (ebnen Politik Weg)
P577 (mod Politik diese)
P578 (mod P576 schließlich)
P579 (für P576 Verständigung)
P580 (mod Verständigung deutsch-französisch)
P581 (mod Verständigung spätere)

P582 ($ Besatzungszone)
P583 (mod Besatzungszone russisch)

P584 (sein Sowjetunion)
P585 (mod sein interessiert)
P586 (wie P584 Frankreich)
P587 (mod wie zunächst)
P588 (an P586 Zusammenarbeit)
P589 (mod Zusammenarbeit umfassend)
P590 (mod umfassend Zonen)
P591 (mod Zonen alle)
P592 (im P588 Interesse)
P593 (mod Interesse Gesamtdeutschland)
P594 (mod interessiert nicht)
P595 (wenn P592 P603)
P596 (mod wenn auch)
P597 (sein interessiert)
P598 (mod interessiert nicht)
P599 (zum P597 Teil)
P600 (aufgrund P599 Überlegungen)
P601 (mod Überlegungen politisch)
P602 (mod Überlegungen andere)
P603 (als P600 Frankreich)

P604 (beginnen Sowjetunion P609)
P605 (auf P604 Faust)
P606 (mod Faust eigene)
P607 (durchsetzen Sowjetunion Veränderungen)
P608 (mod Veränderungen tiefgreifend)
P609 (in P607 Besatzungszone)
P610 (mod Besatzungszone ihre)

P611 (besetzen man Schlüsselpositionen)
P612 (mod besetzen nach-und-nach)
P613 (mit P611 Kader)
P614 (mod Kader sowjetisch)
P615 (und P613 Gruppe)
P616 (mod Gruppe Exilkommunisten)
P617 (mod Exilkommunisten deutsch)
P618 (mod Gruppe eine)
P619 (mod Kader 60000)
P620 (in P615 Länder)
P621 (mod Länder gegründet)
P622 (mod gegründet rasch)
P623 (mod Länder Mecklenburg-Brandenburg-Sachsen-Anhalt-Thüringen-und-Sachsen)

P624 (sein Entnazifizierung Instrumente)
P625 (sein Bildungsreform Instrumente)
P626 (und P624 P625)
P627 (mod Instrumente Umwälzungen)
P628 (mod Umwälzungen ideologisch)
P629 (mod ideologisch gesellschaftlich)
P630 (in P626 Zonen)
P631 (mod P630 anders)
P632 (mod anders als)
P633 (mod Zonen westlich)
P634 (mod Zonen drei)
P635 (mod Zonen andere)
P636 (durchdringen-sollen Instrumente Lebensbereiche)
P637 (mod Lebensbereiche alle)

P638 (sein Enteignungen Anfänge)
P639 (sein Verstaatlichungen Anfänge)
P640 (und P638 P639)
P641 (von P640 Banken)
P642 (von P640 Industrien)
P643 (mod Anfänge Wirtschaft)
P644 (mod Wirtschaft zentralgelenkt)
P645 (mod Wirtschaft staatsmonopolistischen)

P646 ($ Bodenreform Vorstufe)
P647 (zur P646 Kollektivierung)
P648 (mod Kollektivierung Landwirtschaft)

P649 (geschehen Vereinigung)
P650 (mod geschehen freiwillig)
P651 (mod P649 nicht)
P652 (von Vereinigung SPD-und-KPD)
P653 (zur SPD-und-KPD SED)
P654 (feiern Kommunistische Vereinigung)
P655 (als P654 Bündelung)
P656 (mod Bündelung Kräfte)
P657 (mod Kräfte proletarisch)
P658 (sondern P649 P660)
P659 (erzwingen KPD)
P660 (durch P659 Pression)

P661 (so-äußern KPD-Führer-Walter-Ulbricht P664)
P662 (aussehen-müssen es)
P663 (mod P662 demokratisch)
P664 (aber P662 P666)
P665 (haben-müssen wir alles)
P666 (in P665 Hand)

P667 ($ Wiederaufbauhilfe)
P668 (mod Wiederaufbauhilfe USA)
P669 ($ Marshall-Plan)
P670 ($ Bi-Zone)
P671 (und P669 P670)

P672 (verhindern Interessen Absicht)
P673 (mod Interessen unterschiedlich)
P674 (mod Interessen Siegermächte)
P675 (mod Siegermächte vier)
P676 (mod Absicht ursprünglich)
P677 (mod Absicht ihre)
P678 (regieren Siegermächte Deutschland)
P679 (als P678 Ganzes)

P680 (tiefer-werden Graben)
P681 (zwischen P680 Ostzone Westzonen)
P682 (mod Ostzone geführt)
P683 (mod geführt kommunistisch)
P684 (mod Westzonen geführt)
P685 (im P684 Sinne)
P686 (mod Sinne Demokratie)
P687 (mod Demokratie westlich)

P688 (sein es)
P689 (mod P688 folgerichtig)
P690 (vor P688 Hintergrund)
P691 (mod Hintergrund Ost-Westkonflikt)
P692 (mod Ost-Westkonflikt verschärfend)
P693 (mod verschärfend sich)
P694 (und P690 P695)
P695 (angesichts P688 Not)
P696 (mod Not wirtschaftlich)
P697 (mod Not wachsend)
P698 (mod P688 nur)
P699 (dass P694 P700)
P700 (zusammenrücken Zone)
P701 (mod P700 näher)
P702 (mod Zone amerikanische)
P703 (mod Zone britische)

P704 (miteinbeziehen Amerikaner Westzonen)
P705 (in P704 CARE-Programm)

P706 ($ Überlebenshilfe)
P707 (für P706 Bevölkerung)
P708 (mod Bevölkerung hungernd)
P709 (wie P707 P710)
P710 ($ Schulspeisung)
P711 (für P710 Kinder)
P712 (mod Kinder 3,5-Millionen)
P713 (in P711 Zone)
P714 (mod Zone britisch)
P715 (mod Zone amerikanisch)
P716 (mod P710 ebenso)

P717 (zusammenschließen Zone sich)
P718 (mod Zone britisch)
P719 (mod Zone amerikanisch)
P720 (zur P717 Bi-Zone)
P721 (mod P717 1947)
P722 (mod 1947 Anfang)

P723 (einschwenken-sollen Frankreich)
P724 (mod P723 später)
P725 (mod später Jahre)
P726 (mod Jahre zwei)
P727 (mod P723 erst)
P728 (mod P723 auch)
P729 (auf P723 Linie)
P730 (mod Linie diese)
P731 (mod Linie gemeinsame)

P732 (setzen $ Deutsche)
P733 (in P732 Lage)
P734 (mod Lage arbeiten)

P735 (fordern Journalisten P732)
P736 (mod Journalisten amerikanisch)
P737 (fordern Kongressabgeordnete)
P738 (und P735 P737)

P739 (erweisen Marshall-Plan sich)
P740 (mod P739 dafür)
P741 (als P739 Instrument)
P742 (mod Instrument geeignet)
P743 (ausdehnen man Marshall-Plan)
P744 (auf P743 Westzonen)
P745 (mod Westzonen drei)
P746 (mod Westzonen alle)

P747 (fließen US-Dollar)
P748 (mod P747 allein)
P749 (aus P747 Aufbauhilfe)
P750 (für P749 Europa)
P751 (mod Europa ganz)
P752 (mod US-Dollar 1,5-Milliarden)
P753 (in P750 Wirtschaft)
P754 (mod Wirtschaft westdeutsche)
P755 (mod Aufbauhilfe diese)

P756 (werden Marshall-Plan)
P757 (zur P756 Grundlage)
P758 (mod Grundlage Nachkriegswohlstand)

P759 (mod Nachkriegswohlstand westdeutsch)

P760 (ablehnen Sowjetunion Marshall-Plan-Hilfe)
P761 (für P760 Zone)
P762 (mod Zone ihre)

P763 ($ Entstehung)
P764 (mod Entstehung Staaten)
P765 (mod Staaten deutsch)
P766 (mod Staaten zwei)

P767 (verlassen Sowjetunion Kontrollrat)
P768 (im P767 März)
P769 (mod März 1948)
P770 (mod Kontrollrad alliiert)
P771 (zusammentreten der)
P772 (mod P771 danach)
P773 (mod zusammentreten wieder)
P774 (mod wieder nie)

P775 (beendet-sein Vier-Mächte-Verwaltung)
P776 (mit P775 dem)
P777 (mod Vier-Mächte-Verwaltung Gesamtdeutschland)
P778 (mod beendet-sein endgültig)

P779 (einleiten Sowjetunion Schritte)
P780 (einleiten Westalliierten Schritte)
P781 (und P779 P780)
P782 (mod P781 längst)
P783 (mit P782 Verbündete)
P784 (mod Verbündete ihre)
P785 (mod mit zusammen)
P786 (mod Schritte erste)
P787 (zur P783 Gründung)
P788 (mod Gründung Staaten)
P789 (mod Staaten deutsch)
P790 (in P787 Besatzungsgebiet)
P791 (mod Besatzungsgebiet ihr)

P792 (ziehen man Grenze)
P793 (in P792 Deutschland)
P794 (im P793 Zuge)
P795 (mod Zug Krieg)
P796 (mod Krieg kalt)
P797 (zwischen P794 Machblöcken)
P798 (mod Machtblöcke zwei)
P799 (mod Machtblöcke verfeindet)

P800 (senken Vorhang sich)
P801 (mod Vorhang eisern)
P802 (sagen Churchill P800)

P803 (sein-die Gründung Konsequenz)
P804 (mod Gründung Staaten)
P805 (mod Staaten deutsch)
P806 (mod Staaten zwei)
P807 (mod P803 1949)
P808 (aus P803 Jahre)
P809 (mod Jahre vier)
P810 (mod Jahre Besatzungsherrschaft)
P811 (in P808 Deutschland)

P812 (so-sein Staaten)
P813 (mod sein geprägt)
P814 (mod Staaten beide)
P815 (in P812 Grundzüge)
P816 (mod Grundzüge ihre)
P817 (von P815 Siegermächten)
P818 (sein Systeme)
P819 (mod sein vorstellbar)
P820 (mod sein kaum)
P821 (mod sein gegensätzlicher)
P822 (mod Systeme politisch)

P823 (prophezeihen Churchill P828)
P824 (mod P823 1946)
P825 (mod 1946 schon)
P826 (sehen-müssen wir Tatsache)
P827 (ins P826 Auge)
P828 (dass P827 P832)
P829 (sein Deutschland)
P830 (mod Deutschland zwei)
P831 (im P829 Entstehen)
P832 (so-wie P831 P833)
P833 (stehen Dinge)
P834 (mod P833 gegenwärtig)

P835 (sein eins)
P836 (mod P835 organisiert)
P837 (im P835 Interesse)
P838 (mod Interesse russisch)
P839 (mod P835 mehr-oder-weniger)
P840 (organisiert-sein anderes)
P842 (nach P840 dem)
P843 (mod dem Demokratie)
P844 (mod Demokratie westlich)

P845 ($ Ausblick)
P846 ($ Einheit)
P847 (mod Einheit deutsch)
P848 (mod Einheit 1990)

P849 (werden Grenze)
P850 (durch P849 Deutschland)
P851 (für P850 Jahre)
P852 (mod Jahre 40)
P853 (zur P851 Todeszone)
P854 (zwischen P853 Machtblöcke)
P855 (mod Machtblöcke zwei)
P856 (mod Machtblöcke verfeindet)

P857	(schaffen Veränderungen Voraussetzungen)
P858	(im Veränderungen Ostblock)
P859	(schaffen Aufbegehren Voraussetzungen)
P860	(mod Aufbegehren Menschen)
P861	(in Menschen DDR)
P862	(und P858 P859)
P863	(für P862 Vereinigung)
P864	(mod Vereinigung Deutschland)
P865	(unter P863 Vorzeichen)
P866	(mod Vorzeichen demokratisch)
P867	(mod P857 erst)

Note:

$ = missing argument in the text

mod = modified (i.e., specifies a preceding argument or sentence)

