"Reading Strategies

in

Multimedia Learning"

A thesis submitted for the degree of Doctor of Philosophy in Psychology

By Fang Zhao

First supervisor: Emeritus Prof. Dr. Wolfgang Schnotz

Second supervisor: Prof. Dr. Robert Gaschler

Department of Psychology

University of Koblenz-Landau

Landau in der Pfalz, Germany

22 July 2015

Acknowledgement

Many people supported me during my doctoral period at the University of Koblenz-Landau in Germany. Foremost, I would like to give my gratitude to my first supervisor, Dr. Wolfgang Schnotz. He trusts me and supports me through the whole way. When I wrote my first article, he passed his experiences about publication on me. Without "standing on his shoulders", I would not finish my PhD within three years. He is also very generous and gave me chances to go to conferences and stay in Dr. Mary Hegarty and Dr. Richard Mayer's research groups in University of California, Santa Barbara, USA. I often marvel at his passion and dedication to research on various topics, linguistics, physics, mathematics, etc. Even after his retirement, he is still active in doing research and supervising many PhD students. I am incredibly grateful to have Wolfgang as my first supervisor.

I would also like to express my heartfelt gratitude to my second supervisor, Dr. Robert Gaschler for mentoring me. As a young and knowledgeable professor, he is a friend for me rather than a supervisor. He always cheered me up and looked out for me while I had difficulty in my study and life. Robert has his specific research focus on learning and strategy change in skill acquisition in cognitive psychology. I was lucky that I can extend my knowledge in psychology by attending his lectures and visiting his research group in Berlin and in Hagen. I am thankful that he shared his wisdom with me, facilitating me to grow up as a psychologist.

I am also greatly indebted to the useful suggestions from Dr. Mary Hegarty and Dr. Richard E. Mayer from the USA. They were extremely supportive and advised me during my research stay there. I learned not only how coordinate they are to junior researchers but also how passionate they are to advancing the research. Apart from this, thanks for Mary who has continued to hold the SCRAM (Spatial Cognitive Research at Mary's) for 25 years.

I also owe my gratitude to Dr. Inga Wagner, who was instructional to recruit the participants in this study. As an associated member of Upgrade School, I am appreciated that I can make friends with Dr. Heidrun Ludwig, Dr. Axel Zingernagel, Stefan Schumacher, Patrick Löffler and Regina Hoffmann. Thanks for their provision of expertise and methodology support during my time in Landau. Thanks are also due to German Science Foundation (DFG) for the provision of my stipend so that this project can be implemented.

Finally, I would like to thank Robert Szwarc, whom I recently married. Thanks for his steady accompany from England to Germany and technical support (e.g. repaired my dying laptop, fixed my twisted bikes). Thanks for his consolation during the hard time when our relatives passed away. As we are going to start a new journey soon, I know we will support each other side by side and make it through.

Abstract

Placing questions before the material or after the material constitute different reading situations. To adapt to these reading situations, readers may apply appropriate reading strategies. Reading strategy caused by location of question has been intensively explored in the context of text comprehension. (1) However, there is still not enough knowledge about whether text plays the same role as pictures when readers apply different reading strategies. To answer this research question, three reading strategies are experimentally manipulated by displaying question before or after the blended text and picture materials: (a) Unguided processing with text and pictures and without the question. (b) Information gathering to answer the questions after the prior experience with text and pictures. (c) Comprehending text and pictures to solve the questions with the prior information of the questions. (2) Besides, it is arguable whether readers prefer text or pictures when the instructed questions are in different difficulty levels. (3) Furthermore, it is still uncertain whether students from higher school tier (Gymnasium) emphasize more on text or on pictures than students from lower school tier (Realschule). (4) Finally, it is rarely mentioned whether higher graders are more able to apply reading strategies in text processing and picture processing than lower graders.

Two experiments were undertaken to investigate the usage of text and pictures in the perspectives of task orientation, question difficulty, school and grade. For a $2 \times 2(\times 2 \times 2 \times 2)$ mixed design adopting eye tracking method, participants were recruited from grade 5 ($N = 72$) and grade 8 ($N = 72$). In Experiment 1, thirty-six 5^{th} graders were recruited from higher tier (Gymnasium) and thirty-six 5^{th} graders were from lower tier (Realschule). In Experiment 2, thirty-six 8^{th} graders were recruited from higher tier and thirty-six were from lower tier. They were supposed to comprehend the materials combining text and pictures and to answer the questions. A Tobii XL60 eye tracker recorded their eye movements and their answers to the questions.

Eye tracking indicators were analyzed and reported, such as accumulated fixation duration, time to the first fixation and transitions between different Areas of Interest.

The results reveal that students process text differently from pictures when they follow different reading strategies. (1) Consistent with Hypothesis 1, students mainly use text to construct their mental model in unguided spontaneous processing of text and pictures. They seem to mainly rely on the pictures as external representations when trying to answer questions after the prior experience with the material. They emphasize on both text and pictures when questions are presented before the material. (2) Inconsistent with Hypothesis 2, students are inclined to emphasize on text and on pictures as question difficulty increases. However, the increase of focus on pictures is more than on text when the presented question is difficult. (3) Different from Hypothesis 3, the current study discovers that higher tier students did not differ from lower tier students in text processing. Conversely, students from higher tier attend more to pictures than students from lower tier. (4) Differed from Hypothesis 4, 8^{th} graders outperform 5^{th} graders mainly in text processing. Only a subtle difference is found between 5^{th} graders and 8^{th} graders in picture processing.

To sum up, text processing differs from picture processing when applying different reading strategies. In line with the Integrative Model of Text and Picture Comprehension by Schnotz (2014), text is likely to play a major part in guiding the processing of meaning or general reading, whereas pictures are applied as external representations for information retrieval or selective reading. When question is difficulty, pictures are emphasized due to their advantages in visualizing the internal structure of information. Compared to lower tier students (poorer problem solvers), higher tier students (good problem solvers) are more capable of comprehending pictures rather than text. Eighth graders are more efficient than 5^{th} graders in text processing rather than picture processing. It also suggests that in designing school curricula, more attention should be paid to students' competence on picture comprehension or text-picture integration in the future.

Key words: text-picture integration; task orientation; distinction between text and pictures; school tier; grade; science textbook; eye tracking

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Displaying data in pictorial form has a long tradition (Playfair, 1786; Thorndike, 1911; Tilling, 1975; Tukey, 1977). Nowadays verbal presentations and pictorial presentations are widely used in learning materials in school science textbooks (Kress and Leeuwen, 1996; V. R. Lee, 2009). This requires students to apply appropriate strategies to connect text and pictures (i.e. graph, line and bar) and to construct knowledge from them (Rouet, 2001). Abundant research has shown that text processing differs from picture processing. (Glenberg, Meyer, and Lindem, 1987; Levie and Lentz, 1982; J. R. Levin, Anglin, and Carney, 1987; Reid, Briggs, and Beveridge, 1983). This phenomenon can be explained by the different features of text and pictures. As distinguished by Schnotz (2014), text is categorised as a descriptive representation, which conveys meanings by using arbitrary symbols in rule-governed structure. Thus, text has an advantage in conveying abstract meanings. Pictures are regarded as depictive representations, which visualize the information directly or metaphorically. Hence, pictures have priority in displaying the complete information. Generally, descriptive and depictive representations are considered to be complementary (Goodman, 1978). In real life, text and pictures are normally presented together to compensate for their disadvantages in conveying information (Tversky and P. Lee, 1999). Text presents information with various pictorial elements. English language, for example, uses space between every word, punctuation between every sentence and indentation between every paragraph. Pictures are always accompanied by auxiliary text. Imagine a map:

the legend or title of the map is normally presented in written language. Learning from materials combining text and pictures is the core of multimedia learning (Mayer, 1997). As a consequence, this study on multimedia learning mainly focuses on learning with text and pictures.

Task orientation, posing questions *before* or *after* the reading materials, consititute different reading situations. Corresponding to the reading situations, readers may apply different reading strategies (André, 1979; Hamilton, 1985; Rickards, 1979; Rothkopf and Bisbicos, 1967). When readers try to solve a task, they conduct internal and external processing: identifying the relevant information (i.e. the information demanded to solve the task) by mental search. They build mental models (i.e. analogue to the real world) to solve the question from the internal mental search or from the external textual search (Goodman, 1978; Graesser, Gordon, and Brainerd, 1992; Guthrie and Kirsch, 1987; Mosenthal and Kirsch, 1991). In this circumstance, reading may be navigated by the task. The reader may adopt different reading strategies instead of just processing the content and answering the question (Cerdán, Vidal-Abarca, Martinez, Gilabert, and Gil, 2009). Reading strategy caused by task orientation has been intensively investigated in the context of verbal learning (R. C. Anderson and Pearson, 2002; Frase, 1967; Rothkopf, 1964; Rouet and Britt, 2011). However, reading strategy is rarely mentioned in multimedia learning (text vs picture).

The presented study mainly investigates whether the usage of multimedia presentations (text vs. pictures) involves a genuine difference with different reading strategies caused by task orientation. In other words, the study examines whether readers deal with text and pictures in the same way when questions are posed before or after the material from school science textbooks. In order to further explore this topic, the study also concerns with question difficulty, school tier and grade. It examines whether readers have different preference to text or to pictures when questions are at different difficulty levels. Furthermore, it also explores whether students from higher tier school (Gymnasium) differ from students from lower tier school (Realschule) in using text and pictures. Lastly, it is supposed to find out whether students from higher grade differ from students from lower grade in text processing and picture processing. Additionally, the purpose of the study is to inspire the future research and to indicate pedagogical implications for school science curriculum. In short, the study aims at exploring whether text plays different roles from pictures with different reading strategies in the perspectives of question difficulty, school tiers and grade.

The first part of the thesis reviews the relevant literature and proposes research questions and hypotheses. The second part describes the pilot study, which intends to examine the feasibility of the study and to explore certain issues before carrying out a large-scale study. In the third part, Experiment 1 investigates whether 5^{th} graders comprehend text differently from pictures in perspectives of task orientation, question difficulty and school tier. The third part explains whether 8^{th} graders[1] perceive text differently from pictures in the same perspectives. Finally, the comparison between two experiments explores the differences between 5^{th} graders and 8^{th} graders in using text and pictures in order to suggest educational implications.

1.1 Pictures have a longer history than text

Since ancient time, graphics has been used to convey meaning by visualizing the spatial information of objects (e.g. maps) (Tversky, 1995, 2001). The record of graphics is much earlier than language. The primitive people in Frace are considered as the earliest people to draw graphics. They painted hundreds of animals in caves before 30,000 BC (Chauvet, Brunel, and Hillaire, 1996). Many unearthed relics have depictions (i.e. painting on ironware, carving on stones) from ancient cultures. These depictions were

[1] Due to the difficulty of recruiting 8^{th} graders, 7^{th} graders were also recruited (12.5% among 7^{th} graders and 8^{th} graders). Compared with 8^{th} graders, 7^{th} graders did not show any significant difference in intelligence test and fixation patterns. Therefore, 7^{th} graders and 8^{th} graders were emerged into one group "8^{th} graders".

mainly used to record for historical, personal and political information (Gelb, 1963; Schmandt-Besserat, 1992; Goulmas, 1989). The best examples may be the recording of events by the Egyptians. As depictions have advantages in conveying the direct information than language, the recorded information can still be presented today.

Pictures have also been used to convey abstract concepts from ancient time (Friendly, 2005). For instance, hunters use symbols of drawing to remember the way. However, the idea of using pictures to present information that conveys the non-spatial relations of information is rather new. Graphs, for example, were firstly used in late 18^{th} century to depict abstract meaning (Beniger and Robyn, 1978; Carswell and Wickens, 1998; Friendly, 2008). As a pioneer in using graphs, Playfair (1786) presented the economic and political data with X-Y axis, which is still commonly used in graphs. The period of 1860 to 1890 witnessed the enthusiastical usage of graphs in government documents and in scientific books. Due to the enormous increase of using graphics, this period of time is considered as an "Golden Age of graphics" (Funkhouser, 1937, pp. 330).

Regarding directionality, pictures can convey the information not only in a temporal way but also in a spatial way (Tversky, 1991). Through visual elements and spatial relations among them, pictures convey spatial or non-spatial relations of information (Larkin and Simon, 1987; Tversky, 2001; Winn, 1989). However, it is not always easy to comprehend pictures although they visualize the spatial or non-spatial relations of information. Readers need to cognitively process the details and the relationships between elements, in order to understand the pictures (Hegarty and Just, 1993). As a consequence, additional text is presented to guide the processing of pictures.

In contrast, the history of written languages are much shorter than pictures. Many unearthed relics confirm that written languages come from pictures (Tversky, Morrison, and Betrancourt, 2002). Until now, there are still languages like Chinese, which come from highly abstract pictures. The communication records using visual marks can be easily found in ancient cultures. These recorded written remains from primitive people show a system how written languages develop and change into their current forms.

Unlike pictures, which are iconic; most written languages are considered as arbitrary symbols (Chomsky, 1957). The relations of the written languages and the real objects are based on random choices. However, many unearthed relics confirm that written languages come from pictures (Tversky et al., 2002). The communication records using visual marks can be easily found in ancient cultures. The earliest true writing of languages are found in Egypt around 3,200 BC and in China around 1,200 BC (Daniels and Bright, 1996). The recorded written remains from primitive people show a system how written languages develop and change into their current forms (see Figure 1.1)[2]. The written languages develop in a trend to abstract the original pictures by using symbols.