B.2 Thermos (Experiment 2 and 3)

P1 (sein-ein Glas)
P2 (mod Glas Tee)
P3 (mod Glas schön)

P4 (trinken-können Christoph $)
P5 (in P4 Ruhe)
P6 (mod Ruhe alle)

P7 ($ heiß)

P8 (anfassen-können er $)
P9 (mod P8 kaum)

P10 (sein-das Telefon)
P11 (mod P10 nicht)
P12 (wenn P10)

P13 (dran-sein einer)
P14 (erzählen der)
P15 (erzählen $)
P16 (erzählen $)
P17 (und P14 P15 P16)
P18 (mod P13 wieder)

P19 (beenden man Gespräch)
P20 (mod P19 dann)
P21 (wenn P22 P19)
P22 (kaltwerden Tee)

P23 (verhindern-können man das)
P24 (wie P23)

P25 (denken Christoph P26)
P26 (rauskriegen-müssen man das)
P27 (mod P26 doch)

P28 (rauskriegen man das)
P29 (mit P28 Topf)
P30 (mod Topf voll)
P31 (mit P29 Wasser)
P32 (mod Wasser heiß)
P33 (Frage P31)

P34 (drin-sein-die Wärme)
P35 (mod P34 da)
P36 (mod P34 jetzt-noch)
P37 (mod P34 klar)

P38 (reinhängen er noch-was)
P39 (mod P38 da)
P40 (mod P38 dann)
P41 ($ Stoppuhr)
P42 (mod Stoppuhr eine)

P43 (testen-wollen er P44)
P44 (wegnehmen Material Wärme)
P45 (mod Material welches)
P46 (mod Wärme am-wenigsten)

P47 ($ Fingertest)
P48 (nach P47 Sekunden)
P49 (mod Sekunden dreißig)

P50 ($ Löffel)
P51 (aus P50 Metall)
P52 (mod P50 heiß)

P53 (gehen Glas)
P54 (mod P53 so)

P55 (reinwandern Wärme)
P56 (in P55 Gummi)
P57 (mod Wärme gar-keine)
P58 (mod gar-keine fast)

P59 (kalt-bleiben Holz)

P60 (sein das es)

P61 (wegleiten das Wärme)
P62 (mod P61 am-wenigsten)
P63 (mod P61 wohl)

P64 (sein-ein Christoph Mann)
P65 (mod Mann Tat)

P66 (greifen er)
P67 (zu P66 Bohrmaschine)
P68 (mod P66 gleich)
P69 (und P66 P70)
P70 (bohren er Riesenloch)
P71 (in P70 Holzklotz)

P72 (aussehen $)
P73 (mod aussehen praktisch)
P74 (mod P72 irgendwie)

P75 (sein-sollen es)
P76 (was P75)

P77 (sein-sollen es Teetasse)
P78 (aus P77 Holz)

P79 (drinbleiben-müssen Wärme)
P80 (mod P79 doch)
P81 (mod P79 länger)
P82 (mod P79 da)

P83 ($ heiß)

P84 ($ heiß)
P85 (mod P84 lange)
P86 (mod lange nicht)
P87 (mod P84 aber)

P88 (finden Christoph P89)
P89 (gehen-müssen das)
P90 (mod P89 noch-besser)
P91 (überlegen)
P92 (mod P91 mal)

P93 (sein der)
P94 (aus der Holz)
P95 (mod P93 bester)
P96 (mod P93 schon)
P97 (von P93 Gegenstand)
P98 (mod Gegenstand vier)

P99 (geben es etwas)
P100 (mod P99 vielleicht)
P101 (mod P99 aber)
P102 (wegnehmen das)
P103 (von P102 Wärme)
P104 (mod P102 noch-weniger)

P105 ($ Versuch)
P106 (mod P105 neu)

P107 ($ Wasser)
P108 (mod Wasser heiß)
P109 (mod P107 wieder)

P110 ($ Gefäß)
P111 (mod Gefäß zwei)

P112 (vergleichen-wollen Christoph was)
P113 (mod P112 wohl)

P114 (legen er Glas)
P115 (auf P114 Gefäß)
P116 (mod Gefäß ein)
P117 (und P115 P118)
P118 ($ noch-eins)
P119 (mod P118 dann)

P120 (legen er Glas)
P121 (auf P120 Gefäß)
P122 (mod Gefäß anderes)
P123 (mod P120 auch)
P124 (mod Glas ein)
P125 (legen er Korkstückchen)
P126 (mod P125 dann)
P127 (und P125 P128)
P128 (legen er Glas)
P129 (mod Glas zweites)
P130 (mod P128 erst)
P131 (mod erst dann)

P132 (dazwischen-sein Luft)
P133 (mod P132 also)

P134 ($ Buttertest)
P135 (mod P134 jetzt)

P136 (schmelzen Butter)
P137 (mod P136 zuerst)
P138 (wo P136)

P139 (schmelzen Butter)
P140 (mod P139 da)
P141 (mod P139 natürlich)
P142 (leiten man Wärme)
P143 (mod P142 am-besten)
P144 (aus P142 Wasser)
P145 (mod Wasser heiß)
P146 (zu P142 Butter)
P147 (wo P142)

P148 (schmelzen $ sie)
P149 (mod P148 links)
P150 (und P148 P)
P151 (durchkommen Wärme)
P152 (mod P151 nicht)
P153 (wo P151 P154)
P154 (dazwischen-sein Luft)
P155 (mod wo da)
P156 (und P148 P153)

P157 ($ Glas)
P158 (mod Glas ein)
P159 ($ Glasstückchen)
P160 (mod Glasstückchen ein-paar)
P161 (als P159 Abstandhalter)
P162 (zum P161 Boden)
P163 (und P162 P64)
P164 ($ Teeglas)
P165 (mod P164 darein)
P166 (mod P164 jetzt)

P167 (umgeben-sein $)
P168 (von P167 Luft)
P169 (mod P167 nur)
P170 (mod P167 jetzt)

P171 (gucken P173)
P172 (mod P171 mal)
P173 (funktionieren das)
P174 (mod P173 ob)

P175 ($ Test)

P176 (mit P175 Vergleichsteeglas)
P177 (und P176 P178)
P178 (Anfangstemperatur)

P179 ($ $ Grad)
P180 (mod Grad 80-81-82)
P181 (mod 82 gut)

P182 ($ $ Grad)
P183 (mod P182 auch)
P184 (mod P182 da)
P185 (mod Grad 80-81-82)

P186 (laufen Zeit)
P187 (mod P186 also)

P188 ($ $ Stop)
P189 (nach P188 Viertelstunde)
P190 (mod Viertelstunde ein)
P191 (und P189 P192)
P192 (nachmessen $)

P193 ($ $ Grad)
P194 (mod Grad 46)
P195 (ohne P193 Luftpolster)

P196 ($ $ Grad)
P197 (bei P196 Glas)
P198 (im P197 Glas)
P199 (mit P198 Luft)
P200 (mod P199 dazwischen)
P201 (mod Grad 60)

P202 ($ Unterschied)
P203 (mod Unterschied schön)
P204 (mod schön ganz)

P205 (wegwandern-können Wärme)
P206 (mod P205 also)
P207 (mod P205 mehr)
P208 (mod mehr nicht)
P209 (nach P205 unten)
P210 (und P209 P211)
P211 (zur P205 Seite)

P212 (wegwandern-können Wärme)
P213 (nach P212 oben)
P214 (mod P213 aber)

P215 (weggehen Wärme)
P216 (mod P215 noch)
P217 (mod P215 da)

P218 (weggehen Wärme)
P219 (mod Wärme viel)
P220 (mod viel schön)

P221 (mod schön ganz)
P222 (mod P218 zwar)
P223 (denn P218 P225)
P224 (wollen Wärme)
P225 (nach P224 oben)
P226 (mod P224 immer)

P227 (aufhalten-können man die)
P228 (wie P227)

P229 ($ $ Test)
P230 (mod Test zweiter)
P231 (mit P229 Vergleichsteeglas)

P232 ($ $ Test)
P233 (mit P232 Deckel)
P234 (mod P232 diesmal)

P235 ($ Deckel)
P236 (mod Deckel ein)
P237 (mod P236 nur)
P238 (mod nur nicht)
P239 (mod P235 aber)

P240 ($ einer)
P241 (für P240 Glas)
P242 (mod Glas jedes)
P243 (bei P241 Glas)
P244 (im P243 Glas)

P245 (messen $)
P246 (mod P245 wieder)
P247 (nach P245 Viertelstunde)
P248 (mod Viertelstunde ein)

P249 (sein $ Grad)
P250 (mod Grad 45-46)
P251 (mod 45-46 so)
P252 (im P249 Vergleichsteeglas)
P253 (überraschen das)
P254 (mod P253 nicht)
P255 (mod P253 jetzt)

P256 (gucken $ P258)
P257 (bringen Deckel)
P258 (was P257)
P259 (mod P256 mal)
P260 (mod mal jetzt)

P261 ($ Grad)
P262 (mod Grad 70-71-72)

P263 (mehr-sein $ Grad)
P264 (mod Grad 12)
P265 (als P263 P266)
P266 ($ Deckel)

P267 (mod Deckel ohne)
P268 (mod P256 also)

P269 (heiß-sein Tee)
P270 (mod P269 da)
P271 (mod heiß noch)
P272 (mod sein Grad)
P273 (mod Grad 72)

P274 (stolz-sein Christoph)
P275 (mod P274 aber)
P276 (mod P274 trotzdem)
P277 (auf P274 Tasse)
P278 (mod Tasse sein)
P279 (mod Tasse neu)

P280 (geben es P282)
P281 (mod P280 vielleicht)
P282 (verbessern $ noch-was)
P283 (mod P280 aber)

P284 (haben er Idee)
P285 (mod Idee eine)

P286 ($ Alufolie)
P287 (mod Alufolie silbrig-spiegelnd)

P288 (geben $ es)
P289 (in P288 Haushalt)
P290 (mod Haushalt jeder)
P291 (mod jeder fast)

P292 (warm-halten $ Essen)
P293 (mit P292 dem)
P294 (mod P292 auch)
P295 (mod P292 schon)
P296 (mod P292 mal)

P297 (umwickeln-können man Teeglas)
P298 (mod P297 natürlich)
P299 (mod P297 auch)
P300 (mit P297 dem)
P301 (mod Teeglas ein)

P302 (einpacken)

P303 (haben-sollen das Sinn)
P304 (mod Sinn welchen)
P305 (mod P303 aber)
P306 (Frage P303)

P307 (nehmen Christoph Wärmelampe)
P308 (verstehen $ das)
P309 (um-zu P307 P308)
P310 (mod Wärmelampe eine)
P311 (nehmen Christoph Glas)
P312 (ohne P311 Alu)
P313 (mod Glas eins)
P314 (nehmen Christoph Glas)
P315 (mit P314 Alu)
P316 (mod Glas eins)
P317 (und P311 P315)

P318 ($ Buttertest)
P319 (mod P318 jetzt)
P320 (mod P318 wieder)

P321 ($ Ergebnis)
P322 (mod Ergebnis klar)

P323 (ganz-bleiben Butter)
P324 (mit P323 Alufolie)
P325 (weil P324 P326)
P326 (zurückspiegeln Alufolie Wärme)
P327 (mod P326 wieder)

P328 (reinkommen nichts)
P329 (in P328 Glas)

P330 (reinlassen was Wärme)
P331 (mod P330 nicht)
P332 (rauslassen was sie)
P333 (mod P332 nicht)
P334 (mod P332 auch)

P335 (anfassen-können-müssen man Glas)
P336 (mod P335 dann)
P337 (mod P335 jetzt)
P338 (mod P335 doch)
P339 (mod P335 gut)
P340 (mit P335 Tee)
P341 (mod Tee heiß)

P342 (wenn P340 P343)
P343 (drumsein-die Alufolie)

P344 ($ Alufolie)
P345 (mod P344 ohne)