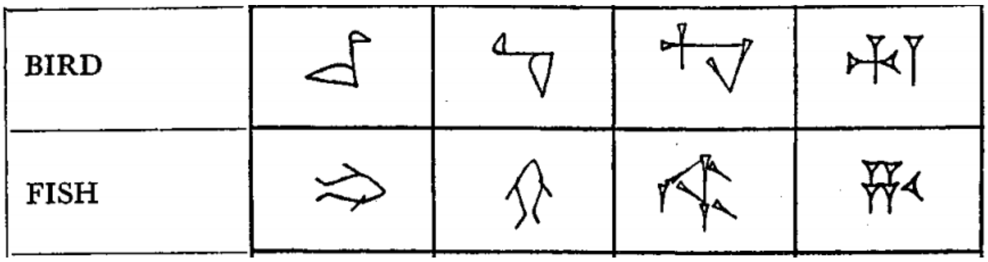


Figure 1.1: Development of Sumerian languages from pictures into highly abstract symbols.

1.2 Pictures are effective in conveying spatial information

Pictures are effective in presenting concrete, abstract or symbolic information (Heiser and Tversky, 2005). Readers can use pictures to abstract the essential information by its spatial structures and conceptual inferences (Larkin and Simon, 1987; Pinker,

[2] This graph is cited from Gelb, I. J. (1963). *A study of writing* (2nd ed., pp.70). Chicago: University of Chicago Press. Copyright by Ignace J. Gelb (1963).

1990). The spatial feature of pictures makes pictures as ideal representations in problem solving. As a landmark in epidemiology (see Figure 1.2), the deaths of cholera in central London was marked by Dr. Snow on the map (Tufte, 1983)[3]. He marked water pumps by using crosses. The complete graph made by him clearly presented that the dead people were mainly scattered around the water pump in Broad Street. Thanks to the help of the graph, he found the cause of cholera- the polluted water pump. Perceiving a picture is like observing an object (Kosslyn, 1976). It is proved by

Figure 1.2: The deaths of cholera and water pump in central London, which was drawn by John Snow in 1854.

neuroimaging data that pictures can activate the same area of the brain as perception

[3] This graph is cited from Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press, pp.24. Copyright by Edward R. Tufte (1983).

(Kosslyn, 1976). When a reader comprehends a graph, s/he creates a visual description of the information by using Gestalt laws (Pinker, 1990). For instance, s/he may group the proximate or adjacent elements in the graph. From the visual description, the reader extracts propositions and create mental model (analogue of information) from the graph. Meanwhile, the extracted mental model will be inherently inferenced by connecting to the prior experience. The constructed mental model can be used to answer questions or solve problems (Kieras and Bovair, 1984).

Pictures can facilitate comprehension and enhance memory (Levie and Lentz, 1982; Denis, 1984; J. R. Levin and Mayer, 1993; Schnotz and Kulhavy, 1994). It can be due to at least 5 function types, that cause advantages in using graphics: attentional (Levie and Lentz, 1982), economic (Tversky et al., 2002; Bauer and Johnson-Laird, 1993), affective (Mayer, 1989; Winn, 1987), cognitive (Yates, 1969; Small, 1997) and compensatory (Winn, 1989) function types. For the attentional function, pictures can be aesthetically appealing. Pictures can motivate readers and guide them to particular points in the text. Pictures can also save words for description and thus be economic. This applies to maps, systems and spatial information that is difficult or confusing to present in words. Besides, pictures can make learning more enjoyable and affective. In the cognitive perspective, pictures can organize information by using spacial elements. Pictures can externalize the underlying structure of the information, which can reduce the cognitive load. Thus, pictures can enhance the processing of information and facilitate memory. Finally, graphics can compensate for the inadequacy of verbal representation by providing the alternative non-verbal representation (Taylor and Tversky, 1992).

Besides, pictures can also enhance visual search by using saliency and spatial metaphors. The salient parts of pictures are effective in guiding readers in visual search, such as colour, contrast and brightness (D. J. Parkhurst and Niebur, 2003). Graphics depend on spatial metaphors when conveying concepts without spatial meaning. The features of graphics can use figures of depiction to convey abstract meanings. Metonymy, for instance, represents abstract concepts through inference. A document folder in Microsoft Word, for instance, represents the function of “opening a folder”.

1.3 Text has advantages in presenting abstract information

In order to convey information with text, readers need to use certain structural or grammatical rules. Take German as an example, the verbs are sometimes presented at the end of the sentence. Mark Twain once wrote about his experience of learning German, “Whenever the literary German dives into a sentence, that is the last you are going to see of him till he emerges on the other side of his Atlantic with his verb in his mouth” [4].

Written languages are normally presented in a temporal way. They convey the information in a linear structure. The new information is displayed after the old information. The sentences are in coherence with the meaning. This structure can assist readers to structure the information in a sequential order (Levelt, 1989). From another perspective, producing information in a sequential order can enhance cognitive processing (Denis and Denhiere, 1990).

Text can present spatial information but it loses the visual and spatial representation of elements in graphics. For instance, it is quite difficult to describe a map or a circuit board with verbal representation. However, text can facilitate the construction of structure. To some extent, text can also be as effective as pictures (Taylor and Tversky, 1992). It can be explained by the study on exploring the spatial power of verbal instructions in installing objects (Tversky, 1991).

[4] Clemens, S. L. (1889). *A Connecticut Yankee in King Arthur’s court* (pp. 195). New York, London: Harper & Brothers.

Although text does not present information with spatial elements, text can also be as effective as pictures in facilitating readers to create situations (Mani and Johnson-Laird, 1982; Taylor and Tversky, 1992). Studies on how readers use and comprehend language demonstrate that readers can use language to construct a situational model or mental model (van Dijk and Kintsch, 1983). Linguists presumed that the mental representation of space is vertical and Euclidean-like (Hirtle and Jonides, 1985). However, it is an misinterpretation about the spatial cognition of language. Compare the following two sentences. (1) Three turtles rested beside a floating log and a fish swam beneath them. (2) Three turtles rested on a floating log and a fish swam beneath them. According to Bransford, Barclay, and Franks (1972), two sentences contain the similar linguistic structures. However, when readers perceive sentence (2), they are likely to think the fish swam also under the log. It suggests that text can be used to create a vivid situation rather than simply vertical space.

1.4 Learning with text and pictures

Abundant research has shown that learning can be generally improved with text and pictures compared to text alone (J. R. Anderson and Biddle, 1975; Glenberg et al., 1987; Levie and Lentz, 1982; J. R. Levin et al., 1987; Reid et al., 1983). This phenomenon is due to the compensatory features of text and pictures in presenting spatial and descriptive information in learning materials (Hegarty and Sims, 1994). Learning from multimedia presentations can build up connections between text and pictures so that learners can understand more deeply than from text or pictures (Mayer, 1997; Mayer, Moreno, Boire, and Vagge, 1999; Mayer, 2009, 2014a). Besides, multimedia presentations can foster comprehension of complex subject, better problem-solving skills and retention of information.

Many theories in multimedia learning address the issue of cognitive processes during learning from text and pictures (Kress and Leeuwen, 1996; Merriënboer and Kirschner, 2007; Rouet, 2001; Sweller, Ayres, and Kalyuga, 2011; Zwaan and Radvansky, 1998). Although they have different foci, all agree with the idea that text and pictures are processed in a fundamentally different way. There are mainly three models suggested how the reader constructs separated mental representations from verbal and pictorial materials. (1) According to the Dual Coding Theory (Paivio, 1986), the reader's cognitive processing is in a hierarchical structure with dual channels. It indicates the referential connections between verbal and pictorial channels. It also assumes that the reader can retrieve information better by using two different channels. (2) Mayer (2005) proposes The Cognitive Theory of Multimedia Learning, in which the reader has limited capacity to establish two separate mental representations from verbal and pictorial information. He summarizes the processes into four steps. (i) Selecting relevant words; (ii) selecting relevant images; (iii) organizing the selected words into a verbal mental model; (iv) organising the selected images into a pictorial mental model; and (v) integrating the verbal model and the pictorial model with prior knowledge into a coherent mental representation. Distinct from other models, (3) the Integrative Model of Text Picture Comprehension (Schnotz and Bannert, 2003) differentiates descriptive (from verbal information) and depictive representation (from pictorial information). It also suggests that readers integrate text and pictures in one mental model, which is later proved by other research (Schüler, Arndt, and Scheiter, 2015). As a consequence, this study is, to a large extent, stimulated by the Integrative Model of TPC.

Although multimedia learning has been a focus for over two decades, learning materials were studied with a limitation on certain tasks in certain subjects. There has also been limited attention given to biology and geography. For instance, the operation of a bicycle tire pump (Mayer and R. B. Anderson, 1992; Tversky, Zacks, Lee, and Heiser, 2000) or an automobile braking system (Johnson and Mayer, 2012; Mayer, 2014a), lightening (Mayer, 2001), pulley system (Hegarty, 1992)or flushing toilet (Eitel, Scheiter, Gerjets, and Kuhl, 2014).

Besides, multimedia learning was mostly studied with elementary school students (Guercin, 2001) and university students (Chiou, Tien, and Lee, 2015; Frase, 1967; Park, Flowerday, and Brünken, 2015; Schüler et al., 2015). Research on secondary school students is still not sufficient. Different from elementary school students and university students, secondary school students are under pubertal period, in which they experience fast body and brain development. "Scores on intelligence tests obtained over several years from the same individual fluctuate most during the period from 12 to 15 years of age" (Slavin, 1997, pp. 96-97). If secondary school students can grasp the correct reading strategy, their learning performance can be enormously enhanced even for the rest of their lives. Therefore, it is necessary and important to study reading strategy in text and pictures comprehension among secondary school students.

1.5 Cognitive processes of question-answering

Reading is usually performed with certain purposes, such as solving a question. Questions can affect text processing. Abundant research has shown that questions fundamentally influence the way of cognitive processing in verbal materials (André, 1979; Hamilton, 1985; Rickards, 1979; Rothkopf and Bisbicos, 1967). Under certain circumstances, adding questions to printed text can induce meaningful processing compared to printed text alone (André and Thieman, 1988; Ausubel, 1968; Dornisch and Sperling, 2006; Hershberger and Terry, 1965; Mayer, 1975; McConkie, Rayner, and Wilson, 1973; Rickards and Di Vesta, 1974; van den Broek, Tzeng, Risden, Trabasso, and Basche, 2001). This is attributed to review or preview function of questions (Frase, 1968b). In this respect, questions can be used as the most compelling stimuli for guiding text comprehension (Rothkopf and Bisbicos, 1967).

Many researchers propose the cognitive processes for question answering with verbal materials (André, 1979; H. H. Clark and E. V. Clark, 1977; Fisher, 1981; Goldman and Durán, 1988). They indicate that questions can influence text processing. The most suggestive models are mathemagenic model, cybernetic model and TRACE model (Rouet, 2006). The mathemagenic model (Rothkopf, 1963) hypothesizes that questions provoke responses in enhancing knowledge acquisition. According to this assumption, the essence of learning lies in reader's responses while reading instead of the provided questions. It suggests that verbal materials can modify the mathemagenic behaviours. The outside stimuli can influence the learner's performance. To some extent, the learner is regarded as a closed- and open-loop system. For closed-loop system, questions play the roles as evaluative criteria to provoke an entry point into cognitive skills. It is demonstrated that questions posed after the materials can partially enhance the appropriate mathemagenic behaviours.

The cybernetic model (Hershberger and Terry, 1965; K. U. Smith and M. F. Smith, 1966) assumes that a reader can use questions to determine whether the criterion of acceptable behaviour (correct answer) is provided in the achieved behaviour (what the reader has learnt). An error is assumed to be signalled, when s/he failed to find the answer after reading the material. The error is negative feedback which the reader uses to control her/his reading behaviour until s/he meets the criterion. The cybernetic model emphasizes reader's control on matching or comparing function of questions (Stolurow, 1961).

Rouet (2006) proposes the TRACE model (task-based relevance and content extraction) for task-solving. It provides detailed processing steps when readers need to solve a task with text. While the readers comprehend the task, they construct the task model, which requires specific information to solve the task. Then, they make decisions to search for information in the text after examining their known information with the task model. If they need to search for the external information, they select the information and assess the available source to identify the relevant information. Later, they probably update a response model, which contains information to solve the task. Afterwards, they examine whether the response model satisfies the needs for solving the task. If the answer is negative, they start another round of searching and assessing.

Although the models in question answering have different emphasizes, they point out that questions can guide readers in text processing. Readers are inclined to construct a criterion from questions, which can largely influence text processing. However, it is rarely mentioned in multimedia learning whether readers comprehend text differently from pictures in regard of questions.

1.6 Effective factors in verbal learning

This study is mainly inspired by the previous research on effective factors in text comprehension and in multimedia learning. It is demonstrated that verbal learning can be influenced by task orientation (Ausubel, 1968; Frase, 1968a; Rickards and Denner, 1978), question difficulty (André, 1979; Graesser et al., 1992), school tier (Cataldo and Oakhill, 2000; Cerdán, Gilabert, and Vidal-Abarca, 2011), and grade (Golden, 1942; van den Broek et al., 2001). Preliminary studies argue that text and pictures are comprehended differently from text alone (Mayer, 2009; Tversky and P. Lee, 1999). It is suggested that task orientation (Zhao, Schnotz, Wagner, and Gaschler, 2014), question difficulty (Hochpöchler et al., 2013), school tier (Schnotz et al., 2010), grade (Schnotz, 2014) can affect multimedia learning. However, with limited sample sizes, research on effective factors in multimedia learning can be biased and is still inadequate. Therefore, my research aims at examining the effects of task orientation, question difficulty, school tier and grade to deepen the understanding of multimedia learning.