P346 ($ vorsicht)

P347 ($ ohne)

P348 (nehmen-können man es)
P349 (mod P348 aber)
P350 (mit P348 Alufolie)
P351 (in P350 Hand)
P352 (mod P348 bequem)
P353 (mod bequem ganz)

P354 (sein-noch Alufolie Bereicherung)
P355 (mod P354 also)

P356 (für P354 Tassenerfindung)
P357 (mod Tassenerfindung Christophs)
P358 (mod Tassenerfindung neu)

P359 (rein-$ $)
P360 (mit P359 dem)
P361 (in P360 Glas)
P362 (mod Glas anderes)
P363 (mit P361 Korkstücken)
P364 (als P363 Abstandhalter)
P365 (für P364 Luftpolster)

P366 (rein-$ $ Tee)

P367 (drauf-$ $ Doppeldeckel)
P368 (mod P367 dann-noch)
P369 (mit P367 Strohalm)
P370 (mod Doppeldeckel praktisch)

P371 ($ super)

P372 (heiß-halten das Tee)
P373 (mod das alles)
P374 (mod P372 jetzt)
P375 (mod heiß schön)

P376 ($ heiß)

P377 (stolz-sein er)
P378 (mod P377 trotzdem)
P379 (mod P377 aber)
P380 (auf P377 Teewarmhaltetasse)
P381 (mod Teewarmhaltetasse seine)
P382 (mod Teewarmhaltetasse neue)

P383 (wissen du)
P384 (mod du Christoph)
P385 (mod P383 aber)
P386 (Frage P383)

P387 (geben es das)
P388 (mod P387 schon)

P389 (hören $)
P390 (mod P389 schon)
P391 (mod P389 mal)
P392 (Frage P389)

P393 ($ Thermoskanne)

P394 (warm-halten die Tee)
P395 (mod P394 auch)
P396 (mod warm schön)

P397 (funktionieren $)
P398 (mod funktionieren ähnlich)
P399 (mod ähnlich ganz)
P400 (wie P397 Erfindung)
P401 (mod Erfindung seine)

P402 (aufschneiden wir eine)
P403 (mod P402 mal)

P404 ($ Flasche)
P405 (mod P405 auch)
P406 (mod Flasche eine)
P407 (in P404 Flasche)

P408 (sein Flasche)
P409 (mod sein verspiegelt)
P410 (mod verspiegelt innen)
P411 (und P408 P)
P412 (sein Luft)
P413 (mod P412 ringsum)
P414 (wegleiten die Wärme)
P415 (mod P414 nicht)

P416 (trinken Christoph)
P417 (mod P416 am-liebsten)
P418 (mod P416 aber)
P419 (mod P416 natürlich)
P420 (aus P416 Teewarmhaltetasse)
P421 (mod Teewarmhaltetasse seine)
P422 (mod Teewarmhaltetasse eigene)

P423 (heiß-bleiben Tee)
P424 (mod P423 wunderbar)
P425 (in P423 der)
P426 (mod P423 auch)

P427 ($ heiß)

P428 (lernen er es)
P429 (mod P428 nicht)

P430 (sein Tee köstlich)
P431 (mod köstlich trotzdem)

Note:

$ = missing argument in the text

mod = modified (i.e., specifies a preceding argument or sentence

B.3 Lightning (Experiment 2 and 3)

P1 (erleben ihr P2)
P2 (blitzen es)
P3 (mod ihr alle)
P4 (mod P1 mal)
P5 (mod mal schon)

P6 (was P9)
P7 (haben-zu-tun Blitz)
P8 (mod Blitz dieser)
P9 (mit P7 Luftballon)
P10 (mod Luftballon ein)
P11 (reiben man den)
P12 (mit P11 Wollpullover)
P13 (mod Wollpullover ein)
P14 (und P12 P15)
P15 (anziehen der Papierkonfetti)
P16 (mod P15 danach)

P17 (sein-das beides gleiches)
P18 (mod P17 ungefähr)

P19 (wie P20)
P20 (bitte)

P21 (sein-ein Blitz)
P22 (was P21)
P23 (und P22 P25)
P24 (blitzen es)
P25 (warum P24)

P26 (haben-die jeder Zutaten)
P27 (mod haben zu-Hause)
P28 (mod Zutaten dazu)

P29 (nehmen ich Luftballon)
P30 (und P29 P32)
P31 (reiben den ich)
P32 (mit P32 Wollpullover)
P33 (mod Luftballon ein)
P34 (mod Wollpullover ein)
P35 (wenn P36 P30)
P36 (herstellen ich Elektrizität)

P37 (sein ich Armin-das-Elektrizitätswerk)
P38 (mod P37 also)

P39 (sehen-können man)
P40 (mod P39 nicht)

P41 (sehen-können man P43)
P42 (mod P41 aber)
P43 (dasein Elektrizität)

P44 (speichern man sie)
P45 (auf P44 Ballonhülle)

P46 (anziehen sie Haare)
P47 (mod Haare lang)
P48 (mod Haare Franziskas)

P49 (wegkönnen Elektrizität)
P50 (mod P49 nicht)
P51 (von P49 Ballonhülle)

P52 (sein-die Elektrizität Hauptzutat)
P53 (für P52 Blitz)
P54 (sein-der Elektrizität Blitz)
P55 (mod P54 aber)
P56 (mod P54 noch-nicht)
P57 (mod Blitz selbst)
P58 (mod Blitz eigentlich)

P59 (nehmen ich Ballon)
P60 (mod P59 jetzt)
P61 (legen ich den)
P62 (in P61 Wasser)
P63 (wenn P64 P59)
P64 (verlassen-können Elektrizität Ballon)

P65 (abfließen sie)
P66 (in P65 Wasser)

P67 (tun ich Ballon)
P68 (mod P67 jetzt)
P69 (mod P67 wieder)
P70 (neben P67 Haare)
P71 (mod Haare Franziskas)
P72 (wenn P73 P70)
P73 (hochkommen Haare)
P74 (mod P73 nicht)

P75 (anziehen man sie)
P76 (mod P75 nicht)
P77 (mod P76 mehr)

P78 (bewegen sich)
P79 (was P78)
P80 (mod P78 da)
P81 (machen Wind P79)
P82 (mod P81 nur)

P83 (haben-wollen ich Elektrizität)
P84 (mod P83 jetzt)
P85 (mod Elektrizität neu)
P86 (auf P83 Luftballon)
P87 (mod P86 drauf)

P88 (wenn P94 P86)
P89 (machen-müssen ich den)
P90 (mod machen-müssen trocken)
P91 (mod P89 wieder)
P92 (mod P89 erst)
P93 (draufreiben ich Elektrizität)
P94 (und P89 P96 P98)
P95 (mod draufreiben da)
P96 (mit P89 Wollupullover)
P97 (mod Elektrizität neu)
P98 (erzeugen ich Elektrizität)

P99 (ausprobieren-können jeder-von-euch das)
P100 (mod P99 allerdings)
P101 (mod ausprobieren selber)
P102 (mod P99 zuhause)

P103 (wehtun $)
P104 (mod P103 nicht)
P105 (und P103 P106)
P106 (passieren es nichts)
P107 (mod P106 bei)

P108 (draufkriegen ich Energie)
P109 (auf P108 Luftballon)
P110 (mod P108 denn)
P111 (mod Lufballon so-ein)
P112 (wie-viel P109)

P113 (abhängen-von das Größe)
P114 (mod P113 bisschen)
P115 (mod bisschen ein)
P116 (und P113 P118)
P117 (reiben ich)
P118 (wie-viel P117)

P119 (sein-die Elektrizität)
P120 (mod P119 hier)
P121 (mod P119 wieder)
P122 (mod Elektrizität gespeichert)
P123 (auf P122 Ballon)

P124 (zeigen-können ich euch)
P125 (mod P124 jetzt)
P126 (passieren)
P127 (was P125)
P128 (wenn P127 P129)
P129 (sein-dieser Speicher)
P130 (mod sein voll)
P131 (mod voll zu)

P132 (sein Pullover)
P133 (mod Pullover mein)
P134 (was P132)
P135 (mod P132 eben)
P136 (sein das Bänder)
P137 (mod Bänder diese)
P138 (mod P136 jetzt)
P139 (mod Bänder beide)
P140 (einschalten ich die)

P141 (aneinander-reiben die)
P142 (herstellen die Elektrizität)
P143 (und P141 P142)

P144 (sein Kugel das)
P145 (sein Luftballon)
P146 (mod P145 eben)
P147 (mod Luftballon mein)
P148 (was P145)

P149 (sein Speicher)
P150 (auf P149 Kugel)
P151 (mod Speicher dieser)
P152 (mod sein voll)
P153 (mod voll zu)
P154 (und P149 P158)
P155 (kommen ich)
P156 (mit P155 Metallspitze)
P157 (mod Metallspitze eine)
P158 (in P156 Nähe)
P159 (wenn P162 P154)
P160 (wegspringen Elektrizität)
P161 (mod P160 irgendwann)
P162 (von P160 Kugel)
P163 (mod P149 jetzt)

P164 (sein-ein das Blitz)
P165 (mod Blitz klein)

P166 (erzeugen Band Elektrizität)
P167 (mod P166 also)
P168 (mod Elektrizität so-viel)
P169 (durch P166 Reibung)
P170 (dass P169 P171)
P171 (entstehen Blitz)
P172 (mod Blitz ein)
P173 (mod Blitz klein)

P174 (reiben-am $ Himmel)
P175 (was P174)
P176 (mod P174 aber)
P177 (mod P174 denn)

P178 (reiben niemand)
P179 (mod P178 dran)
P180 (mod dran da)
P181 (mod P178 doch)

P182 (reiben Mensch)
P183 (mod Mensch kein)
P184 (mod P182 dran)

P185 (sehen man P189)
P186 (im P185 Zeitraffer)
P187 (rasen Wolken)
P188 (über P187 Himmel)
P189 (wie P188)

P190 (sein-die das Luftschichten)
P191 (vorbeibewegen die sich)
P192 (mod vorbeibewegen aneinander)

P193 (reiben Eiskristalle sich)
P194 (mod Eiskristalle klein)
P195 (mod klein winzig)
P196 (an P193 Luft)
P197 (und P196 P198)
P198 (herstellen Eiskristalle Elektrizität)
P199 (mod P198 dadurch)

P200 (entstehen Elektrizität)
P201 (für P200 Blitz)
P202 (mod P200 also)
P203 (durch P200 Reibung)
P204 (mod Reibung Luftschichten)

P205 (blitzen es)
P206 (wann P205)

P207 (nehmen wir Luftballon)
P208 (mod P207 wieder)
P209 (mod Luftballon unser)
P210 (mod P207 da)

P211 (draufreiben wir Elektrizität)
P212 (mod P211 nicht)
P213 (mod P211 aber)
P214 (mod P211 jetzt)
P215 (mod P211 da)
P216 (mod draufreiben außen)
P217 (füllen wir den)
P218 (mit P217 Wasser)
P219 (sondern P211 P218)

P220 (gutgehen das Weile)
P221 (mod Weile eine)
P222 (vollsein Speicher)
P223 (mod P222 irgendwann)
P224 (mod irgendwann aber)
P225 (und P222 P226)
P226 (platzen Ballon)
P227 (mod P226 dann)

P228 (sagen-können man P229)
P229 (entladen Ballon sich)
P230 (mod P228 auch)
P231 (mod entladen blitzartig)

P232 (blitzen es)
P233 (speichern-können Himmel Elektrizität)
P234 (mod Elektrizität keine)
P235 (mod keine mehr)
P236 (wenn P232 P233)

P237 (passieren $)
P238 (mod passieren genau)
P239 (was P237)
P240 (blitzen es)
P241 (wenn P239 P240)

P242 (zeigen-können wir das)
P243 (mod P242 euch)
P244 (reisen-müssen wir)
P245 (nach P244 München)
P246 (zur P244 Universität)
P247 (mod Universität technisch)
P248 (um P246 P242)