(1) Task orientation: question before vs. after material

In this study, task orientation is mainly an issue of posing the question before or after the material. Task orientation (location of questions) was studied in text comprehension as early as 1920s (Distad, 1927). It is indicated that task orientation can cause genuine differences in verbal learning and processing activities (Rickards and Di Vesta, 1974). The reader normally selects and processes the information based on the question demands in task-oriented reading. The demands are defined as *task model* (Rouet, 2006), in which the reader constructs the mental model to meet the requirements of the task model.

According to the TRACE model of question solving (Rouet, 2006), reading situation can be different when questions are placed before or after the material. When a question is displayed before the material, reading may heavily focus on meeting the requirement of questions. Thus, reading can be relatively selective. When a question is placed after the material, reading can be rather general at the beginning. The question is answered based on the existed knowledge of the material and when the mental model about the material meets the requirement of the question. If not, a search for the required information starts, until the requirements for the question is fulfilled. Therefore, the reader may apply different reading strategies to comprehend the materials.

Task orientation or location of questions can influence verbal learning. However, not enough knowledge has shown whether readers use text in the same way as pictures when question is posed before or after the material. Reading is rather general when question is placed after the material (Rothkopf and Bisbicos, 1967). Before the question is displayed, readers may start to construct the mental model of the material. After the question is posed, they may also identify the connections that have not been recognized in the initial reading. In this respect, the acquisition and retention of information can be reinforced by placing questions after the material (Rickards and Di

Vesta, 1974; van den Broek et al., 2001). Posing questions after the materials affects the processing behaviour (Frase, 1967; Rothkopf, 1966). It involves a forward process (establishing or searching for proper reading strategy) and a backward process (reviewing the material with previous experience for answering questions). However, other researchers (Boyd, 1973; Rothkopf and Billington, 1974) argue that questions placed after the materials influences the backward process (reviewing material to answer the question) rather than forward process (eliciting appropriate reading strategy). Nevertheless, they all agree that reading is relatively general when question is posed after the material.

In contrast, reading is comparatively selective when question is posed before the material. The pre-posed questions can guide readers in selecting the required information. Besides, questions can also activate readers' prior experience on the topic (Robertson, Black, and Lehnert, 1985; Rothkopf, 1963). Studies on social science textbooks suggest that readers' prior knowledge can be triggered when questions are placed at the beginning of the chapters or materials (Stolurow, 1961; Washbourne, 1929). According to Rothkopf and Coke (1963), questions can have specific and general facilitative effects on learning. Learner is a guided control system. The reader can respond discriminately to the material with appropriate skills. The control pattern contains the discriminate response to the stimuli (material). The reader corresponds very sensitively to the feedbacks from the instruction. In other words, questions presented at the beginning can provoke the relevant schema (structured knowledge about concepts) from prior knowledge before reading (Mayer, 2008). Their constructed representations are more selective and amenable to the questions (Robertson et al., 1985; van Oostendorp and Goldman, 1999). In short, reading is rather selective when question is posed before the material.

Previous studies argue that text processing is fundamentally different from picture processing (Mayer, 2009; Tversky and P. Lee, 1999). It is rarely mentioned how a reader answers questions by using text and pictures. Preliminary studies (Schnotz, Ludewig, et al., 2014; Zhao et al., 2014) reveal that text and pictures play different roles in cognitive processing. Text mainly guides general reading. It is mainly served as the tool to build the initial mental model and to update the demanded information from questions. Conversely, pictures mainly guide selective reading. They assist the mental model construction and serve as convenient tools to locate the required information from questions. Since this has been rarely investigated, my research mainly focuses on how text, and picture elements are used with different task-orientations.

(2) Question difficulty: easy vs. difficult question

Reading comprehension may be influenced by task orientation. However, text can be processed differently when required questions are at different difficulty levels. Questions can guide the reader to have different levels of text processing with easy questions and difficult questions (Allen, 1970; Ausubel, 1963; Hamilton, 1985; Hunkins, 1968; Rickards, 1979; Tenenberg, 1969; Watts and R. Anderson, 1971). When questions are presented at different difficulty levels, graphics can also be processed at different levels (Bertin, 1967; Pinker, 1990). However, it is arguable whether readers prefer to text or pictures when the questions are easy or difficult. Furthermore, question difficulty has been thoroughly discussed in the domain of verbal learning instead of multimedia learning.

Comprehending verbal materials with easy questions is generally different from with difficult questions. The main difference lies in the reader's depth of processing (André, 1979). The depth of processing suggests that the reader processes information in a hierarchical stages (R. C. Anderson, 1970; Craik and Lockhart, 1972; Lockhart and Craik, 1990). It varies from superficial processing of facts to deep processing of meaning. The depth of cognitive processing can be summarized into three levels (DeLeeuw and Mayer, 2008; Mayer, 2008). (1) *Extraneous processing*, in which a reader processes the irrelevant information to the learning goal. If a reader comprehends the materials in extraneous processing, s/he processes the information superficially. (2) *Intrinsic or essential processing*, in which the reader processes the crucial information to understand the material. (3) *Germane or generative processing*, in which the reader reorganizes the information and relates it with the previous knowledge. If a reader processes the information in intrinsic or essential processing and germane or generative processing, s/he analyses the verbal materials semantically and cognitively at a deeper degree. S/he may process the information by elaborating with her/his prior knowledge.

Difficult questions may facilitate deeper processing and require more information than easy questions (Ausubel, 1963; Briggs and Reed, 1943; Cerdán et al., 2009; Dornisch and Sperling, 2006; English, Welborn, and Killian, 1934; Ozgungor and Guthrie, 2004; Vidal-Abarca and Sanjose, 1998; Wixson, 1983). According to Quest Model (Graesser et al., 1992), easy-level questions majorly test rote learning or single concepts (recall of facts). They engender specific review of the materials. In contrast, difficult questions test the information more than from the direct memory or broad conceptual structures (beyond the recall of facts). It produces both a subsume and a review process. Factual questions contain less cognitive processing than questions requiring meaningful learning. By processing the materials at a deeper level, the reader may integrate them with their prior knowledge and re-construct the information (Tulving and Madigan, 1970). Besides, the reader may pay more attention to the materials when the questions are difficult. Frase and his colleagues propose “directed attention hypothesis” from a series of studies on question difficulty levels in verbal learning (Frase, 1969a, 1969b, 1970a, 1970b, 1972). They find that the reader draws more attention to more information from the verbal materials when the questions are difficult. Consequently, the reader may have the deeper understanding of the materials and pay more directed attention with difficult questions than with easy questions.

However, it is questionable whether the reader can truly generate deep processing when question is difficult. The reader processes the materials faster with easy questions than difficult questions based on matching simple concepts (Rouet, 2001). Comparatively, difficult questions may require the reader to search for a large amount of information (Cerdán et al., 2009). From this point of view, difficult questions can increase reader's frustration and decrease her/his motivation (André, 1979). In most circumstances, difficult questions would result to reader's poor performance. The reader may give up before they perform deep processing. As a result, the reader may not necessarily have deeper processing and may perform worse with difficult questions than easy questions.

In multimedia learning, readers tend to focus more frequently on text when the questions are easy (Hochpöchler et al., 2013; Schnotz, Ludewig, et al., 2014). They tend to focus more frequently on pictures when the questions are difficult. It may due to the distinct features between text and pictures in conveying information. Text has an advantage in describing abstract information. Pictures have priority in visualizing the internal structure. Hence, pictures are suggested as external scaffolds for mental model construction and updates (Eitel et al., 2014). Pictures can also facilitate understanding the difficult materials (Carney and J. R. Levin, 2002). As the questions become more and more difficult, the reader may rely more and more on pictures rather on text to locate the corresponding information (J. R. Levin and Mayer, 1993). Concerning the question difficulty, readers may engage in different processing with easy and difficult questions. Due to the lack of research on question difficulty in multimedia learning, this study also aims at exploring whether readers have different preference to text or pictures.

(3) School tier: good vs poorer problem solver

In Rhineland Palatinate in Germany, secondary schools are divided into higher school tier (Gymnasium) and lower school tier (Realschule). Students from higher school tier are considered to be good problem solvers, which have rigorous academic training for universities (Clements, 2013). Students from lower school tier are regarded as poorer problem solvers, which will have general training for technical colleges. However, the placement of students in these school tiers is far from perfect. The distributions of good or poorer students are overlapping. Lower tier students may perform above the average of the higher tier students. Besides, the school tiers are often influenced by other social factors, such as education or wealth of the parents (Schneider, 2007). In this study, lower tier students are defined as "poorer problem solvers". Yet, it does not mean that the lower tier students are all poorer problem solvers. It indicates that the lower tier students are on average less skilled in academic problem solving than higher tier students.

Good (Gymnasium students) or poorer problem solvers (Realschule students) may have different reading patterns in text processing. Good problem solvers are suggested to process text more efficiently than poorer problem solvers. However, it is still rarely mentioned whether good problem solvers differ from poorer problem solvers in using text and pictures when question is presented before or after the material. In verbal reading, a poorer problem solver normally locates information less efficiently and processes information less deeply than a good problem solver (Cataldo and Oakhill, 2000; Hannon and Daneman, 2004). Regarding the reading patterns, the poorer problem solver often cannot previously elaborate the meaning of single concepts to the global coherent text. S/he is also less able to discard the distracting information than the good problem solver (Cerdán et al., 2011). Opposite to the poorer problem solver, the good problem solver hardly mismatches the information for answering the question. Therefore, the poorer problem solver performs more slowly and worse than the good problem solver.

Furthermore, poorer problem solvers differ from good problem solver in processing text and pictures. It is disputable whether poorer problem solvers benefit more from pictures than good problem solvers (Hochpöchler et al., 2013; Levie and Lentz, 1982;

Schnotz et al., 2010). Some studies show that poorer problem solvers benefit more from pictures than good problem solvers (Cooney and Swanson, 1987; F. Goldberg, 1974; Mastropieri and Scruggs, 1989; Rusted and Coltheart, 1979). It can be explained that poorer problem solvers are not able to access the meaning of words efficiently. They move frequently to the pictures and examine the figures in the pictures during reading than the good problem solvers (Rusted and Coltheart, 1979). They depend on both text and pictures to build the initial mental model. Thus, they may benefit more from multimedia presentations (Mayer, 2001).

Whereas, other studies show the opposite findings (Bluth, 1973; P. E. Parkhurst, 1974) or no difference between two groups (Peeck, 1974). Good problem solvers are more likely to use pictures for mental model construction. They use pictures as scaffolds for initial mental model construction. Afterwards, a picture is applied as an easy access to require mental model updates. Compared to the poorer problem solvers, the good problem solvers invest less time on text and more time on pictures when the questions become more difficult (Hochpöchler et al., 2013; Schnotz, Ludewig, et al., 2014). It is suggested that pictures can be used as external scaffolds for mental model construction and updates (Eitel et al., 2014). Owing to their capability of using pictures, the good problem solvers outperform the poorer problem solvers.

However, it is still inadequate to answer the question whether good problem solvers differ from poorer problem solvers in usage of text and pictures with different task orientation (Ozgungor and Guthrie, 2004). A preliminary study (Washbourne, 1929) assumes that task orientation may be less effective to poorer problem solvers than good problem solvers. Both types of readers may benefit more when questions are posed prior to than after the verbal material. Regarding the difficulty level of questions, it is reported that difficult questions have more effect on poorer problem solvers than good problem solvers (Shavelson, Berliner, Ravitch, and Loeding, 1974). Therefore, this study aims at investigating whether task orientation influences good readers and poorer readers.

(4) Grade: lower vs. higher grader

With a cross-sectional design, this study explores whether there might be an increase in multimedia skills with age. For example, children become better at integrating text and pictures when they grow older. Readers from different grades may have different reading patterns with multimedia presentations (Seretny and Dean, 1986; Tal, Siegel, and Maraun, 1994; van den Broek et al., 2001). The accuracy increases from lower grade to higher grade (Golden, 1942). It takes readers from lower grade than readers from higher grade longer time to comprehend the materials (Schnotz, 2014). It seems that young readers have difficulty in establishing a coherent representation when questions are posed after the material. Young readers' working memory capacity is possibly occupied with superficial level of processing, such as semantic recognition, syntactic decomposition (Carpenter, Miyake, and Just, 1994; Oakhill and Yuill, 1996). Understanding the materials may interfere with integrating the demands from questions, searching for corresponding materials and organizing the information. In contrast, older readers are able to automatically process the information and they have enough space in working memory to comprehend the questions (van den Broek et al., 2001). They can possibly benefit from questions during reading because questions can direct their attention to the requested information.

In addition, younger students have different reading patterns on text and pictures from older students. Younger students tend to pay more attention to text than older students (Hannus and Hyönä, 1999; Hochpöchler et al., 2013). Conversely, they fixate less time on pictures than older students. It seems that young students are less able to amend the processing of pictures to the demand of tasks. As the questions become more and more difficult, young students still have similar performance on pictures. In contrast, older students are more sensitive to identify the demands of tasks. They are more skilled at using pictures to fulfil the tasks (Hartman, 2001). Younger readers may not know how to deal with the pictures. They may be less able to perform the suitable strategy for picture processing (Hasselhorn, 1996). As age is rarely mentioned in multimedia learning, this study also focuses on how younger readers differ from older readers in using text and pictures.

1.7 Reading strategies triggered by task orientation

In order to explain differences between reading strategies when learning from texts, Rickards and Denner (1978) suggest a distinction between general and selective processing. General processing deals with the global thematic coherence of the text, whereas specific processing focuses on unique information required for specific purposes. There seems to be an inherent conflict between these two kinds of processing, as Rickards and Denner find that pre-posed questions can lead to a highly selective processing, but at the expense of global understanding of the text. As far as it is concerned, this distinction between two different kinds of processing has not been applied to the combined processing of text and pictures, yet.

In line with Rickards and Denner (Rickards and Denner, 1978), this study differentiates between *general coherence-formation processing* and *selective goal-oriented processing* of text and pictures. The two kinds of processing are not meant as a strategy dichotomy. Instead, coherence-formation processing and goal-oriented processing are strategy components that can be combined. As time and processing resources are limited, however, the two components cannot be both maximized at the same time. Instead, they can obtain different emphasis in the process of TPC. Thus, there is a continuum with a primarily general coherence-formation processing at the one end and a primarily selective goal-oriented processing at the other end. Depending on the specific learning situation, different kinds of processing will be combined into a suitable strategy. In the following, I will consider three different reading strategies:

(a) If a reader receives a text with pictures without a specific goal in mind, s/he will put the emphasis on general coherence-oriented processing. That is, s/he will try to construct coherent mental representations based on the available information. This kind of processing is defined as the *goal-free processing*.

(b) If a reader has first processed a text with pictures without a specific goal in mind (i.e., applied a goal-free processing) and then gets access to the specific questions to be answered, s/he will put emphasis on selective goal-oriented processing and focus on task-relevant information. This kind of processing is considered as the *delayed goal-oriented processing*.

(c) If a reader is presented specific questions before receiving a text with pictures to be used for answering these questions, s/he will put more emphasis on selective goal-oriented processing. However, although processing is goal-directed from the very beginning on, some coherence-oriented processing is also required because the reader needs some understanding of what the text and the pictures are about. This kind of processing is defined as the *initial goal-oriented processing*.

1.8 Aims and objectives

This study is supposed to examine whether readers use text differently from pictures when they apply different reading strategies. The reading strategies are activated by locating the question either before or after materials. Three other factors are investigated to further explore reading strategy in multimedia learning. To be specific, this study explores whether readers have different preference to text or pictures when question difficulty increases. With cross-sectional design, it compares the usage of text and pictures between good problem solvers and poorer problem solvers. Due to the difficulty of distinction between good problem solvers and poorer problem solvers, this study compares students from higher tier (Gymnasium) and from lower tier (Realschule). Finally, this study discusses whether students' age influences the way of text and picture processing. To simplify the design, students from different grades are recruited and compared.

Four research questions are proposed in respects of task orientation, question difficulty, school tier and grade. In Experiment 1, the effects of task orientation, question difficulty and school tiers are investigated among 5th graders. In Experiment 2, the effects of the same factors are explored among 8th graders. Finally, the usage of text and pictures is compared between 5th graders and 8th graders with different reading strategies.

1. How do readers deal with text and deal with pictures in different forms of reading?

It is hypothesised that task orientation influences comprehension with multimedia presentations at the cognitive level instead of the outcome level (i.e. accuracy). Based on the findings in verbal learning, it was presumed that readers may have similar accuracy results. On the one hand, readers consolidate the acquired information and adapt it to the questions when question is posed after the verbal material (van den Broek et al., 2001). On the other hand, the relevant knowledge structure can be activated when question is posed before the verbal material (Goldman and Durán, 1988). As suggested from the pilot study (Zhao et al., 2014), text mainly guides mental model construction. Hence, readers are assumed to draw more attention to text with goal-free processing (no question yet). Pictures may guide selective reading. When the questions are illustrated, readers may tend to emphasize on pictures with delayed goal-oriented processing (question after material). In contrast, they seem to have rather balanced attention to text and pictures with initial goal-orientated processing, although text is mainly focused. Besides, readers' reading routine is also suggested that pictures are focused earlier than text regardless of reading strategies (Mason, Pluchino, and Ariasi, 2014; Zhao et al., 2014). It is presumed that building mental model is primarily guided by text and using mental model is predominately guided by pictures. Likely, text mainly guides general reading; whereas, pictures primarily guide selective reading.

2. Does preference of readers for text vs. pictures depend on question difficulty?

Readers are expected to have more accurate answers with easy questions than with difficult questions. They may pay more attention to the materials when asked to answer difficult questions rather than easy questions. Text distinguishes from pictures in different features and functions (Stalbovs, Scheiter, and Gerjets, 2015). Text has an advantage in conveying abstract information. Pictures have priority in transferring spatial information by visualizing the structure. Thus, readers may focus mainly on text with easy questions. They may focus mainly on pictures with difficult questions. They may integrate text and pictures less frequently with easy questions than difficult questions.

3. What are the differences in text processing and picture processing between students from different school tiers?

Students from higher school tier (good problem solvers) may outperform readers from lower school tier (poorer problem solvers) in accuracy and reading strategy. Students from higher tier can more automatically process the semantic and syntactic levels of verbal materials than students from lower tier (Rusted and Coltheart, 1979). Students from higher tier are possibly more able to adapt pictures to the demands of questions than students from lower tier. Thus, they possibly have more accurate answers than students from lower school tier. In other words, they probably process the verbal materials more efficiently than students from lower school tiers. They may process the pictorial materials more intensively than students from lower school tiers. Therefore, the key of being a good problem solver may lie in efficient processing of text and pictures.

4. What are the differences in text processing and picture processing between students from different grades?

Students from higher grade (8^{th} graders) may outperform students from lower grade (5^{th} graders) in accuracy and multimedia comprehension. As students learn more from lower grade to higher grade, they may perform tasks more accurately (Golden, 1942). Students from higher grade have less difficulty in decoding and encoding the verbal information than students from lower grade (Oakhill and Yuill, 1996). Hence, students from higher grade may process the verbal information more automatically than students from lower grade. However, they are more skillful in using pictures than students from lower grade (Hartman, 2001). Therefore, students from higher grade may invest less time on text but more time on pictures than students from lower grade.

Chapter 2

Pilot Study

A pilot study (Zhao et al., 2014) was performed in order to examine the feasibility of the study and to modify the deficiency of the design. It aims to explore whether the usage of text differs from the usage of picture when readers follow different strategies of knowledge acquisition. In a within-subjects design using eye tracking, 17 secondary school students comprehended blended text and picture materials with three different processing. (1) Goal-free processing, which requires students to process text and picture purposelessly. (2) Delayed goal-oriented processing, which requires them to gather information to answer the question (which explains the task) provided after prior experience with text and picture. (3) Initial goal-oriented processing, which requires them to comprehend text and picture to solve the task equipped with the prior information about the question from the beginning. Eye tracking data showed that text and picture play different roles in these processing conditions. (1) The results are in line with the assumption that text (rather than picture) is more likely used to construct mental models in goal-free processing of text and picture. (2) Students seem to primarily rely on the picture to answer the question after the prior experience with the material with delayed goal-oriented processing. (3) Text and picture are both used heavily when the question is presented first, enabling students to selectively process question-relevant aspects of the material at first contact.

2.1 Method

2.1.1 Participants

Seventeen participants from secondary schools in Germany were included in this study (M = 13 years, SD = 3.4 years). Eleven participants were male and six were female. Participants were tested on spatial ability and verbal ability (Heller and Perleth, 2000). They were marginally above average of the norm sample in spatial ability (average T of 53.65; SD = 7.39) and verbal ability (average T of 55.59; SD = 9.08).

2.1.2 Materials

In a previous study, 60 text-picture units were selected from textbooks about geography and biology with 288 test questions. The units and questions were tested with 1060 students in grades 5 to 8 according to Item-Response Theory including DIF (Differential Item Functioning)-analyses for gender, grade and school. Additionally, a rational task analysis was performed on questions (Schnotz et al., 2011). In order to answer the questions correctly, participants need to process both text and pictures. As the pilot study adopted eye-tracking method, only a subset of the text-picture units was used for pragmatic reasons. Each participant received six text-picture units. The units were selected in a way that the type of image (realistic pictures vs. graphs) and the level of difficulty (easy vs. medium vs. difficult) were balanced throughout the experiment. Participants were randomly distributed to different task orders to control for sequencing effects. The selected units (see A) and their average difficulties (beta-values in terms of Item-Response Theory) were as follows:

1) Banana trade: easy level ($\beta = -0.95$) containing 95 words; realistic image in geography.

2) Legs of insects: easy level ($\beta = -0.75$) containing 59 words; realistic image in biology.

3) Auditory: medium level ($\beta = 0.10$) containing 122 words; graph in biology.

4) Pregnancy: medium level ($\beta = 0.39$) containing 136 words; realistic image in biology.

5) Map of Europe: difficult level ($\beta = 1.37$) containing 136 words; map in geography.

6) Savannah: difficult level ($\beta = 1.47$) containing 170 words; graph in geography.

2.1.3 Stimulating specific strategies

Learning materials were designed to study whether text processing differs from picture processing regarding task orientation. Students from secondary school were instructed to read materials combined with text and picture in biology and geography (see Figure 2.1)[1]. The materials were designed with incongruent information in text and picture to avoid the difficulty in processing of duplicated information (Heckler and Childers, 1992). For instance, the readers are required to process the information from text AND picture to answer the question. The questions are instructed either before or after the material. This might have an effect on how much participants look at text and at picture and how much fluctuate between text and picture.

Each participant was instructed to process six text-picture units under different conditions in order to gather eye tracking indicators of text vs. picture processing under three different processing conditions. Three units were presented without any information about the question to be solved afterwards. This was expected to stimulate a goal-free processing (see panel a in Figure 2.2). After this first phase of processing, the task appeared on the screen and participants were asked to solve the task. Participants could then re-read the text and re-observe the picture under the guidance of the task. This second phase of processing was expected to stimulate a delayed goal-oriented strategy. The other three units were presented when the participants had the

[1] BiTe Project is short for Bild-Text-Integration (translated as Picture-Text-Integration), which investigates the integrative ability of pupils with text and picture and how this ability is developed and can be encouraged since 2008.

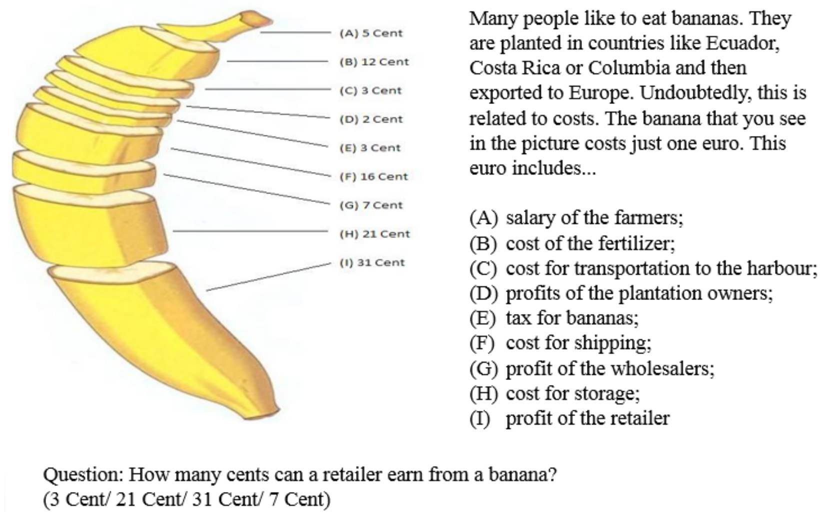


Figure 2.1: An example of infographics Banana Trade in this research (created by BITE Project).

task to be solved already in mind. Participants read the question first. After participants had read the task, the corresponding text and pictures appeared on the screen. Therefore, participants could explore the text and the pictures under the guidance of the question from the very beginning on. This was expected to stimulate an initial goal-oriented processing (see panel b in Figure 2.2).

Delayed goal-oriented and initial goal oriented processing were with a higher proportion of goal-oriented processing (see Figure 2.3). Hence, this study will focus on comparisons between (1) goal-free processing vs. delayed goal-oriented processing; (2) goal-free processing vs. initial goal-oriented processing.

2.1.4 Procedure

Experiments were conducted (each taking about 45 min) individually in a lab environment with the permission of students' parents. Materials and instructions were in German (native language of the research participants). After being informed about the purpose of the study, participants took the paper-pencil IQ tests and watched an instruction video. Participants were seated at 60-65 cm distance from the 24-inch monitor of the eye tracker positioned vertically at the eye level. A 5-point calibra-

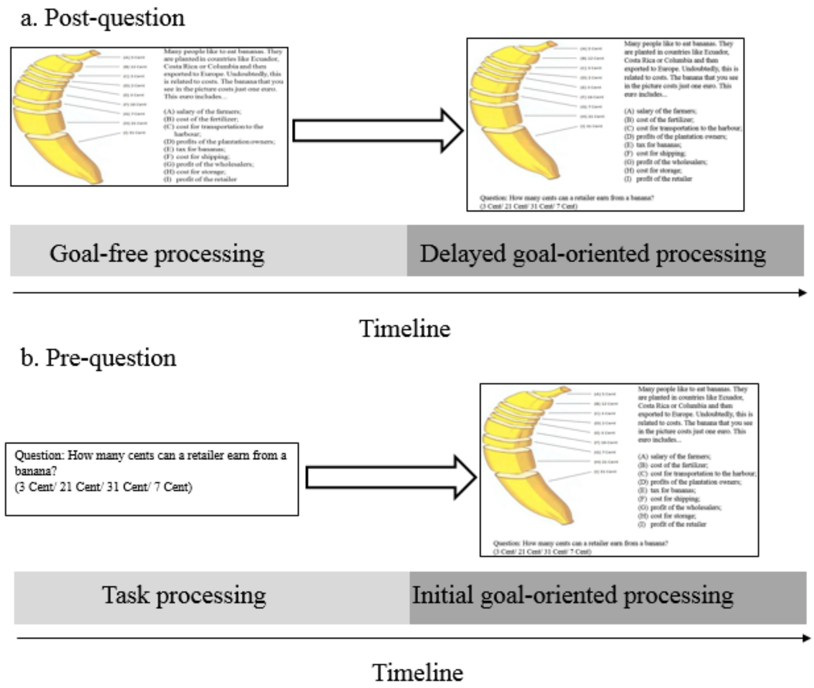


Figure 2.2: Design overview for the three reading processing. The timeline indicates the time that text, picture and question appeared on the screen. Panel a illustrates the post-question condition, in which question is posed after the material. Panel b displays the pre-question condition, in which question is posed before the material.