P249 (erzeugen-können die Blitze)
P250 (mod erzeugen-können künstlich)

P251 (angucken-können man Blitze)
P252 (mod P251 sich)
P253 (mod Blitze künstlich)
P254 (mod Blitze solche)
P255 (mod angucken-können genau)
P256 (mod P251 nur)

P257 (würde-warten-wollen man darauf)
P258 (in P257 Natur)
P259 (haben-müssen man Geduld)
P260 (wenn P259 P258)

P261 ($ Jahr)
P262 (mod Jahr ein)
P263 (oder P261 noch-länger)
P264 (mod P261 vielleicht)

P265 (brauchen man Monstrum)
P266 (an P265 Kamera)
P267 (filmen $ Blitz)
P268 (um P266 P267)
P269 (mod Monstrum so-ein)
P270 (mod P265 auch)
P271 (mod Blitz so-ein)

P272 (sein das so-schwer)
P273 (dass P272 P274)
P274 (tragen-müssen Leute es)
P275 (mod Leute zwei)

P276 (aussehen Kamera so)
P277 (von P276 innen)
P278 (mod Kamera diese)

P279 (zeigen-können die Blitz)
P280 (mod zeigen-können langsamer)
P281 (mod langsamer 40000mal)
P282 (als P279 P284)
P283 (passieren er)
P284 (in P283 Wirklichkeit)

P285 (aushalten-würde Filmmaterial Geschwindigkeit)
P286 (mod Filmmaterial kein)
P287 (mod Filmmaterial der-Welt)
P288 (mod Geschwindigkeit diese)

P289 (bedienen-sich man Trick)
P290 (mod Trick ein)
P291 (mod P289 deswegen)

P292 (sein Spiegel da)
P293 (drehen Spiegel sich)

P294 (umlenken der Bild)
P295 (von P294 Blitz)
P296 (auf P294 Filmmaterial)
P297 (einspannen $ Filmmaterial)
P298 (am P297 Rand)
P299 (mod einspannen außen)
P300 (und P298 P301)
P301 (bewegen Filmmaterial sich)
P302 (mod P301 nicht)
P303 (mod nicht überhaupt)

P304 (hinkommen es da)

P305 (treffen man Punkt)
P306 (mod Punkt richtig)
P307 (mod treffen auch)
P308 (mit P305 Kamera)
P309 (damit P312 P308)
P310 (blitzen es)
P311 (wo P310)
P312 (sein-eine Blitzanlage da)

P313 (kommen Blitz)
P314 (von P313 oben)
P315 (mod oben da)

P316 (einschlagen Blitz)
P317 (mod P316 nachher)
P318 (in P316 Haus)
P319 (mod Haus dieses)
P320 (mod P316 dann)
P321 (mod einschlagen unten)

P322 (treffen-würde man Blitz)
P323 (in P322 Natur)
P324 (mod P322 sonst)
P325 (mod treffen-würde nie)
P326 (mod Blitz so-ein)
P327 (mod P322 sonst)
P328 (mod P322 ja)

P329 (wissen man P331)
P330 (blitzen es)
P331 (wo P330)
P332 (mod P329 ja)
P333 (mod P329 gar-nicht)

P334 (reintun man Optik)
P335 (mod P334 vorher-noch)

P336 (kommen Signal)
P337 (mod P336 dann)
P338 (mod Signal ein)

P339 (sein das Elektrizität)
P340 (rausmüssen alle)

P341 (sein das gefährlich)

P342 (werden es spannend)
P343 (mod P342 jetzt)

P344 (speichern-müssen $ Elektrizität)
P345 (mod Elektrizität so-viel)
P346 (mod mal zuerst)
P347 (mod P344 jetzt)
P348 (mod speichern mal)
P349 (dass P344 P350)
P350 (blitzen-können es)

P351 (machen-können man das)
P352 (in P351 Universität-München)

P353 (einschalten $ Kamera)
P354 (mod P353 dann)
P355 (hochfahren $ Kamera)
P356 (und P353 P357)
P357 (mit P355 Geschwindigkeit)
P358 (mod Geschwindigkeit ihre)
P359 (wenn P362 P356)
P360 (sehen-können man Dinge)
P361 (mod Dinge so-tolle)
P362 (in P360 Blitz)
P363 (sehen man Dinge)
P364 (in P363 Leben)
P365 (mod sehen noch-nie)

P366 (haben Blitz Geheimnis)
P367 (mod Blitz so-ein)
P368 (mod P366 eigentlich)
P369 (mod Geheimnis drei)
P370 (gehen Blitz so-schnell)

P371 ($ Geheimnis)
P372 (mod Geheimnis erstes)

P373 (sehen man Blitz)
P374 (mod P373 bevor)
P375 (sein-eine Vorentladung)
P376 (mod P375 vorher)

P377 (bilden die Schlauch)
P378 (mod Schlauch so-ein)

P379 (nachrutschen Blitz)
P380 (mod P379 dann)
P381 (mod P379 erst)
P382 (mod Blitz eigentlich)

P383 ($ Geheimnis)
P384 (bei P383 Mal)
P385 (mod Mal nächstes)
P386 (mod P384 jetzt)
P387 (mod Geheimnis zweites)

P388 (denken man P390)
P389 (mod P388 immer)
P390 (sausen Blitz)
P391 (von P390 oben)
P392 (nach P391 unten)

P393 (sein das so)
P394 (mod P393 gar-nicht)

P395 (kommen Teil)
P396 (von P395 oben)
P397 (mod Teil Blitz)

P398 (gehen er Viertel)
P399 (mod Viertel drei)
P400 (mod Viertel Weg)
P401 (wenn P403 P398)
P402 (entgegenkommen einer)
P403 (von P402 unten)

P404 (leuchten es)
P405 (mod P404 dann)
P406 (mod P404 erst)

P407 ($ Geheimnis)
P408 (mod Geheimnis drittes)
P409 (mod P407 jetzt)
P410 (weil P407 P411)
P411 (sein es so-schön)

P412 ($ Achtung)
P413 (mod Achtung gleich)

P414 (nachrutschen Blitz)
P415 (mod nachrutschen ruckweise)
P416 (in P414 Zug)
P417 (mod Zug schnell)
P418 (mod Zug ein)
P419 (mod P416 nicht)
P420 (wie P416 P421)
P421 (sehen man das)
P422 (mod P421 immer)

P423 (leuchten es)
P424 (mod P423 dann)

P425 (sein-die das Hauptentladung)

P426 (passieren dasselbe)
P427 (mod dasselbe genau)
P428 (mod passieren blitzschnell)
P429 (bei P426 Blitz)
P430 (mod Blitz jeder)
P431 (mod P426 immer)
P432 (wenn P433 P431)
P433 (sein-ein Gewitter)

P434 (denken ihr das)
P435 (Frage P434)

Note:

$ = missing argument in the text

mod = modified (i.e., specifies a preceding argument or sentence

C

Experimental Filmtranscripts

C.1 Thermos – Version 1

Ein schönes Glas Tee. Kann Christoph in aller Ruhe trinken. Aber es ist heiß. Kann er kaum anfassen. Wenn das Telefon nicht wäre. Ist wieder einer dran, der erzählt und erzählt und erzählt. Und wenn das Gespräch dann beendet ist, dann ist der Tee kalt geworden. Wie kann man das verhindern? Das muss doch raus zu kriegen sein, denkt Christoph. Mit einem Topf voll mit heißem Wasser? Da ist die Wärme jetzt noch drin, klar. Dann hängt er da vier Gegenstände hinein: einen Metalllöffel, ein Glasröhrchen, ein Stück Gummi und ein Stück Holz. Dazu holt er noch ne Stoppuhr. Nach 30 Sekunden der Fingertest. *DER LÖFFEL AUS METALL HEIß. GLAS, MHM, GEHT SO. DAS GUMMI IST NOCH KÄLTER. UND DAS HOLZ, IST GANZ KALT GEBLIEBEN. DAS IST ES. (IMPLICIT INFORMATION)* (naming word IR context: Leiten)

Christoph ist ein Mann der Tat. Gleich greift er zur Bohrmaschine und bohrt ein Riesenloch in einen Holzklotz. Sieht irgendwie praktisch aus. Und was soll es sein? Eine Teetasse – aus Holz.. Aber der Tee bleibt trotzdem nicht lange heiß, *WEIL DIE WÄRME AUS DER HOLZTEETASSE ENTWEICHT. (EXPLICIT INFORMATION)** Das muss noch besser gehn, findet Christoph, mal überlegen. Von den vier Gegenständen war der aus Holz schon der beste. Aber es muss noch etwas anderes geben.

NEUER VERSUCH: CHRISTOPH FÜLLT HEIßES WASSER IN ZWEI REAGENZGLÄSER. (IMPLICIT INFORMATION) (naming word IR context: Vergleich) Auf eines der Reagenzgläser legt er ein Glasplättchen und dann darauf noch eins. Und auf das andere Glasgefäß legt er auch ein Glasplättchen, darauf dann Korkstückchen und dann erst das zweite Glasplättchen. *ZWISCHEN DEN BEIDEN GLASPLÄTTCHEN IST ALSO LUFT. (EXPLICIT INFORMATION)* (naming word IU context: Spannung) Und jetzt, der Buttertest. Dazu gibt Christoph jeweils ein Stückchen Butter auf die Glasplättchen. Wo schmilzt die Butter zuerst? *DIE BUTTER IST AUF DEN GLASPLÄTTCHEN OHNE KORKSTÜCKCHEN DAZWISCHEN GESCHMOLZEN. AUF DEN GLASPLÄTTCHEN MIT DEN KORKSTÜCKCHEN ALS ABSTANDHALTER DAGEGEN NICHT. (IMPLICIT INFORMATION)* (naming word IR context: Isolieren)

Ein neuer Versuch: Ein großes Glas, dahinein ein paar Korkstückchen als Abstandshalter zum Boden und darauf ein kleineres Teeglas. Mal gucken, ob das funktioniert. Ein einfaches Vergleichsteeglas, um zu testen, ob der Tee im Glas-im-Glas heißer bleibt. Die Anfangstemperatur: 80, 81, gute 82 Grad im Vergleichsteeglas. Und da, 80, 81, auch im Glas-im-Glas 82 Grad. Also, die Zeit läuft. Nach einer Viertelstunde, stopp und nachmessen. Im Vergleichsteeglas 46 Grad. Bei dem Glas-im-Glas 60 Grad. Ganz schöner Unterschied. Was passiert mit der Wärme? Nach unten und zur Seite wird die Wärme nicht mehr weggeleitet*. ABER NOCH NACH OBEN. (EXPLICIT INFORMATION)* (naming word IU context: Anziehung) Hhhhm.

Zweiter Test mit dem Vergleichsteeglas und dem Glas-im-Glas. Diesmal bekommt aber jedes der beiden Gläser im Glas-im-Glas noch ein Stück aus dünner Pappe darauf. Wieder nach einer Viertelstunde messen. Im Vergleichsteeglas, das überrascht jetzt nicht, wieder so 45, 46 Grad. *UND JETZT MAL GUCKEN, WAS DIE PAPPE GEBRACHT HAT. 70, 71, 72 GRAD IM GLAS-IMGLAS! ALSO, 12 GRAD MEHR ALS OHNE PAPPE. (IMPLICIT INFORMATION)* (naming word IR context: Deckel) Der Tee ist noch heiß. Christoph ist stolz auf seine neue Tasse. Vielleicht gibt's aber noch was zu verbessern. Er hat ne Idee.

Christoph holt Alufolie. Gibt's in fast jedem Haushalt. Natürlich kann man damit auch ein Teeglas umwickeln. Aber, welchen Sinn soll das haben? Um das zu verstehen, nimmt Christoph eine Wärmelampe und bestrahlt damit ein Teeglas, das mit Alufolie umwickelt ist und ein Glas ohne Alufolie. Und jetzt, wieder der Buttertest. Dazu gibt Christoph jeweils ein Stückchen Butter in die beiden Gläser. Klares Ergebnis. Im Glas ohne Alufolie schmilzt die Butter. Im Glas mit Alufolie bleibt die Butter ganz, weil die Wärme weder ins Glas rein, noch raus kommt, *DA DIE ALUFOLIE DIE WÄRME ZURÜCKSPIEGELT (EXPLICIT INFORMATION)* **(naming word IU context: Einschlag)** Dann müsste man doch jetzt das Glas mit der Alufolie drumrum und dem heißen Tee drinnen gut anfassen können. So ist es auch. Das Glas ohne Alufolie kann man dagegen kaum anfassen, weil es die Wärme aus dem Tee angenommen hat. Damit wird der Tee im Glas ohne Alufolie auch schneller kalt.

Christoph nimmt ein kleines Teeglas mit Alufolie umwickelt und stellt es in das große Glas, mit den Korkstücken als Abstandshalter hinein. Dann füllt er Tee hinein. Und dann noch den praktischen Doppeldeckel mit dem Strohhalm drauf. Super! Das alles hält den Tee jetzt schön heiß. Christoph ist stolz auf seine neue Teewarmhaltetasse.

Aber Christoph, weißt du? Das gibt's schon. Schon mal gehört? Thermoskanne? Ja, die hält den Tee auch schön warm. Funktioniert ganz ähnlich wie deine Erfindung. Wir ham mal eine aufgeschnitten. Auch eine Flasche in der Flasche.

Aber, Christoph trinkt natürlich am liebsten aus seiner eigenen Teewarmhaltetasse. Darin bleibt der Tee auch wunderbar heiß.

Note: Sentences printed in italics and capital letters refer to the type of information *(EXPLICIT VERSUS IMPLICIT INFORMATION)*. Sentences printed in bold refer to type of naming word and to the naming word presented **(IR versus IU context: naming word presented)**.

*As the two experimental films revealed an uneven number of identified LCBI positions, at one LCBI position in the thermos film a naming task was only implemented in its IR context without presenting a naming word from the lightning film in its IU context. Accordingly, in the lightning film, one naming word from the thermos film was presented in its IU context at a control position that did not correspond to a LCBI position.

C.1 Thermos – Version 2

Ein schönes Glas Tee. Kann Christoph in aller Ruhe trinken. Aber es ist heiß. Kann er kaum anfassen. Wenn das Telefon nicht wäre. Ist wieder einer dran, der erzählt und erzählt und erzählt. Und wenn das Gespräch dann beendet ist, dann ist der Tee kalt geworden. Wie kann man das verhindern? Das muss doch raus zu kriegen sein, denkt Christoph. Mit einem Topf voll mit heißem Wasser? Da ist die Wärme jetzt noch drin, klar. Dann hängt er da vier Gegenstände hinein: einen Metallllöffel, ein Glasröhrchen, ein Stück Gummi und ein Stück Holz. Dazu holt er noch ne Stoppuhr. Nach 30 Sekunden der Fingertest. Der Löffel aus Metall … heiß. Glas, mhm, geht so. Das Gummi ist noch kälter. Und das Holz, ist ganz kalt geblieben. Das ist es. *DAS LEITET DIE WÄRME AM WENIGSTEN (EXPLICIT INFORMATION)* (naming word IU context: Elektrizität)

Christoph ist ein Mann der Tat. Gleich greift er zur Bohrmaschine und bohrt ein Riesenloch in einen Holzklotz. Sieht irgendwie praktisch aus. Und was soll es sein? Eine Teetasse – aus Holz.. *ABER DER TEE BLEIBT TROTZDEM NICHT LANGE HEIß. (IMPLICIT INFORMATION)* (naming word IR context: Entweichen) Das muss noch besser gehn, findet Christoph, mal überlegen. Von den vier Gegenständen war der aus Holz schon der beste. Aber es muss noch etwas anderes geben.

Neuer Versuch: Christoph füllt heißes Wasser in zwei Reagenzgläser. *ER WILL WOHL WAS VERGLEICHEN (EXPLICIT INFORMATION)* (naming word IU context: Reibung) Auf eines der Reagenzgläser legt er ein Glasplättchen und dann darauf noch eins. *UND AUF DAS ANDERE GLASGEFÄß LEGT ER AUCH EIN GLASPLÄTTCHEN, DARAUF DANN KORKSTÜCKCHEN UND DANN ERST DAS ZWEITE GLASPLÄTTCHEN. (IMPLICIT INFORMATION)* (naming word IR context: Luft) Und jetzt, der Buttertest. Dazu gibt Christoph jeweils ein Stückchen Butter auf die Glasplättchen. Wo schmilzt die Butter zuerst? Die Butter ist auf den Glasplättchen ohne Korkstückchen dazwischen geschmolzen. Auf den Glasplättchen mit den Korkstückchen als Abstandhalter dagegen nicht. *HIER HAT DIE LUFT ZWISCHEN DEN GLASPLÄTTCHEN ALS ISOLATOR GEWIRKT (EXPLICIT INFORMATION)* (naming word IU context: Entladung)

Ein neuer Versuch: Ein großes Glas, dahinein ein paar Korkstückchen als Abstandshalter zum Boden und darauf ein kleineres Teeglas. Mal gucken, ob das funktioniert. Ein einfaches Vergleichsteeglas, um zu testen, ob der Tee im Glas-im-Glas heißer bleibt. Die Anfangstemperatur: 80, 81, gute 82 Grad im Vergleichsteeglas. Und da, 80, 81, auch im Glas-im-Glas 82 Grad. Also, die Zeit läuft. Nach einer Viertelstunde, stopp und nachmessen. Im Vergleichsteeglas 46 Grad. Bei dem Glas-im-Glas 60 Grad. Ganz schöner Unterschied. *WAS PASSIERT MIT DER WÄRME? NACH UNTEN UND ZUR SEITE WIRD DIE WÄRME NICHT MEHR WEGGELEITET. (IMPLICIT INFORMATION)* (naming word IR context: Oben) Hhhhm.

Zweiter Test mit dem Vergleichsteeglas und dem Glas-im-Glas. Diesmal bekommt aber jedes der beiden Gläser im Glas-im-Glas noch ein Stück aus dünner Pappe darauf. Wieder nach einer Viertelstunde messen. Im Vergleichsteeglas, das überrascht jetzt nicht, wieder so 45, 46 Grad. Und jetzt mal gucken, was die Pappe gebracht hat. 70, 71, 72 Grad im Glas-imGlas! Also, 12 Grad mehr als ohne Pappe. *DER DECKEL AUS PAPPE HAT VERHINDERT, DASS DIE NACH OBEN AUFGESTIEGENE WÄRME ENTWEICHT (EXPLICIT INFORMATION).* Der Tee ist noch heiß. Christoph ist stolz auf seine neue Tasse. Vielleicht gibt's aber noch was zu verbessern. Er hat ne Idee.

Christoph holt Alufolie. Gibt's in fast jedem Haushalt. Natürlich kann man damit auch ein Teeglas umwickeln. Aber, welchen Sinn soll das haben? Um das zu verstehen, nimmt Christoph eine Wärmelampe und bestrahlt damit ein Teeglas, das mit Alufolie umwickelt ist und ein Glas ohne Alufolie. Und jetzt, wieder der Buttertest. Dazu gibt Christoph jeweils ein Stückchen Butter in die beiden Gläser. Klares Ergebnis. *IM GLAS OHNE ALUFOLIE SCHMILZT DIE BUTTER. IM GLAS MIT ALUFOLIE BLEIBT DIE BUTTER GANZ, WEIL DIE WÄRME WEDER INS GLAS REIN, NOCH RAUS KOMMT. SPIEGELN (IMPLICIT INFORMATION)** * Dann müsste man doch jetzt das Glas mit der Alufolie drumrum und dem heißen Tee drinnen gut anfassen können. So ist es auch. Das Glas ohne Alufolie kann man dagegen kaum anfassen, weil es die Wärme aus dem Tee angenommen hat. Damit wird der Tee im Glas ohne Alufolie auch schneller kalt.

Christoph nimmt ein kleines Teeglas mit Alufolie umwickelt und stellt es in das große Glas, mit den Korkstücken als Abstandshalter hinein. Dann füllt er Tee hinein. Und dann noch den praktischen Doppeldeckel mit dem Strohhalm drauf. Super! Das alles hält den Tee jetzt schön heiß. Christoph ist stolz auf seine neue Teewarmhaltetasse.

Aber Christoph, weißt du? Das gibt's schon. Schon mal gehört? Thermoskanne? Ja, die hält den Tee auch schön warm. Funktioniert ganz ähnlich wie deine Erfindung. Wir ham mal eine aufgeschnitten. Auch eine Flasche in der Flasche.

Aber, Christoph trinkt natürlich am liebsten aus seiner eigenen Teewarmhaltetasse. Darin bleibt der Tee auch wunderbar heiß.

Note: Sentences printed in italics and capital letters refer to the type of information *(EXPLICIT VERSUS IMPLICIT INFORMATION)*. Sentences printed in bold refer to type of naming word and to the naming word presented (**IR versus IU context: naming word presented**).

*As the two experimental films revealed an uneven number of identified LCBI positions, at one LCBI position in the thermos film a naming task was only implemented in its IR context without presenting a naming word from the lightning film in its IU context. Accordingly, in the lightning film, one naming word from the thermos film was presented in its IU context at a control position that did not correspond to a LCBI position.

C.2 Lightning – Version 1

Ihr habt alle schon mal erlebt, dass es blitzt. Was hat aber dieser Blitz mit einem Luftballon und einem Wollpullover zu tun? Es ist beides ungefähr das Gleiche.

WENN MAN EINEN WOLLPULLOVER NIMMT UND BEWEGT DEN AUF EINEM LUFTBALLON AUF UND AB, DANN KOMMT ANSCHLIEßEND DIESES PAPIERKONFETTI NACH OBEN, WENN MAN DEN BALLON IN DIE NÄHE HÄLT (IMPLICIT INFORMATION) (naming word IR context: Reibung) Das Gleiche passiert mit Franziskas langen Haaren: auch sie werden, wie das Papierkonfetti *VOM BALLON ANGEZOGEN (EXPLICIT INFROMATION),* wenn man ihn in ihre Nähe hält. (naming word IU context: Entweichen) Wenn man den Ballon jetzt nimmt und legt den ins Wasser, dann kommen die Haare anschließend nicht mehr hoch, sie werden nicht mehr angezogen. Was sich da bewegt, macht nur der Wind.

DURCH DIE REIBUNG VON PULLOVER UND BALLON, WURDE DIE HAUPTZUTAT FÜR DEN BLITZ ERZEUGT. DIE KANN MAN ZWAR NICHT SEHEN, ABER SIE IST AUF DER BALLONHÜLLE GESPEICHERT UND ZIEHT FRANZISKAS LANGE HAARE AN. (IMPLICIT INFORMATION) (naming word IR context: Elektrizität) Wenn man den Ballon aber ins Wasser legt, dann kann die Elektrizität ins Wasser abfließen. Um neue Elektrizität erzeugen zu können, muss der Ballon erst wieder trocken sein. Wie viel Elektrizität kriegt man denn auf so nen Luftballon drauf? Hängt ein bisschen von der Größe ab, und wie viel man reibt. Jetzt können wir euch zeigen was passiert, wenn dieser Speicher zu voll ist.

Unser Pullover, das sind jetzt diese beiden Bänder, die eingeschaltet werden. Die reiben aneinander und stellen Elektrizität her. Oben sind sie mit einer Metallkugel verbunden. Diese Kugel ist jetzt das, was eben unser Luftballon war. Und wenn der Speicher auf der Kugel jetzt zu voll ist, und man kommt mit ner Metallspitze in ihre Nähe, dann springt irgendwann die Elektrizität von der Kugel weg. Und das ist ein kleiner Blitz, *DER DURCH DIE ZUVOR AUFGEBAUTE SPANNUNG ENTSTANDEN IST. (EXPLICIT INFORMATION)* (naming word IU context: Luft) Das Band hat also durch Reibung so viel Elektrizität erzeugt, dass ein kleiner Blitz entsteht.