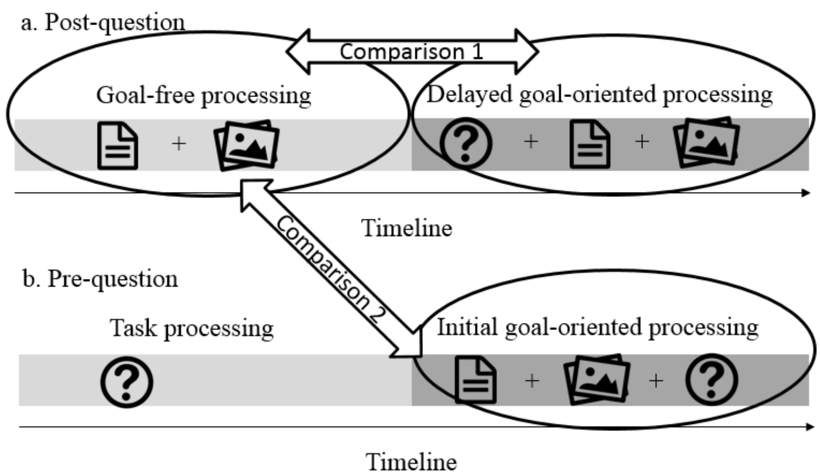


Figure 2.3: Comparisons between goal-free and goal-oriented processing in this study. The timeline indicates the time that text, picture and question appeared on the screen. The icons under delayed goal-oriented processing represent: question, text and picture.

tion was conducted before participants read the material. Once the calibration was successful, the experiment would start. The experiment included a warm-up phase and a main test. The aim of the warm-up phase was to give participants the opportunity to get familiar with the eye tracking system and with using keyboard and mouse for turning pages and answering the questions. A Tobii XL60 eye tracker was used to record eye movements at the rate of 60 Hz. The system can compensate for head movements and thus provides relatively precise data from young participants and a comfortable testing situation. For each unit, the strategy was indicated by a written instruction presented upfront. Finally, participants were thanked and rewarded with €12 for taking part in our experiment.

2.1.5 Scoring

The AOIs were drawn manually using Tobii Studio software. Each task had three separated AOIs: the picture, the text and the question. The average picture AOI for the six materials covered 25.14% of the screen; and the average text AOI covered 27.72%; the average question AOI occupied 18.55%. Participants obtained one point when they answered a question correctly. They could obtain a maximum score of six points (i.e., accuracy rate of 100%) and a minimum score of zero points.

2.1.6 Eye tracking method

Past research has established the basic assumption that cognitive process is mirrored by eye movement indicators, (e.g. Gaschler, Marewski, and Frensch, 2015; Godau, Wirth, Hansen, Haider, and Gaschler, 2014). The positions where a person's eyes look at are related to cognitive processing according to the eye-mind assumption and the immediacy assumption (Just and Carpenter, 1980). The eye-mind hypothesis assumes that eye movements can reflect information processing. The immediacy hypothesis states that the processing of information is immediate and happens directly after it is perceived. Although the quality of the eye tracking is disputable (Wass, Forssman, and Leppänen, 2014), it can provide the spatial and temporal patterns of attention (van Gog and Scheiter, 2010). By providing the position (word, picture), the duration (how long) and the sequence (firstly look at where and later look at where), eye tracking enlightens the cognitive processes during reading. Fixation counts and fixation duration, i.e. accumulated fixation counts and the fixation time at a particular Area of Interest (AOI) are associated with the depth of cognitive processing and spatial distribution of attention (Hegarty, 1992; Rayner, 1998). Time to the first fixation of an AOI (i.e. latency until an AOI is fixated for the first time) indicates readers' interest in the AOI (J. H. Goldberg and Kotval, 1999). Transitions between AOIs (i.e., frequency of the eye movements from one AOI to the other) can mark the integration of contents presented in the AOIs (Johnson and Mayer, 2012). Eye movements is suggested to be associated with learning outcomes (Mason, Tornatora, and Pluchino, 2013). Higher integrators of text and picture tended to outperform the lower integrators.

2.2 Results

Participants had an average of 59% accurate answers to the questions (SD = 25%) with goal-free processing combined with delayed goal-oriented processing and an average of 53% accurate answers (SD = 36%) with initial goal-oriented strategy, $t(16) = 0.65$, $p = .52$, $d = 0.19$. The average time for comprehending text and pictures and answering the question was 1.14 seconds per word (SD = 0.39 second per word) from a range of 1.71 seconds per word to 0.61 seconds per word.

Most importantly, the different experimental conditions of processing strategies differed in eye fixations on text vs. picture. Considering the variation of reading speed between participants and the limited number of participants, I decided to use proportion measures for capturing the relative weight of text vs. picture (i.e. proportion of fixation counts on text, proportion of fixation time on text, proportion of visit counts on text and proportion of visit duration on text (Holmqvist et al., 2011). For example, the proportion of fixation counts on text was calculated by dividing this count by the sum of fixation counts on text and picture. Due to the dependency of text and picture, data for text were only shown with these indicators to avoid redundancy. Thus, if I refer to high proportion of fixations on text, this implies a low proportion on the picture and vice versa.

In order to explore whether the usage of text differs from the usage of picture with different strategies, we conducted two one-way repeated-measures multivariate analysis of variance (MANOVA) with four eye-tracking indicators in three reading conditions. As there were two types of task-driven strategies, two MANOVAs were performed: (1) goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material), and (2) goal-free processing vs. initial goal-oriented processing (question before material). The eye-tracking indicators included the proportion of fixation counts and fixation duration on text, the proportion of number of visit and visit duration on text, the time to the first fixation on text and picture and the number of transitions between text and picture.

2.2.1 Goal-free processing vs. delayed goal-oriented processing

In order to check whether the goal-free processing (no question yet) differs from the delayed goal-oriented processing (question after material) in fixations on text vs. picture, we performed a MANOVA and univariate analyses (ANOVAs) with the eye tracking indicators listed in Table 2.1. The goal-oriented processing and delayed goal-oriented processing differed significantly with respect to the relative weight on text (rather than picture) across the eye tracking indicators, $F(7,26) = 14.10$, $p < .001$, ${\eta_p}^2 = .79$. As reported below, separate ANOVAs confirmed the difference between the two strategies for indicators like fixation counts and fixation duration, visit counts and visit duration and time to the first fixation.

(1) Fixation indicators during text-picture comprehension

The ANOVA revealed that the proportion of fixation counts on texts (rather than on picture) was significantly higher with goal-oriented processing than with delayed goal-oriented, $F(1,32) = 56.51$, $p < .001$, ${\eta_p}^2 = .64$. Proportion of fixation counts on texts and proportion of accumulated fixation duration on texts were correlated by $r = .97$. Thus, unsurprisingly the proportion of accumulated fixation duration on text was higher with goal-oriented processing than with delayed goal-oriented processing, $F(1,32) = 70.41$, $p < .001$, ${\eta_p}^2 = .69$.

(2) Visit indicators during text-picture comprehension

Participants visited the text AOI (rather than the picture AOI) more often with the goal-free processing than with the delayed goal-oriented processing. The proportion of number of visits on text was higher when participants engaged in initial coherence-formation processing rather than in delayed goal-oriented processing, $F(1,32) = 7.93$, $p = .008$, ${\eta_p}^2 = .20$. Participants had a higher proportion of accumulated visit duration on text with goal-free processing than with delayed task-oriented processing, $F(1,32) = 14.24$, $p = .001$, ${\eta_p}^2 = .31$.

(3) Time to first fixation on text and picture

Participants fixated within a shorter time on texts with the initial coherence-formation strategy than with the delayed goal-oriented strategy, $F(1,32) = 9.43$, $p = .004$, ${\eta_p}^2 = .23$. They also fixated more quickly on pictures with the goal-free processing than with the delayed goal-oriented processing, $F(1,32) = 4.96$, $p = .033$, ${\eta_p}^2 = .13$. Pictures were fixated slightly quicker than texts: $F(1,32) = 0.87$, $p = .358$, ${\eta_p}^2 = .03$ for the goal-oriented processing; $F(1,32) < 1$, for the delayed goal-oriented processing.

(4) Transitions between text and picture

The transitions between text and picture did not show a robust difference between goal-oriented processing and delayed goal-oriented processing, $F(1,32) = 2.34$,

p = .136, $\eta_p{}^2$ = .07. With the delayed goal-oriented processing, participants had 27% of transitions (SD = 10.6%) between text and picture, 25% of transitions (SD = 10.4%) between text and question and 48% of transitions (SD = 13.7%) between picture and question. In short, participants transferred their eyes slightly more often between text and picture with goal-free processing than with delayed goal-oriented processing. When questions were illustrated, participants mainly transferred their attention between picture and question.

Table 2.1: Mean and standard deviations of eye tracking indicators in different goal-oriented strategies.

Eye tracking indicators	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
% Fixation counts on text	80.11 (12.81)	47.87 (12.19)	64.95 (16.4)
% Fixation time on text	81.27 (12.99)	44.00 (12.73)	66.93(15.72)
% Visit counts on text	48.29 (9.13)	39.64 (8.79)	46.16 (8.47)
% Visit time on text	66.72 (20.27)	44.88 (12.61)	69.31 (11.91)
Time to first fixation on text (sec)	2.11 (1.44)	6.60 (5.85)	2.21 (1.92)
Time to first fixation on picture (sec)	1.57 (1.92)	5.02 (6.09)	0.87 (2.07)
Average number of transitions between text and picture	21.59 (13.4)	14.94 (14.47)	8.53 (16.18)

2.2.2 Goal-free processing vs. initial goal-oriented processing

Eye tracking indicators also showed that the goal-free processing (i.e. general processing without task) differed from the initial goal-oriented strategy (i.e. task was presented at the beginning) in terms of text and picture processing. The MANOVA showed the significant effect of goal-orientation on eye tracking indicators, $F(7,26) =$ 3.49, p = .009, $\eta_p{}^2$ = .48. Separate ANOVAs yielded effects of strategy condition on fixation indicators but not on visit indicators.

(1) Fixation indicators during text-picture comprehension

There was a significantly higher proportion of fixation counts on text (rather than on picture) with the goal-free processing than with the initial goal-oriented processing, $F(1,32) = 9.03$, $p = .005$, ${\eta_p}^2 = .22$. As proportion of counts and of duration of fixations were highly correlated ($r = .84$), this was mirrored by a similar effect on proportion of fixation duration on text, $F(1,32) = 8.41$, $p = .007$, ${\eta_p}^2 = .21$.

(2) Visit indicators during text-picture comprehension

Participants had a similar pattern of results in visiting text and picture for both experimental strategy conditions. No difference was detected for the proportion of visit counts and visit duration on text between the goal-free processing and the initial goal-oriented processing, $Fs(1,32) < 1$.

(3) Time to first fixation on text and picture

We did not find any difference between the experimental strategies for latency of first fixation on text or picture, $Fs(1,32) < 1.03$. Participants fixated the picture marginally sooner than the text with initial task-oriented processing, $F(1,32) = 3.79$, $p = .06$, ${\eta_p}^2 = .11$.

(4) Transitions between text and picture

Transitions between text and picture did not differ when participants followed the goal-free processing vs. the initial goal-oriented processing, $F(1,32) = 1.48$, $p = .23$, ${\eta_p}^2 = .04$. For the initial goal-oriented processing, participants had 45% (SD = 13%) of transitions between text and picture, 20% (SD = 7.6%) between text and question and 35% (SD = 17.5%) between picture and question. In brief, participants transferred their eyes dominantly between text and picture, secondarily between picture and question and lastly between text and question.

2.3 Discussion

The pilot study provides first methodological tools and results to specify the distinction between general and specific processing proposed by Rickards and Denner (1978) for processing of text vs. picture in mixed material. General processing (i.e., when the question is not known yet) deals with the global thematic coherence of the material. Pre-posed questions can lead to a highly selective processing. In order to apply this account to the processing of mixed material (text and picture), the current study examined whether text processing differs from picture processing and whether this difference is moderated by the strategies used by the learner. Specifically, this study compared (1) goal-free processing (no question yet) with delayed goal-oriented strategy (question after material) and (2) goal-free processing with initial goal-oriented processing (question before material). A higher emphasis on text relative to picture was expected for the goal-free processing. Pictures were assumed to lead to higher values in fixation indicators with the delayed goal-oriented processing.