Aber was reibt denn am Himmel? Im Zeitraffer sieht man, wie die Wolken über den Himmel rasen. Das sind die Luftschichten, die sich aneinander vorbeibewegen. Und winzig kleine Eiskristalle reiben sich an der Luft und stellen dadurch die Elektrizität her. Die Elektrizität für den Blitz entsteht also durch die Reibung der Luftschichten. Und wann blitzt´s? Da nehmen wir wieder unsern Luftballon. Reiben aber jetzt da nicht außen Elektrizität drauf, sondern wir füllen den mit Wasser. *DAS GEHT NE WEILE GUT, ABER IRGENDWANN IST DER SPEICHER VOLL, UND DANN PLATZT DER BALLON. UND WENN DER HIMMEL KEINE ELEKTRIZITÄT MEHR SPEICHERN KANN, DANN BLITZT´S. (IMPLICIT INFORMATION)* (naming word IR context: Entladung)

Und was passiert genau wenn´s blitzt? Um euch das zeigen zu können mussten wir nach München reisen, zur technischen Universität. Die können Blitze künstlich erzeugen. Und nur solche künstlichen Blitze kann man sich genau angucken. Wollte man darauf in der Natur warten, müsste man Geduld haben. Vielleicht ein Jahr oder noch länger. Und um so einen Blitz zu filmen, braucht man auch so ein Monstrum an Kamera. Das ist so schwer, dass es zwei Leute tragen müssen. Von innen sieht diese Kamera so aus. Die kann einen Blitz 40 000 mal langsamer zeigen, als er in Wirklichkeit passiert. Und diese Geschwindigkeit würde kein Filmmaterial der Welt aushalten. Deswegen bedient man sich eines Tricks. Da ist ein Spiegel, der sich dreht. Und der lenkt das Bild von dem Blitz auf das Filmmaterial um, das außen am Rand eingespannt ist und sich überhaupt nicht bewegt. Da

kommt´s hin. Und damit man mit der Kamera auch den richtigen Punkt trifft, wo´s blitzt, ist da eine Blitzanlage und ein kleines Modelhaus. Aus der Blitzanlage kommt der Blitz nachher raus. *UND DER WIRD DANN IN DAS HAUS EINSCHLAGEN.(EXPLICIT INFORMATION)* **(naming word IU context: Oben)** Sonst in der Natur würd man so nen Blitz ja nie treffen. Man wüßt ja gar nicht wo es blitzt. Vorher noch die Optik rein. Dann kommt ein Signal. Das ist Elektrizität, alle müssen raus. **(naming word IU context: Spiegeln)*** Das ist gefährlich. Und jetzt wird´s spannend. Jetzt muss zuerst mal so viel Elektrizität gespeichert werden, dass es blitzen kann. Das kann man in der Universität München machen. Und wenn dann die Kamera eingeschaltet ist und hochgefahren ist mit ihrer Geschwindigkeit, dann kann man in dem Blitz so tolle Dinge sehen, die man noch nie im Leben gesehen hat. So ein Blitz, der ja so schnell geht, hat eigentlich drei Geheimnisse. Erstes Geheimnis: bevor man den Blitz sieht, ist vorher eine Vorentladung. Die bildet so nen Schlauch. Und dann rutscht der eigentliche Blitz erst nach. Und jetzt beim nächsten mal das zweite Geheimnis: Man denkt ja immer der Blitz saust von oben nach unten. Das ist gar nicht so. Ein Teil des Blitzes kommt von oben. Und wenn er drei Viertel des Weges gegangen ist, kommt von unten einer entgegen. Und dann leuchtet´s erst. Und weil´s so schön war jetzt das dritte Geheimnis: Achtung gleich! Der Blitz rutscht ruckweise nach, nicht in einem schnellen Zug, wie man das immer sieht. Und dann leuchtet´s. Das ist die Hauptentladung. Und genau dasselbe passiert blitzschnell, bei jedem Blitz, immer dann, wenn ein Gewitter ist. Hättet ihr das gedacht?

Note: Sentences printed in italics and capital letters refer to the type of information *(EXPLICIT VERSUS IMPLICIT INFORMATION)*. Sentences printed in bold refer to type of naming word and to the naming word presented **(IR versus IU context: naming word presented)**.

*As the two experimental films revealed an uneven number of identified LCBI positions, at one LCBI position in the thermos film a naming task was only implemented in its IR context without presenting a naming word from the lightning film in its IU context. Accordingly, in the lightning film, one naming word from the thermos film was presented in its IU context at a control position that did not correspond to a LCBI position.

C.2 Lightning – Version 2

Ihr habt alle schon mal erlebt, dass es blitzt. Was hat aber dieser Blitz mit einem Luftballon und einem Wollpullover zu tun? Es ist beides ungefähr das Gleiche.

Wenn man einen Wollpullover nimmt und bewegt den auf einem Luftballon auf und ab, dann kommt *DURCH DIESE REIBUNG (EXPLICIT INFORMATION)* anschließend dieses Papierkonfetti nach oben, wenn man den Ballon in die Nähe hält. (naming word IU context: Deckel) *DAS GLEICHE PASSIERT MIT FRANZISKAS LANGEN HAAREN: AUCH SIE KOMMEN WIE DAS PAPIERKONFETTI NACH OBEN, WENN MAN DEN BALLON IN IHRE NÄHE HÄLT. (IMPLICIT CONTEXT)* (naming word IR context: Anziehung) Wenn man den Ballon jetzt nimmt und legt den ins Wasser, dann kommen die Haare anschließend nicht mehr hoch, sie werden nicht mehr angezogen. Was sich da bewegt, macht nur der Wind. Durch die Reibung von Pullover und Ballon, wurde die Hauptzutat für den Blitz erzeugt. Die kann man zwar nicht sehen, aber *DIE ELEKTRIZITÄT (EXPLICIT INFORMATION)* ist auf der Ballonhülle gespeichert und zieht Franziskas lange Haare an. (naming word IU context: Vergleich) Die Elektrizität kann nicht von der Ballonhülle weg. Wenn man den Ballon aber ins Wasser legt, dann kann die Elektrizität ins Wasser abfließen. Um neue Elektrizität erzeugen zu können, muss der Ballon erst wieder trocken sein. Wie viel Elektrizität krieg ich denn auf so nen Luftballon drauf? Hängt ein bisschen von der Größe ab, und wie viel man reibt. Jetzt können wir euch zeigen was passiert, wenn dieser Speicher zu voll ist.

Unser Pullover, das sind jetzt diese beiden Bänder, die eingeschaltet werden. Die reiben aneinander und stellen Elektrizität her. Oben sind sie mit einer Metallkugel verbunden. Diese Kugel ist jetzt das, was eben unser Luftballon war. *UND WENN DER SPEICHER AUF DER KUGEL JETZT ZU VOLL IST, UND MAN KOMMT MIT NER METALLSPITZE IN IHRE NÄHE, DANN SPRINGT IRGENDWANN DIE ELEKTRIZITÄT VON DER KUGEL WEG. UND DAS IST EIN KLEINER BLITZ. (IMPLICIT INFORMATION)* (naming word IR context: Spannung) Das Band hat also durch Reibung so viel Elektrizität erzeugt, dass ein kleiner Blitz entsteht.

Aber was reibt denn am Himmel? Im Zeitraffer sieht man, wie die Wolken über den Himmel rasen. Das sind die Luftschichten, die sich aneinander vorbeibewegen. Und winzig kleine Eiskristalle reiben sich an der Luft und stellen dadurch die Elektrizität her. Die Elektrizität für den Blitz entsteht also durch die Reibung der Luftschichten. Und wann blitzt´s? Da nehmen wir wieder unsern Luftballon. Reiben aber jetzt da nicht außen Elektrizität drauf, sondern wir füllen den mit Wasser. Das geht ne Weile gut, aber irgendwann ist der Speicher voll, und dann platzt der Ballon. Und wenn der Himmel keine Elektrizität mehr speichern kann, *DANN KOMMT ES ZU EINER ENTLADUNG DES SPEICHERS (EXPLICIT INFORMATION)* und es blitzt. (naming word IU context: Leiten)

Und was passiert genau wenn´s blitzt? Um euch das zeigen zu können mussten wir nach München reisen, zur technischen Universität. Die können Blitze künstlich erzeugen. Und nur solche künstlichen Blitze kann man sich genau angucken. Wollte man darauf in der Natur warten, müsste man Geduld haben. Vielleicht ein Jahr oder noch länger. Und um so einen Blitz zu filmen, braucht man auch so ein Monstrum an Kamera. Das ist so schwer, dass es zwei Leute tragen müssen. Von innen sieht diese Kamera so aus. Die kann einen Blitz 40 000 mal langsamer zeigen, als er in Wirklichkeit passiert. Und diese Geschwindigkeit würde kein Filmmaterial der Welt aushalten. Deswegen bedient man sich eines Tricks. Da ist ein Spiegel, der sich dreht. Und der lenkt das Bild von dem Blitz auf das Filmmaterial um, das außen am Rand eingespannt ist und sich überhaupt nicht bewegt. Da

kommt´s hin. *UND DAMIT MAN MIT DER KAMERA AUCH DEN RICHTIGEN PUNKT TRIFFT, WO´S BLITZT, IST DA EINE BLITZANLAGE UND EIN KLEINES MODELHAUS. AUS DER BLITZANLAGE KOMMT DER BLITZ NACHHER RAUS. (IMPLICIT INFORMATION)* **(naming word IR context: Einschlagen)** Sonst in der Natur würd man so nen Blitz ja nie treffen. Man wüßt ja gar nicht wo es blitzt. Vorher noch die Optik rein. Dann kommt ein Signal. Das ist Elektrizität, alle müssen raus. **(naming word IU context: Isolieren)*** Das ist gefährlich. Und jetzt wird´s spannend. Jetzt muss zuerst mal so viel Elektrizität gespeichert werden, dass es blitzen kann. Das kann man in der Universität München machen. Und wenn dann die Kamera eingeschaltet ist und hochgefahren ist mit ihrer Geschwindigkeit, dann kann man in dem Blitz so tolle Dinge sehen, die man noch nie im Leben gesehen hat. So ein Blitz, der ja so schnell geht, hat eigentlich drei Geheimnisse. Erstes Geheimnis: bevor man den Blitz sieht, ist vorher eine Vorentladung. Die bildet so nen Schlauch. Und dann rutscht der eigentliche Blitz erst nach. Und jetzt beim nächsten mal das zweite Geheimnis: Man denkt ja immer der Blitz saust von oben nach unten. Das ist gar nicht so. Ein Teil des Blitzes kommt von oben. Und wenn er drei Viertel des Weges gegangen ist, kommt von unten einer entgegen. Und dann leuchtet´s erst. Und weil´s so schön war jetzt das dritte Geheimnis: Achtung gleich! Der Blitz rutscht ruckweise nach, nicht in einem schnellen Zug, wie man das immer sieht. Und dann leuchtet´s. Das ist die Hauptentladung. Und genau dasselbe passiert blitzschnell, bei jedem Blitz, immer dann, wenn ein Gewitter ist. Hättet ihr das gedacht?

Note: Sentences printed in italics and capital letters refer to the type of information *(EXPLICIT VERSUS IMPLICIT INFORMATION)*. Sentences printed in bold refer to type of naming word and to the naming word presented **(IR versus IU context: naming word presented)**.

*As the two experimental films revealed an uneven number of identified LCBI positions, at one LCBI position in the thermos film a naming task was only implemented in its IR context without presenting a naming word from the lightning film in its IU context. Accordingly, in the lightning film, one naming word from the thermos film was presented in its IU context at a control position that did not correspond to a LCBI position.