Eye tracking indicators were analyzed to reveal the processing of text and picture, according to the eye-mind theory and the immediacy theory. Our results confirmed general differences of text vs. picture processing (McNamara, 2007). Importantly, eye tracking indicators of relative emphasis on text (rather than on picture) differed among the experimentally induced processing strategies. For the goal-free processing, participants primarily fixated on text rather than on picture. As fixation count and fixation duration have been linked to the depth of cognitive processing and distribution of attention (Rayner, 1998), the result suggest the primary usage of text during mental model construction in initial coherence-formation processing. The same is true for visit counts and visit duration. As visit indicators have been linked to the importance and information value of the AOIs (Jacob and Karn, 2003), participants possibly consider text to be important and informative with goal-free processing. Participants also needed little time to proceed to text and picture (i.e. time to the first fixation) with goal-free processing. This indicator has been linked to participants' interest on text and picture when reading is general (J. H. Goldberg and Kotval, 1999). In addition, participants had frequent transitions between text and picture with goal-free processing. According to the Integrative Model of TPC, participants may establish their mental model by integrating text and picture. This is consistent with our assumption that participants intensely processed the content to establish an initial mental model with goal-free processing.

For the delayed goal-oriented processing, picture was mainly used to scaffold question solving after the initial construction of the mental model. The results from fixation counts and fixation duration were consistent with the idea that pictures are primarily used when participants need to answer the question with delayed goal-oriented processing (cf. Hochpöchler et al., 2013). They had a high amount of visit counts and visit duration on pictures with delayed goal-oriented processing. It seems that participants considered pictures important and informative when they were asked to solve tasks after the initial construction of the mental model. Besides, data also revealed that participants perceived pictures sooner than texts with both processing strategies. This can be explained by pictures attracting readers' attention (Mayer, 1989; Tversky, 2001; Winn, 1989). Transition data showed that they focused their main attention on picture and question, less attention on text and picture and the least on text and question when they followed delayed goal-oriented processing. Participants seemed to primarily use the picture to scaffold question answering. They paid less attention on text and picture because they have already constructed the initial mental model and they just need to further construct or update this mental model in order to answer the question. They paid the least attention on text and question, which implies that the text also helps participants to answer the question. Pictures possibly serve as a tool for question solving with delayed goal-oriented processing and text may be used for building and updating the mental model. These results correspond to the assumption of unequal usage of text and pictures in the Integrative Model of TPC. According to this model, pictures are mainly processed as an external representation to solve questions.

With the initial goal-oriented processing, participants invested a large amount of time on pictures because participants may have used pictures as an external tool to scaffold question answering. They visited more frequently and spent longer time on text than on picture. This might be explained by coherence-formation being supported by the text. In our design, participants need to process both text and picture to get the correct answer. Although participants were instructed with questions in initial task-oriented processing, they still needed to understand the text and the picture, thus also constructing a mental model. Therefore, text and pictures were both used with initial task-oriented processing. Also, time to the first fixation suggested that text and pictures drew participants' interest with initial task-oriented processing. Learners might integrate text and picture to build the initial mental model, as assumed in the Integrative Model of TPC. Similar patterns were shown for transitions between text and pictures with initial goal-oriented processing. Participants transferred their attention most frequently between text and picture, less between picture and question and the least between text and question with the initial goal-oriented processing. On the one hand, participants integrated text and picture with initial goal-oriented processing. On the other hand, more transitions between picture and question than between text and question support our assumption that picture is primarily used to solve the question. This result also corresponds to the assumption that text is more likely to be used for general coherence-formation processing and picture is especially used for specific task-oriented processing.

Summarizing the results, the pilot study suggested that text processing and picture processing differ substantially when learners are exposed to text-and-picture material. The differences are moderated by processing strategies triggered by context factors such as presentation of the question prior to vs. after the first exposure to the text-and-picture material. Likely, text is primarily used for coherence-formation processing. It assists learners to comprehend the content of the materials, which generate an initial mental model or coherent semantic representation. Picture is likely used for task-oriented specific processing. When learners have constructed the initial mental model, picture is mainly used as an external representation to update the mental model and to answer the question. When learners have the tasks beforehand, picture might serve mainly to scaffold the initial mental model construction. Future studies should provide more evidence for the link between (a) differences in fixation patterns elicited by different processing strategies and (b) the formation and usage of mental models integrating text and picture information.

In conclusion, returning to the hypotheses posed at the beginning of this study, it is now possible to state that text processing differs fundamentally from picture processing and that this difference is moderated by different reading strategies. More specifically, this study suggested that text is mainly used to build the mental model with coherence-formation general strategy. Comparatively, picture is more likely to guide readers to solve questions with task-oriented selective strategy. Similar to the previous studies on text comprehension, reading strategies also influence the comprehension of text and picture. Our findings expand the Rickards and Denner, 1978) account of global vs. selective processing to the domain of mixed (picture and text) materials. The results suggest that eye tracking indicators can play a major role in assessing and scaffolding text and picture comprehension. Eye tracking indicators might be used to assess whether the learner is following an approach suitable to the current context factors (i.e. presentation of the question prior to vs. after the first exposure to the text-and-picture material). Based on such assessment, interventions guiding visual attention to areas relevant at the current processing stage can involve salient visual cues presented online during task processing (cf. Rouinfar, Agra, Larson, Rebello, and Loschky, 2014).

Chapter 3

Experiment 1

This experiment is based on the published pilot study (Zhao et al., 2014). It is supposed to deepen the understanding of the effective factors in multimedia learning. Based on the study in text processing, this experiment is mainly concerned with the influence of task orientation on learning with text and pictures. It also explores whether the reader prefers text or pictures when questions are of different difficulty levels. Furthermore, it investigates whether the readers from higher school tiers differ from readers from lower school tiers in text processing and picture processing.

According to the hypotheses, it is assumed that a reader may pay a large amount of attention to text when s/he is required only to comprehend the materials. However, s/he may focus more frequently on pictures when s/he is required to answer questions. The reader is expected to focus more frequently on text with easy questions, whereas s/he may focus more frequently on pictures with difficult questions. Besides, students from higher school tier may process the text more efficiently and pictures more intensively than students from lower school tier.

The results are in line with Hypothesis 1 on task orientation. Readers emphasize on text to comprehend the materials; whereas, they fixate mainly on pictures to answer the questions. For question difficulty, it is found that readers spend much more time on text and on pictures with difficulty questions than with easy questions. However, the assumed pattern from Hypothesis 2 is not shown. The increase of attention on pictures is more than on text when question is difficult. Different from Hypothesis 3, no difference is shown in text processing between students from higher school tier and lower school tier. Conversely, students from higher school tier process the pictorial information more intensively than students from lower school tier.

3.1 Method

Based on the hypotheses, participants were recruited from higher school tier (Gymnasium) and lower school tier (Realschule) in 5^{th} grader. Their accuracy and eye movements were recorded during reading. Task orientation, question difficulty and school tier showed different degrees of influence on multimedia comprehension.

3.1.1 Participants from Grade 5

Seventy-seven 5^{th} graders in Germany participated in Experiment 1 based on power analysis to ensure the adequacy of the sample size. According to power analysis (Faul, Erdfelder, Buchner, and Lang, 2009), minimal 44 participants were needed for effect size $F = .25$, $\alpha = .05$, $\beta = .95$ for two groups and three measurements. Data from 72 participants ($M = 11.4$ years, $SD = 0.6$ years) were analysed because 5 participants had problems with calibration on the eye-tracker and were unmotivated in reading. Forty-four (61.1%) participants were male and twenty-eight (38.9%) were female. Thirty-six were students from Gymnasium (higher school tier) and thirty-six were from Realschule (lower school tier).

Intelligence tests for 5^{th} graders were used to check their spatial and verbal abilities (Heller and Perleth, 2000). Students from Gymnasium had better scores than average in spatial ability (average T of 52.97; $SD = 9.22$) and verbal ability (average T of 51.06; $SD = 5.88$). Students from Realschule had slightly worse scores than average in spatial ability (average T of 47.28; $SD = 9.34$) and verbal ability (average T of 46.86; $SD = 5.37$). They took part in the experiment because they were interested in the study and they could receive € 12 as reward.

3.1.2 Mixed design

Experiment 1 was a $2(\times 2 \times 2 \times 2)$ mixed design with *school tier* (Gymnasium vs. Realschule) as between subject variable; *task orientation* (post-question vs. pre-question), *question difficulty* (level 1 vs. level 2/3), and *multimedia representation* (text vs. pictures) as within-subjects variables. Students from Gymnasium and Realschule were separated into two groups. Students from each group were randomly distributed to learn six tasks. Dependent variables were accuracy of solving the questions and eye tracking indicators. Four eye tracking indicators were reported, including number of fixation, accumulated fixation duration, time to the first fixation and transitions between Areas of Interest (AOIs).

3.1.3 Eye tracking apparatus

A Tobii XL60 eye tracker was used to build the experiment and to record eye movements at the rate of 60 Hz. The materials were displayed with the 24-ich eye tracker. The system can compensate for head movements and thus provides relatively precise data from young participants and a comfortable testing situation. A second monitor was connected to the eye tracker in order to check the synchronised eye movements and the position of eyes during the experiment. When the eye movements were not displayed on this monitor, the experimenter would promptly modify participants' position or the height of the eye tracker. Later, the recorded eye movements were exported from Tobii Studio 2.3.1, e.g. accumulated fixation duration, time to the first fixation. However, eye tracking indicators like transitions between AOIs cannot be exported directly from Tobii Studio. Thus, R Core Team (2015) was used to calculate the transitions between AOIs. All the data were analysed combined with R and SPSS based on the needs.

3.1.4 Materials from science textbooks

Six text-picture units out of sixty were chosen in the experiment for the sake of eye tracking method. The materials with combined text and pictures were the same as in the pilot study. It aims at assessing students' skills in integrating text and pictures in a large scale. The information in the text and the pictures was incongruent to avoid difficulty of processing (Heckler and Childers, 1992). Differential Item Functioning (DIF) analysis and rational task analysis on questions were performed to ensure the reliability of the units and questions (Schnotz et al., 2011). In Experiment 1, each participant needed to comprehend 6 text-picture materials (see in Appendix A). The materials varied in difficulty (beta-values in terms of DIF analysis), text length, image type (realistic pictures vs. graphs) and subject (geography vs. biology).

Each text-picture material had three problem-solving questions. However, only one out of three questions was selected for each text-picture material (see Table 3.1). Based on Wainer (1992)'s hierarchies of structure mapping, questions were designed to measure reader's understanding of the presented material. Level 1 question dealt with processing of elements. It had an average beta value of -1.15 on the Rasch scale. Level 2 and level 3 questions required a reader to process a relationship between elements in the material. The relationship was indicated but not directly displayed in the material. They had an average beta value of +0.57 and +1.39, respectively. Besides, participants needed to comprehend both text and pictures to answer questions correctly. Level 1 had 1.25 text-picture mapping in terms of rational task analysis; level 2 had 1.5; level 3 had 3.00.

Each participant was asked to learn three consecutive text-picture materials in post-question condition and three consecutive text-picture units in pre-question condition. Text, pictures and question were relatively close together to avoid cognitive strain (Rickards and Di Vesta, 1974). They were also not overlapped to avoid data redundancy and unnecessary interpretation difficulty in eye tracking (Holmqvist et al.,

Table 3.1: Each material has one question. Number 1 behind the materials refers to an easy question; 2 refers to a medium question; 3 means a difficult question. Materials (M) in dark grey areas are in post-question position and materials in light grey areas are in pre-question position.

Seq	M 1	M 2	M 3	M 4	M 5	M 6
1	Insect1	Pregnancy3	Europe2	Banana1	Auditory3	Savannah2
2	Europe3	Insect2	Fetus1	Savannah3	Banana2	Auditory1
3	Pregnancy2	Europe1	Insect3	Auditory2	Savannah1	Banana3
4	Banana1	Auditory3	Savannah2	Insect1	Pregnancy3	Europe2
5	Savannah3	Banana2	Auditory1	Europe3	Insect2	Pregnancy1
6	Auditory2	Savannah1	Banana3	Pregnancy2	Europe1	Insect3
7	Insect1	Pregnancy3	Europe2	Banana1	Auditory3	Savannah2
8	Europe3	Insect2	Fetus1	Savannah3	Banana2	Auditory1
9	Pregnancy2	Europe1	Insect3	Auditory2	Savannah1	Banana3
10	Banana1	Auditory3	Savannah2	Insect1	Pregnancy3	Europe2
11	Savannah3	Banana2	Auditory1	Europe3	Insect2	Pregnancy1
12	Auditory2	Savannah1	Banana3	Pregnancy2	Europe1	Insect3

2011). The materials were presented in 12 sequences (see Table 3.1) to counterbalance the question difficulty and reading strategy. In post-question condition, three materials were presented without any information about the question to be solved afterwards (see panel a in Figure 3.1). It included two phases of processing: goal-free processing and delayed goal-oriented processing. With goal-free processing, participants viewed only text and pictures. Then they followed delayed goal-oriented processing where they could re-read the text and re-observe the pictures under the guidance of the question. In pre-question condition, three units were displayed with the question to be solved at the beginning (panel b in Figure 3.1). It also contained two phases of processing, including task processing and initial goal-oriented processing. For task processing, participants viewed only the question so they have the question in mind before processing the units. For initial goal-oriented processing, participants could explore the text and the pictures under the guidance of the question from the very beginning on. Task processing will not be discussed due to the focus was on comprehending with text and pictures in this study.