C.3 Naming words

Word set (Version)	Naming words	Film
1	Reibung Elektrizität Entladung	Lightning
1	Entweichen Luft Oben Spiegeln	Thermos
2	Anziehung Spannung Einschlagen	Lightning
2	Leiten Vergleich Isolieren Deckel	Thermos

D

Offline Comprehension Tests

D.1 Offline comprehension test history (Experiment 1)

Retention questions

1.) Was waren die Hauptziele der Alliierten, die auf der Konferenz von Jalta im Februar 1945 beschlossen wurden?

4 D's: Denazifizierung, Demilitarisierung, Demontage, Demokratisierung

2.) Welche der vier Siegermächte war auf der Konferenz von Potsdam im Sommer 1945 nicht anwesend?

Frankreich

3.) Was wurde auf der Konferenz von Potsdam 1945 beschlossen?

Einteilung Deutschlands in 4 Besatzungszonen und Teilung Berlins in 4 Sektoren

4.) Wie hieß das oberste Regierungsorgan, das von den Alliierten eingerichtet wurde?

Alliierter Kontrollrat

5.) Was setzten die Amerikaner durch, um den Sowjets den Abtransport von Maschinen und Geräten im Ruhrgebiet zu verwehren?

Jede Besatzungsmacht sollte nur Zugriff auf ihrer eigene Zone haben

6.) Wen oder was fürchteten die Briten mehr als ein wiedererstarktes Deutschland?

Die russische Gefahr/den Kommunismus

7.) Was taten die Briten, um den Wiederaufbau in Deutschland voranzubringen?

Drosselten die Reparationen und erhöhten die Lebensmittelrationen

8.) Wen oder was forderte Frankreich für das Ruhrgebiet?

Eine internationale Kontrolle der ehemaligen Waffenschmiede Deutschlands

9.) Was hatten die französische und die amerikanische Besatzungspolitik gemeinsam?

Aufbau einer gezielten Schul- und Jugendpolitik, Aufbau eines langfristigen Demokratisierungsprozesses

10.) Was tat Frankreich im Widerspruch zum Potsdamer Abkommen?

Kapselte seine Besatzungszone von den anderen ab und gliederte das Saargebiet in das Währungs- und Wirtschaftsgebiet Frankreichs ein

11.) Wer oder was trennte laut Churchill die französische Zone von den anderen Besatzungszonen?

Ein seidener Vorhang

12.) Welche Besatzungsmächte schlossen sich zur Bi-Zone zusammen?

USA und GB

13.) Welche Besatzungsmacht verließ das oberste Regierungsorgan der Alliierten?

Sowjetunion

14.) Wann wurden die beiden deutschen Staaten BRD und DDR gegründet?

1949

15.) Welche zwei Parteien vereinigten sich im April 1946 zur SED?

KPD und SPD

16.) Die SPD stimmte der Vereinigung von SPD und KPD zu, weil sie gegen den von den westlichen Besatzungszonen geförderten Kapitalismus war.

Richtig, die Vereinigung von SPD und KPD wurde zwar von der KPD erzwungen, dennoch stimmte die SPD zu, um eine breitere Front gegen den Kapitalismus zu bilden

17.) Der Alltag für die Deutschen in der französischen Besatzungszone gestaltete sich als sehr angenehm.

Falsch, die Franzosen führten eine harte Besatzungspolitik durch

18.) Das Ziel der amerikanischen Besatzungspolitik war die Unterdrückung der Deutschen.

Falsch, die Deutschen spürten die Besatzer in dieser Zone weniger als in den anderen Zonen. Ziel der amerikanischen Besatzungspolitik war die Demokratisierung der Zone

Global bridging inferences questions

1.) Inwiefern unterschieden sich die Siegermächte in ihrer Auffassung von den Reparationen?
USA und GB: Zurückhaltung, F und Sowjetunion: intensive Demontagen

2.) Was war kennzeichnend für die einzelnen Besatzungspolitiken der vier Siegermächte? Nennen Sie pro Besatzungsmacht zwei charakteristische Merkmale für die Besatzungspolitik!

USA: Demokratisierung, Schülerverwaltung, erste Bürgermeister- und Gemeinderatswahlen, Schaffung der vier Bundesländer Bayern, Hessen, Württemberg-Baden und Bremen

GB: Ziel: Zerschlagung des preußischen Militarismus, Deutschland als Bollwerk gegen den Militarismus zu zerschlagen, Einsatz für den Wiederaufbau, Drosselung der Reparationen und Erhöhung der Lebensmittelrationen, Tolerierung der neu gegründeten deutschen Parteien, Behandlung Deutschlands wie britische Kolonien mit wenig Spielraum zur Eigenverantwortung

Frankreich: bestimmt vom starken Sicherheitsbedürfnis, wollte ein wirtschaftlich schwaches Deutschland mit zerstückelten wirtschaftlich schwachen Einzelstaaten, verhinderte jede zonenübergreifende Einrichtung, forderte internationale Kontrolle des Ruhrgebietes, der ehemaligen Waffenschmiede Deutschlands, kapselte das Saargebiet im Widersprich zum Potsdamer Abkommen von den anderen ab und gliederte das Saargebiet in die Währungs- und Wirtschaftszone Frankreichs ein, „seidene Vorhang", durch gezielte Schul- und Kulturpolitik einen langfristigen Demokratisierungsprozess einleiten

UDSSR: Enteignungen und Verstaatlichung von Banken und Industrien, zentralmonopolistische Wirtschaft, Bodenreform als Vorstufe zur Kollektivierung der Landwirtschaft, Vereinigung SPD & KPD

3.) Wie hieß der Plan zur Wiederaufbauhilfe der USA?
Marshall-Plan

4.) Mit welcher Metapher bezeichnete Churchill die Grenze zwischen den westlichen und östlichen Gebieten bzw. Besatzungsmächten?
Eiserner Vorhang

5.) Es war schon vor 1949 abzusehen, dass Deutschland in zwei Staaten geteilt werden sollte.
Richtig, Churchills Aussage 1946, unvereinbare Besatzungspolitiken von Ost und West, Marshallplan-Hilfe nur im Westen, unterschiedliche wirtschaftliche Entwicklungen

6.) Die Alliierten hatten nach der Kapitulation Deutschlands ein einheitliches Interesse, Deutschland wieder aufzubauen.

Falsch, konnten sich auf der Potsdamer Konferenz nur auf die 4 D's und die Nürnberger Prozesse einigen

7.) Die Amerikaner erkannten, dass nur ein wirtschaftlich starkes Deutschland in Verbindung mit einem starken wirtschaftlichen Europa ein wirksames Bollwerk gegen den „Sowjetkommunismus" bilden konnte und setzten sich im Gegensatz zu Frankreich und der Sowjetunion für den wirtschaftlichen Wiederaufbau Deutschlands ein.

Richtig, die USA wollte gegen den Sowjetkommunismus vorgehen und wollte dass Deutschland ein stabiler Handelspartner fue die USA wird, im Gegengewicht zu der Sowjetunion

8.) Die vier Besatzungsmächte planten von Anfang an, unterschiedliche Besatzungspolitiken für Deutschland.

Falsch, sie hatten unterschiedliche Interessen, planten aber zunächst ein einheitliches Rahmenprogramm – Konferenz von Potsdam, 1945- und einigten sich auf die 4 D's, ursprüngliche Absicht, Deutschland als Ganzes zu regieren

9.) Die Amerikaner führten parallel zur Bodenreform in der Sowjetischen Besatzungszone eine Bodenreform in ihrer Besatzungszone durch.

Falsch, Amerikaner wollten Demokratisierung und Kapitalismus, keine Bodenreform

10.) Frankreich veränderte seine Besatzungspolitik zwischen 1945 und 1949, da es sein Separierungsstreben nicht gegen die anderen drei Besatzungsmächte durchsetzen konnte.

Richtig, zunaechst wollte Frankreich sich von den anderen drei Besatzungsmaechten abkapseln und annektierte im Widerspruch zu dem Potsdamer Abkommen das Saargebiet; Frankreich erkannte, dass es gegen die anderen Besatzungsmaechte sich nicht durchsetzen konnte und aenderte seine Besatzungspolitik, so trat es zum Beispiel der Bi-Zone bei

11.) Die britische und die amerikanische Besatzungspolitik unterschieden sich vor allem in dem Ausmaß der den deutschen zugestandenen Eigenverantwortlichkeit.

Richtig, die Amerikaner wollten die Deutschen eigenverantwortlich handeln lassen, die Briten hingegen wollten die Faeden selber in der Hand behalten, aehnlich ihrer Verwaltungspraxis der Kolonien

12.) Es kam u. a. zur langjährigen Teilung Deutschlands, da sich die Besatzungspolitiken sowie die dahinter stehenden politischen Systeme so sehr voneinander unterschieden, dass daraus keine einheitliche Wirtschafts-, Staats und Gesellschaftsordnung mehr entstehen konnte.

Richtig, die unterschiedlichen Interessen der Besatzungsmaechte verhinderten ihre urspruengliche Absicht, Deutschland als Ganzes zu regieren.

13.) Der Marshall-Plan sollte die Deutschen von Amerika abhängig machen.

Falsch, der Marshall-Plan war eine Überlebenshilfe für die hungernde Bevölkerung und sollte die Bevölkerung wirtschaftlich unterstützen

14.) Frankreich gliederte das Saargebiet in das Wirtschafts- und Währungsgebietes Frankreichs ein, um seine eigene Vormachtstellung unter den Besatzungsmächten auszubauen.

Falsch, Frankreich hatte keine Vormachtstellung unter den Besatzungsmächten und hatte Angst vor einer zentralen, zonenübergreifenden Einrichtung. Frankreich wollte seine eigene Position durch die Eingliederung des Saarlandes stärken.

Elaborative inferences questions

1.) Unter anderem führte die Abschottung der einzelnen Besatzungszonen zu einer Versorgungkrise.
Richtig, da die sowjetische Zone sich von den westlichen Zonen abkapselte und keine Hilfe aus dem Marshall-Plan annahm, kam es zu einer Versorgungskrise

2.) Die Sowjetunion trat im März 1948 u. a. aus dem Alliierten Kontrollrat aus, um die Gründung von zwei deutschen Staaten zu ermöglichen.
Falsch, die Sowjetunion trat aus, weil die Vier-Mächte-Verwaltung endgültig gescheitert war und kein gemeinsamer Nenner zwischen den östlichen und westlichen Besatzungsmächten mehr bestand.

3.) Da Stalin an der Kontrolle des Ruhrgebietes beteiligt werden wollte, setzte er sich für eine wirtschaftliche Einheit Deutschland zunächst stark ein.
Richtig, Stalin setzte sich für *eine wirtschaftliche Einheit Deutschlands stark ein, war aber nicht bereit, auf seine hohen finanziellen Forderungen zu verzichten.*

4.) Der Marshall-Plan förderte indirekt die Entstehung von zwei deutschen Staaten.
Richtig, da die Sowjetunion Hilfe aus dem Marshall-Plan fuer ihre Besatzungszone ablehnte und die drei weistlichen Besatzungszoenen die Hilfe annehmen, entwickelten die jeweiligen Besatzungszonen uneinheitlich

5.) Frankreich trat erst 1949 der Bi-Zone bei, da Frankreich aus Angst vor Großbritannien und den USA nicht von Beginn an mit ihnen zusammenarbeiten wollte.
Falsch, Frankreich wollte eine zonenübergreifende wirtschaftliche Einrichtung verhindern

6.) Die Amerikaner änderten ihre Besatzungspolitik und solidarisierten sich mit den Deutschen, um u. a. den Kommunismus weltweit einzudämmen.
Richtig, die Amerikaner fuerchteten den sowjetischen Kommunismus und wollten ein kapitalistisches Gegengewicht schaffen

7.) Ziel der Bodenreform im September 1945 war die Schaffung lebensfähiger landwirtschaftlicher Betriebe.
Falsch, die Bodenreform war eine Vorstufe zur Kollektivierung der Landwirtschaft und der Anfang des Kommunismus in der DDR

Note: Correct answers are printed in italics.