Take Figure 2.1 as an example, a reader comprehended Banana Trade either in post-question or in pre-question condition. Placing questions before or after the ma-

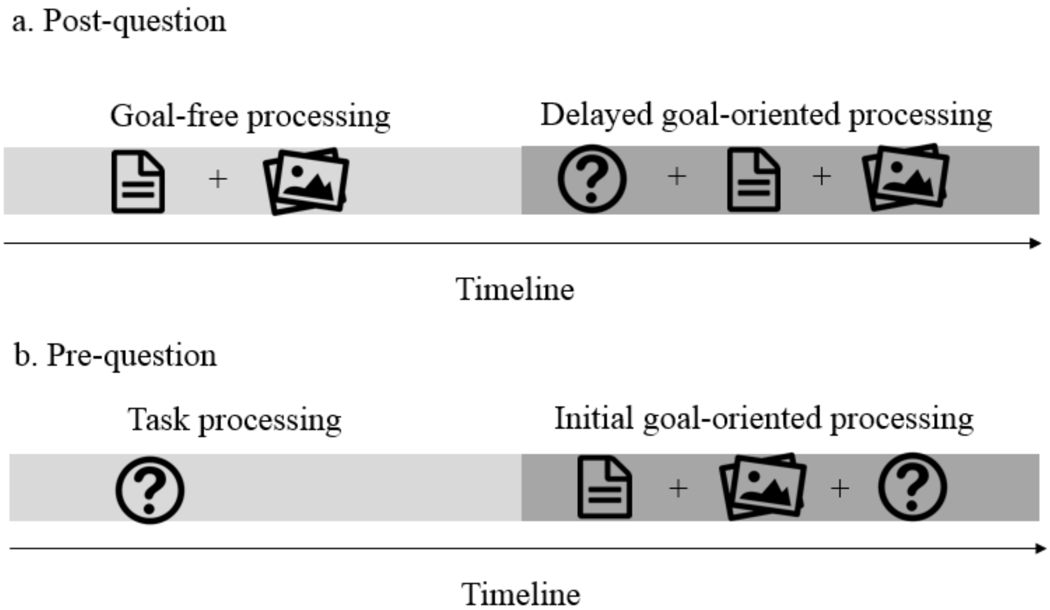


Figure 3.1: Two sequences of displaying text, pictures and question on the screen. Panel a is post-question condition, in which question is displayed after the text and pictures. Panel b is pre-question condition, in which question is shown prior to the text and pictures. The icons in delayed goal-oriented processing represent question, text and pictures.

terials can constitute different reading situations. The reading situations can guide readers to establish a strategy or processing to comprehend the materials.

In the post-question condition, text and pictures were displayed first and question was displayed later (see panel a in Figure 2.2). It triggers (1) goal-free processing in which the reader was required to read the material without the question. (2) Delayed goal-oriented processing in which the reader was required to answer the question after exploring the material. In pre-question condition, question was shown first and text and pictures were shown later (see panel b in Figure 2.2). It yielded also two reading strategies: (3) task processing in which the reader was required to read the question. (4) Initial goal-oriented processing in which the reader was required to search for the answer in the material after having explored the question. Due to the interest in text and picture comprehension, the analysis on (3) task processing was omitted in this study.

The reader was permitted to answer one of the three problem-solving questions for each text-picture material. The difficulty of three questions varied based on Wainer's hierarchies of structure mapping (Wainer and Thissen, 1981; Wainer, 2005). Level 1 was easy and required processing of facts. For the illustrated example (see Figure 2.1, level 1 question was "How many cents can a retailer earn from a banana?" The reader needed to identify "(I) profits of the retailers" in the text and map it to the picture to obtain the correct answer "31 Cent". Level 2 was medium and required processing of relations. Level 2 question was "For whom do people pay the least amount of money if they buy a banana?" The reader needed to compare the information from the picture "(D) 2 Cent" and map it to the text "(D) profits of the plantation owners". Level 3 was difficult and required processing between relations. Level 3 question was "If we compare the farmers, retailers and wholesalers, then who earns the most, who earns less and who earns the least from a banana". The reader needed to extract the information of "farmers", "retailers" and "wholesalers" from the text and the picture and compare their relations. Then they had the correct answer "Retailers earn the most, wholesalers earn less and farmers earn the least from a banana".

3.1.5 Procedure

The experiment was conducted (each data acquisition taking about 45 min) individually in a lab environment with the permission of students' parents. Materials and instructions were in German, which was the participants' native language. After being informed about the purpose of the study, participants took the paper-pencil spatial ability and verbal ability tests and watched an instruction video. Afterwards, they were seated at 60-65 cm distance from the 24-inch monitor of the eye tracker positioned vertically at the eye level. Modified by pilot study, a 9-point calibration was conducted before participants read the material. The 9-pint calibration was supposed to improve the quality of recording the eye movements. Once the calibration was successful, the eye tracking experiment would start.

Participants were randomly assigned to one of the 12 sequences (see Table 3.1). The experiment included a warm-up phase and a main test. The aim of the warm-up phase was to give participants the opportunity to get familiar with the eye tracking system and with using keyboard and mouse for turning pages and answering the questions. Once the page was turned, it was impossible to turn back. Depending on the reading strategy, a question from three difficulty levels was posed to the participants either before or after the initial processing of text and pictures. For each material, the post- and pre-question conditions were indicated by a written instruction presented upfront. Finally, participants were thanked and rewarded with € 12 for taking part in the experiment.

3.1.6 Scoring

Participants obtained one point when they answered a question correctly. They could obtain a maximum score of six points (i.e., accuracy rate of 100%) and a minimum score of zero points. The AOIs were drawn manually using Tobii Studio software. Each material had three separated AOIs: the picture, the text and the question. The text from six materials covered on average 27.72% of the screen; text covered 25.14%; questions occupied 18.55%.

3.2 Results

According to the theories outlined in the introduction, text is expected to guide general reading; whereas pictures may guide selective reading. When the question difficulty increases, readers' attention changes from text to pictures. Readers from higher tier (Gymnasium) are assumed to emphasize less on text but more on pictures than readers from lower tier (Realschule). Data on accuracy, eye movements and correlation between them are reported in three perspects: task orientation, question difficulty and school tier. The accuracy data present the correct answers from 5th graders from higher school tier (Gymnasium) and lower school tier (Realschule). The data from eye movements contain two comparisons because two processing involved higher amount of goal-oriented processing. (1) Goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material). (2) Goal-free processing vs. initial goal-oriented processing (question before material). Finally, correlations between accuracy and eye movements report how well the eye movement data can predict the accuracy. In other words, it explores whether higher values in the eye tracking indicators indeed require for integration of text and graphics.

The results from fixation patterns suggested that text processing differed fundamentally from picture processing with different reading strategies among 5th graders. Consistent with Hypothesis 1, text was mainly emphasised in goal-free processing (no question yet). Pictures were primarily focused in delayed goal-oriented processing (question after material). Participants fixated text and pictures rather balanced in initial goal-oriented processing although text was mainly fixated. Differed from Hypothesis 2, the attention on text and on pictures raised as question difficulty increased. The increase of attention on pictures was more than on text with difficult questions. Inconsistent with Hypothesis 3, higher tier students had similar fixation patterns as lower tier students in text processing. Conversely, higher tier students fixated on pictures longer than lower tier students. In addition, the correlation data revealed that eye tracking indicators indeed suggested the integration of text and pictures.

3.2.1 Question 1: How do readers deal with text and deal with pictures in different forms of reading?

Participants are presumed to have similar proportion of accuracy with post-question condition (reading followed by question) and pre-question condition (reading guided by question). However, they may focus their attention mainly on text when only text and pictures are displayed. They may pay frequent attention to pictures when they are asked to answer questions. The results on task orientation are in line with the hypothesis. Participants had similar accuracy with post-question and pre-question conditions. They focused more frequently on text when they comprehended the material without knowing the question. Conversely, they fixated more frequently on pictures when they required to answer the question.

The measure of accuracy was the average scores for answering six displayed materials. Participants had similar scores in comprehending the materials with post-question condition (M = 45%, SD = 31%) and with pre-question condition (M = 45%, SD = 29%), $F < 1$, ns. However, participants fixated on text and pictures differently with different reading strategies. Four eye tracking indicators were analysed to answer the research questions, i.e. accumulated number of fixation, accumulated fixation duration, time to the first fixation and number of transitions. As number of fixation was highly correlated with accumulated fixation duration, r = .93, only accumulated fixation duration was reported to avoid redundancy.

Two repeated-measures analysis of variance (ANOVAs) were conducted to examine whether the usage of text differed from pictures in different reading processing. As two processing contained a higher proportion of goal-oriented processing, two comparisons were conducted between: (1) goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material); (2) goal-free processing vs. initial goal-oriented processing (question before material).

3.2.1.1 Goal-free processing vs delayed goal-oriented processing

Eye tracking indicators were analyzed based on the eye-mind hypothesis and immediacy hypothesis (Just and Carpenter, 1980). As illustrated in Table 3.2, separate ANOVAs revealed robust differences on fixation patterns between goal-free processing (no question yet) and with delayed goal-oriented processing (question after material). Participants fixated considerably longer on text but significantly shorter on pictures with goal-free processing than with delayed goal-oriented processing. They also viewed text quicker but pictures slower with goal-free processing than with delayed goal-oriented processing. They transferred their attention more frequently between text and pictures with goal-free processing. Conversely, they transferred their attention most frequently between pictures and questions with delayed goal-oriented processing.

Table 3.2: Acumulated fixation duration, time to the first fixation on text and pictures and transitions between text, pictures and question with three reading processing among 5th graders.

Grade 5	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Text	56.64 (20.77)	8.80 (6.69)	49.87 (27.16)
Picture	10.08 (5.65)	14.41 (11.88)	19.76 (13.60)
	b. Time to first fixation (sec)		
Text	2.75 (1.87)	7.38 (7.60)	4.08 (4.49)
Picture	8.61 (14.74)	5.62 (5.21)	0.95 (3.53)
	c. Number of fixation		
Text-Picture	26.59 (11.87)	11.87 (10.06)	26.49 (17.45)
Text-Question	NA	10.75 (8.00)	12.41 (9.10)
Picture-Question	NA	24.66 (15.50)	23.85 (13.90)

(1) Accumulated fixation duration during text-picture integration

The ANOVA showed that participants fixated significantly longer on text with goal-free processing than with delayed goal-oriented processing (see Figure 3.2), $F(1,71) = 341.85$, $p < .001$, ${\eta_p}^2 = .83$. In contrast, participants fixed significantly shorter on pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 7.20$, $p = .009$, ${\eta_p}^2 = .09$. With goal-free processing, participants fixated much longer on text than on pictures, $F(1,71) = 323.32$, $p < .001$, ${\eta_p}^2 = .82$. With delayed goal-oriented processing, they fixated much shorter on text than on pictures, $F(1,71) = 20.16$, $p < .001$, ${\eta_p}^2 = .22$.

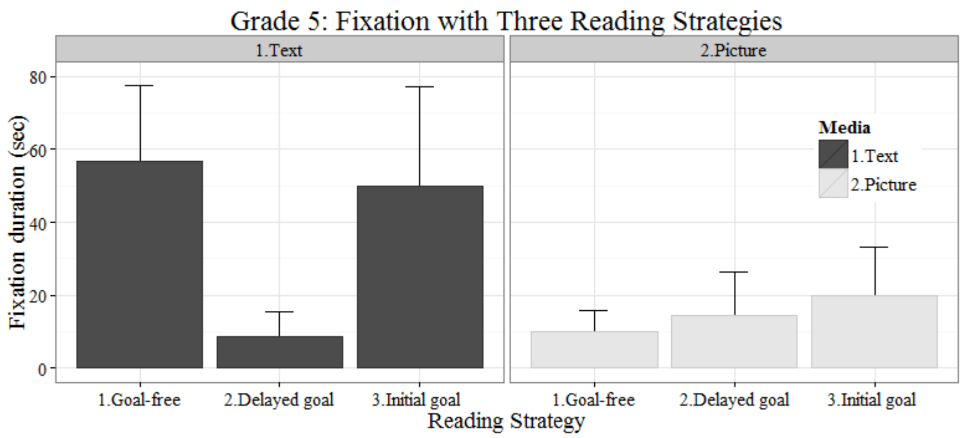


Figure 3.2: Accumulated fixation duration on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders. Error bars indicate between subjects standard error of the mean.

(2) Time to the first fixation during text-picture integration

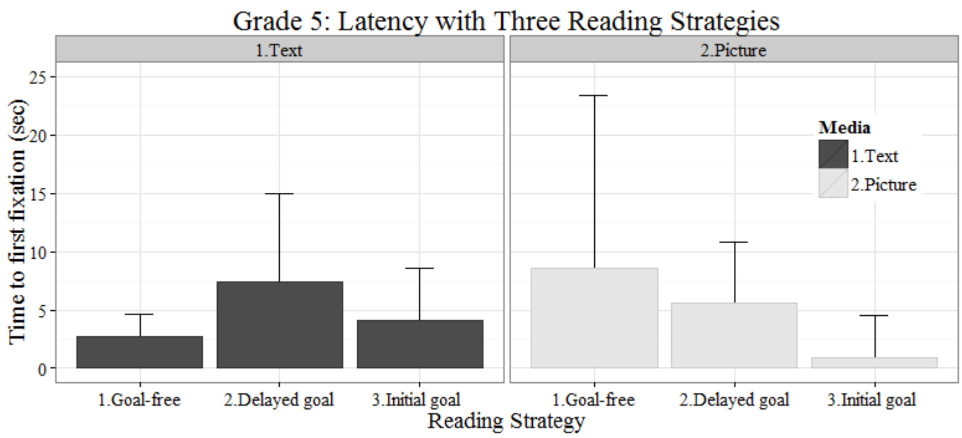


Figure 3.3: Time to the first fixation on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders. Error bars indicate between subjects standard error of the mean.