D.2 Offline comprehension test physics (Experiment 3)

Thermos

Leiten

Frage 1: Behälter aus verschiedenen Materialien (Holz, Gummi, Glas, Metall etc.) sind unterschiedlich gut geeignet, um heiße Flüssigkeiten warm zu halten. Warum ist dies so? Heiße Flüssigkeit erkaltet, weil ihre Wärme an die Umgebung abgegeben wird, wobei Wärme jedoch nur von Materie geleitet werden kann…

Antwort 1: *…und Materialien sich in ihrer Wärmeleitfähigkeit unterscheiden (Metall ist z.B. ein besserer Wärmeleiter als Holz).*

Antwort 2:…und Materialien sich in ihrer Dichte unterscheiden (Metall hat z.B. eine höhere Dichte als Holz).

Antwort 3:…und Materialien sich in ihrem spezifischen Gewicht unterscheiden (Metall hat z.B. ein höheres spezifisches Gewicht als Holz).

Antwort 4:...und Materialien sich in ihrer elektrischen Leitfähigkeit unterscheiden (Metall hat z.B. eine höhere elektrische Leitfähigkeit als Holz)

Entweichen

Frage 2: Stellen wir uns eine offene Teetasse aus z.B. aus Holz vor. In ihr bleibt der Tee nicht lange heiß. Warum nicht?

Antwort 1: Pro Quadratzentimeter entweicht der Großteil der Wärme aus dem Tee, indem der Tee die Wärme an das ihn umgebende Holz abgibt. Der Tee wird kalt.

Antwort 2: Der Tee erkaltet, weil seine Moleküle mit der Zeit immer langsamer wurden und die mittlere Bewegungsintensität der Moleküle einer Flüssigkeit ihrer Temperatur entspricht.

Antwort 3*: Pro Quadratzentimeter entweicht der Großteil der Wärme aus dem Tee, indem der Tee die Wärme in die Luft über ihm gibt. Die warme Luft steigt auf und der Tee wird kalt.*

Antwort 4: Der Tee wird kalt, weil das Holz der Teetasse die Wärmestrahlung, die vom Tee ausgeht, absorbiert.

Luft

Frage 3: Stellen wir uns einen Glasbehälter vor, auf dem zwei Glasplättchen liegen, die durch kleine Korkstückchen voneinander getrennt sind. Was befindet sich hauptsächlich zwischen den beiden Glasplättchen? Zwischen den beiden Glasplättchen befindet sich hauptsächlich…

Antwort 1: …ein Vakuum
Antwort 2: …*normale Luft*
Antwort 3: …Luft mit einem besonders hohen Sauerstoffanteil
Antwort 4: …Kohlenstoffdioxid

Isolieren

Frage 4: …Stellen wir uns heißes Wasser in zwei Glasbehältern vor. Auf dem einen Glasbehälter liegt ein Glasplättchen. Auf dem anderen Glasbehälter liegen zwei Glasplättchen, die durch kleine Korkstückchen voneinander getrennt sind. Legt man ein Stück Butter jeweils auf das obere Glasplättchen, so schmilzt die Butter auf dem einzelnen Glasplättchen, auf den beiden Glasplättchen mit den Korkstückchen dazwischen als Abstandshalter dagegen nicht. Warum nicht?

Antwort 1: Zwischen den Glasplättchen wirkt eine Kapillarkraft, die eine Luftströmung erzeugt, durch die die Butter von unten gekühlt wird.
Antwort 2: Durch die doppelte Glasscheibe entsteht eine Art Spiegel, der die von unten aufsteigende Wärmestrahlung reflektiert und so als Isolator verhindert, dass die Butter erwärmt wird.
Antwort 3: Durch den Abstand zwischen den beiden Glasplättchen kann sich die erwärmende Luft zwischen ihnen zur Seite ausdehnen und wird so abgeführt, ohne die weiter oben liegende Butter zu erwärmen.
Antwort 4: *Durch die Korkstückchen ist Luft als Isolator zwischen den beiden Glasplättchen. Deshalb wird die Wärme nicht bis zur Butter weitergeleitet.*

Oben

Frage 5: In ein offenes Glas wird heißes Wasser eingefüllt. Nach 15 Minuten ist das heiße Wasser jedoch um 20 Grad abgekühlt. Was passiert mit der Wärme?

Antwort 1: *Die warme Flüssigkeit erwärmt die sich im Glas befindende Luft, die nach oben steigt und damit aus dem Glas entweicht.*
Antwort 2: Da sich Wärme immer nur nach oben ausbreiten kann, entweicht die Wärme nach oben aus dem Glas.
Antwort 3: Da Wärme sich nicht durch Luft ausbreiten kann, entweicht sie über das Glas in die Umgebung.
Antwort 4: Die Wärme wandert ins Glas und befindet sich daher nicht mehr im Wasser.

Deckel

Frage 6: Stellen wir uns zwei Gläser vor. In beiden Gläsern ist heißes Wasser. Auf das eine Glas wird eine Pappe gelegt. Auf das andere nicht. Das Wasser im Glas ohne Pappe ist nach 15 Minuten um 12 Grad kälter als das Wasser im Glas mit Pappe. Warum?

Antwort 1: Der Deckel aus Pappe wird von der äußeren Umgebungsluft erwärmt und leitet die Wärme ins Glasinnere. So bleibt das Wasser länger warm.

Antwort 2: *Der Deckel aus Pappe wirkt als Isolierung und verhindert, dass die Wärme aus dem Glas entweichen kann.*

Antwort 3: Unter der Pappe sammelt sich heißer Wasserdampf. Die Wärme des Wasserdampfs wird wieder zurück ins Wasser geleitet und das Wasser bleibt somit warm.

Antwort 4: Die helle Pappe wirkt wie ein Spiegel und reflektiert die aus dem Wasser aufsteigende Wärmestrahlung zurück ins Wasser. Das Wasser bleibt somit warm.

Reflektieren

Frage 7: Ein Glas wird mit Alufolie umwickelt und vor eine Wärmelampe (Infrarotlicht) gestellt. Ein zweites Glas wird ohne Alufolie vor die Wärmelampe gestellt. Wenn man jetzt jeweils ein Stück Butter in das Glas legt, schmilzt die Butter in dem Glas ohne Alufolie, in dem Glas mit Alufolie dagegen nicht. Warum?

Antwort 1: Die Alufolie schafft ein Luftpolster um das Glas mit der Butter, das wie ein Isolator wirkt. So wird die Wärme nicht an die Butter weitergeleitet.

Antwort 2: Die Butter ist kalt. Die Kälte kann aus dem Glas nicht entweichen, weil die Kälte von der Alufolie reflektiert wird.

Antwort 3: *Die Alufolie reflektiert die Wärme von der Wärmelampe und verhindert so, dass Wärme an die Butter in dem Glas weitergeleitet wird.*

Antwort 4: Die Alufolie schafft ein Luftpolster um das Glas mit der Butter, da wie ein Isolator wirkt. So kann die Kälte von der Butter nicht aus dem Glas raus und die Butter schmilzt nicht.

Lightning

Reibung

Frage 8: Mit einem Wollpullover kann man einen Luftballon elektrostatisch aufladen. Wie funktioniert das?

Antwort 1: *Durch Reibung von Pullover an Ballon treten Elektronen vom Pullover auf den Ballon über, weil die Elektronen vom Pullover weniger stark festgehalten werden als vom Ballon. Durch den Elektronenaustausch entsteht eine elektrostatische Aufladung.*

Antwort 2: Durch Bewegung von Pullover an Ballon entsteht kinetische Energie. Diese Energie beruht auf einer erhöhten Bewegungsintensität der Moleküle, die ein Elektronenungleichgewicht zwischen Ballon und Pullover bewirkt. Das führt zur elektrostatischen Aufladung.

Antwort 3: Da der Ballon und der Pullover eine unterschiedliche Elektronendichte aufweisen, genügt ein gegenseitiges Berühren der beiden, um einen Austausch von Elektronen hervorzurufen und damit zu elektrostatischer Aufladung zu führen.

Antwort 4: Durch Reibung von Pullover an Ballon entsteht ein Reibungswiderstand, der zu einer Potenzialdifferenz zwischen Ballon und Pullover führt. Um dies auszugleichen, werden Elektronen ausgetauscht und es kommt zu elektrostatischer Aufladung.

Anziehung

Frage 9: Wenn man einen mit Elektrizität geladenen Ballon an die langen Haare einer Person hält, richten sich die Haare in Richtung des Ballons auf. Warum?

Antwort 1: Der Luftballon und die Haare sind beide negativ geladen und ziehen sich daher an.

Antwort 2: Der Luftballon ist durch seine höhere Dichte wärmer als die Haare und bewirkt dadurch, dass die Haare sich aufrichten.

Antwort 3: Der Luftballon erhitzt sich durch die Aufladung mit Elektrizität. Dadurch steigt die erhitzte Luft in der Umgebung des Ballons auf und die kalten Haare bewegen sich auf den Luftballon zu.

Antwort 4: *Der Luftballon und die Haare haben eine unterschiedliche elektrische Aufladung und ziehen sich daher an.*

Elektrizität

Frage 10: Welche „Hauptzutat“ ist unabdingbar für die Entstehung eines Blitzes?

Antwort 1: Elektromagnetismus
Antwort 2: Wasser
Antwort 3: *Elektrizität*
Antwort 4: Temperaturunterschiede von Luftschichten

Spannung

Frage 11: Wenn man bei einem Bandgenerator die Kugel auflädt und dann eine Metallspitze in die Nähe hält, entsteht ein kleiner Blitz. Warum?

Antwort 1: *Die elektrische Spannung, die zwischen Kugel und Metallspitze entsteht, entlädt sich.*
Antwort 2: Es entsteht magnetische Aufladung, durch die die Elektronen auf der Kugel abgestoßen werden und zur Metallspitze wandern.
Antwort 3: Es entsteht magnetische Aufladung, durch die die Elektronen auf der Metallspitze von der Kugel angezogen werden.
Antwort 4: Die Oberflächenspannung der Kugel erhöht sich stark, so dass die Elektronen in die Metallspitze abgegeben werden.

Entladen

Frage 12: Fallende Regentropfen laden sich durch Reibung an der Luft mit Elektrizität auf. Dabei kann ein Blitz entstehen. Welche der folgenden vier Begründungen für die Blitzentstehung ist richtig?

Antwort 1: *Die Speicherkapazität der Regentropfen für Elektrizität ist begrenzt und wenn die Speicherkapazität erschöpft ist, kommt es zur Entladung in Form eines Blitzes.*
Antwort 2: Fallende Regentropfen vergrößern ihr Volumen. Wenn sie eine kritische Größe erreicht haben, teilen sie sich und ein Blitz entsteht.
Antwort 3: Sobald die fallenden Regentropfen eine kritische Entfernung zur Erde unterschritten haben, kommt es zur Entladung in Form eines Blitzes.
Antwort 4: Fallende Regentropfen erwärmen sich durch Erreichen von wärmeren Luftschichten. Wenn eine kritische Temperatur und damit eine kritische Bewegungsintensität der Moleküle erreicht wird, kommt es zu einem Blitz.

Einschlagen

Frage 13: Warum wirkt eine geerdete Metallspitze, die über ein Gebäude hinausragt, als Blitzableiter?

Antwort 1: Der Blitzableiter ist positiv geladen und zieht daher den negativ geladenen Blitz an.

Antwort 2: *Der Blitz sucht sich im Allgemeinen einen gut leitenden Weg zur Erde. Daher schlägt er bevorzugt in hochragende metallische oder feuchte Gegenstände ein.*

Antwort 3: Damit ein Blitz einschlagen kann, wird ein Gegenstand mit kleiner Oberfläche benötigt, in die sich die Elektronen, die sich in der Spitze des Blitzes konzentrieren, auf engstem Raum entladen können.

Antwort 4: Die Spitze des Blitzableiters ist magnetisch und zieht daher den Blitz an.

Note: Correct answers are printed in italics.

The naming words from Experiment 2 are presented above the question number. As you can see, the naming word appeared twice in the answer alternatives: once in the correct answer and once in a distractor.