Participants perceived text considerably sooner with goal-free processing than with delayed goal-oriented processing (see Figure 3.3), $F(1,71) = 25.33$, $p < .001$, ${\eta_p}^2 = .26$. They fixated marginally slower on pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 3.34$, $p = .07$, ${\eta_p}^2 = .05$. Text was fixated much quicker than pictures with goal-free processing, $F(1,71) = 10.99$, $p = .001$, ${\eta_p}^2 = .13$. In contrast, pictures were fixated slightly quicker than text with delayed goal-oriented processing, $F(1,71) = 2.45$, $p = .12$, ${\eta_p}^2 = .03$.

(3) Number of transitions during text-picture integration

Participants transferred their attention frequently between text and pictures with goal-free processing (M = 26.59, SD = 11.87) (see Figure 3.4). They had M = 11.87 (SD = 10.06) transitions with delayed goal-oriented processing (M = 11.87, SD = 10.06). With goal-free processing, participants had 43.7% (SD = 7.4%) transitions from text to pictures and 55.3% (SD = 7.4%) transitions from pictures to text. With delayed goal-oriented processing, they had 23.1% (SD = 11.1%) transitions between text to pictures, 24.0% (SD = 13.6%) transitions between text and question, 53.0% (SD = 14.8%) transitions between pictures and question.

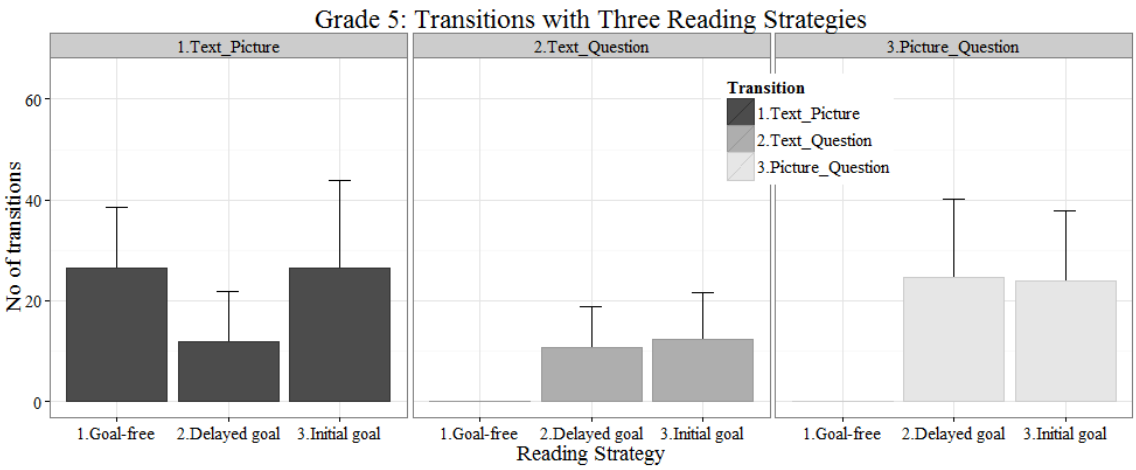
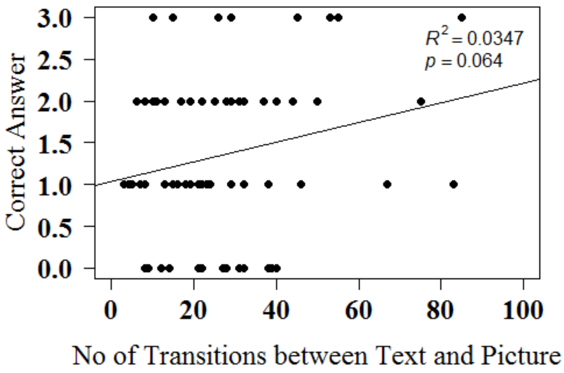
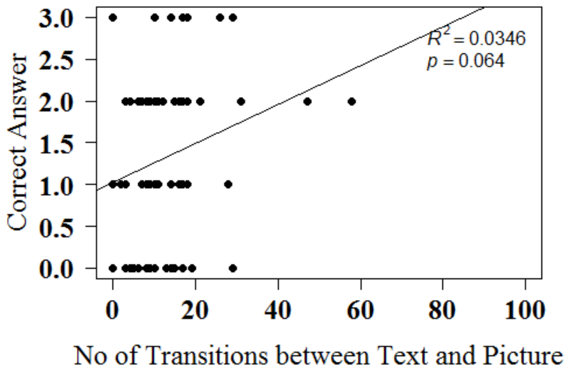


Figure 3.4: Transitions among text, pictures and question in goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders. No transitions are illustrated between text and question, pictures and question because no question is displayed in goal-free processing. Error bars indicate between subjects standard error of the mean.

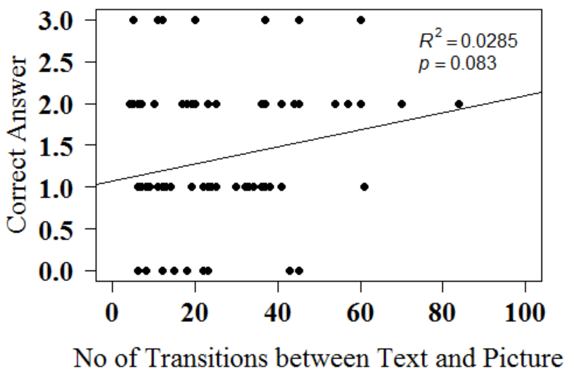
For goal-free processing, there was no robust correlation between accuracy and transitions between text and pictures (see Figure 3.5), $r(72) = .18$, $p = .14$. However, a significant correlation was found in delayed goal-oriented processing between accuracy and transitions between text and pictures, $r(72) = .24$, $p = .05$. In other words, participants, who transited more frequently between text and pictures, had more correct answers in delayed goal-oriented processing.



(a) Goal-free processing



(b) Delayed goal-oriented processing



(c) Initial goal-oriented processing

Figure 3.5: Correlations between accuracy and transitions between text and pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders. X-axis refers to the number of transitions between text and pictures. Y-axis refers to the number of correct answers. The dots represent single participants and the lines represent the trends of the correlations. R^2 and p value show the variability of the data.

3.2.1.2 Goal-free processing vs initial goal-oriented processing

The eye fixation patterns were also significantly different between goal-free processing (no question yet) and initial goal-oriented processing (question before material). The ANOVAs with different eye tracking indicators disclosed a main effect of task orientation on text and pictures. However, no effect was revealed from latency of fixation on text or pictures and on transitions between text and pictures.

(1) Accumulated fixation duration during text-picture integration

Participants fixated shorter on text with goal-free processing than with initial goal-oriented processing (see Figure 3.2), $F(1,71) = 6.43$, $p = .01$, ${\eta_p}^2 = .08$. Contradictorily, they fixated remarkably longer on pictures with goal-free processing than with initial goal-oriented processing, $F(1,71) = 33.48$, $p < .001$, ${\eta_p}^2 = .32$. With initial goal-oriented processing, participants fixated longer on text than on pictures, $F(1,71) = 90.66$, $p < .001$, ${\eta_p}^2 = .56$.

(2) Time to the first fixation during text-picture integration

The ANOVA showed that text was fixated significantly quicker with goal-free processing than with initial goal-oriented processing (see 3.3), $F(1,71) = 6.15$, $p = .02$, ${\eta_p}^2 = .08$. Picture was focused much slower with goal-free processing than with initial goal-oriented processing, $F(1,70) = 18.86$, $p < .001$, ${\eta_p}^2 = .21$. Participants fixated on text much slower than on pictures with initial goal-oriented processing, $F(1,71) = 19.77$, $p < .001$, ${\eta_p}^2 = .22$.

(3) Number of transitions during text-picture integration

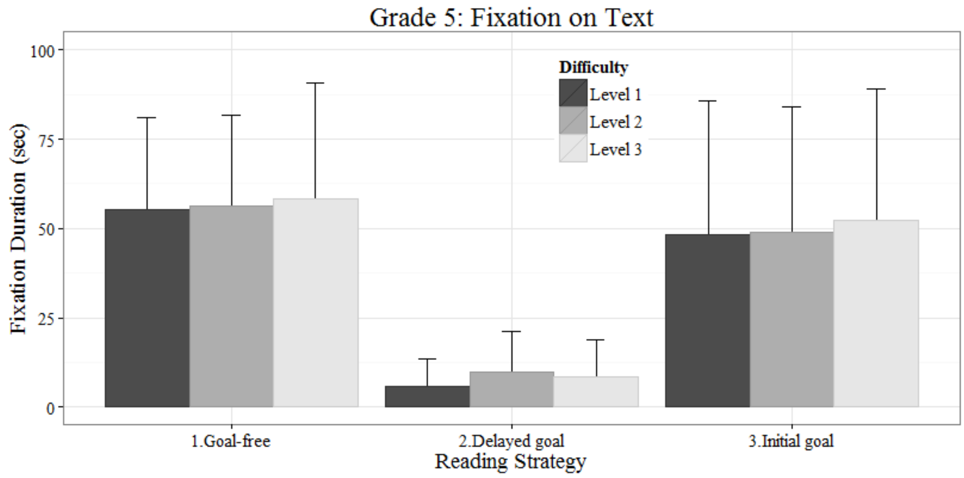
Figure 3.4 illustrated that participants had similar transitions between text and pictures with goal-free processing ($M = 26.59$, $SD = 11.87$) and with delayed goal-oriented processing ($M = 26.49$, $SD = 17.45$). With initial goal-oriented processing, participants had 40.5% ($SD = 14.9\%$) transitions between text and pictures, 20.5% ($SD = 11.9\%$) transitions between text and question, 39.1% ($SD = 14.3\%$) transitions between pictures and question. No robust correlation was found between accuracy and transitions between text and pictures in initial goal-oriented processing (see Figure 3.5), $r(72) = .16$, $p = .18$.

3.2.2 Question 2: Does preference of readers for text vs. pictures depend on question difficulty?

For question difficulty, participants are assumed to obtain more correct answers with easy questions than difficulty questions. Text is expected to be fixated longer than pictures with easy questions than difficult questions. Pictures are assumed to be emphasised with difficult question due to the easy access of information in pictures (Pinker, 1990). Unsurprisingly, participants had higher accuracy with easy questions than difficult questions. Text was fixated similarly with easy and difficult questions. Picture was focused much longer with difficult question than with easy question. Nevertheless, the presumed pattern was not shown: participants tended to perceive text similarly no matter the presented question was easy or difficult. In contrast, they perceived pictures quicker with easy question than with difficult question.

Level 2 and level 3 questions were combined together into level 2/3 questions, owing to their similar design, accuracy and fixation patterns. Level 2 and level 3 questions were designed to compare relations of elements in the materials. Without any doubt, no difference was shown between accuracy of level 2 questions ($M = 40\%$, $SD = 37\%$) and level 3 questions ($M = 35\%, SD = 33\%$), $t(71) = .80$, $p = .43$. Besides, there was no difference for level 2 and level 3 questions in fixation duration across all the reading strategies (3.6), $F < 1$, ns. Consequently, level 2 and level 3 questions were merged into level 2/3 questions.

Participants had better scores with easy questions ($M = 60\%$, $SD = 35\%$) than with difficult questions ($M = 37\%$, $SD = 24\%$), $F(1,71) = 22.74$, $p < .001$, ${\eta_p}^2 = .24$. Paralleled to task orientation, two repeated-measures ANOVAs were performed to check whether question difficulty influences participants' processing with verbal and pictorial information. (1) Goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material). (2) Goal-free processing vs. initial goaloriented processing (question before material). As there was no question in goal-free processing, it was considered as baseline for the comparisons.



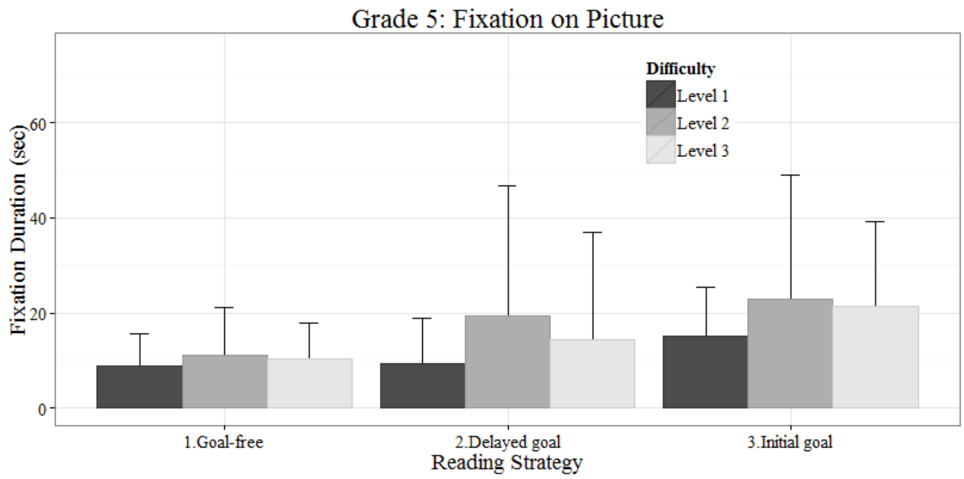


Figure 3.6: Accumulated fixation duration on text (upper graph) and on pictures (lower graph) from 5th graders for solving level 1 (easy), level 2 (medium) and level 3 (difficult) questions with goal-free processing, delayed goal-oriented processing and initial goal-oriented processing. Error bars indicate between subjects standard error of the mean.

3.2.2.1 Goal-free processing vs delayed goal-oriented processing

The (2 × 2) ANOVAs were performed with *question difficulty* (level 1 vs. level 2/3) and *task orientation* (goal-free vs. delayed goal-oriented processing) as within subject factors to check whether readers have preference to text or pictures as question difficulty increases (see Table 3.3). Text and pictures were fixated longer with difficult questions (processing of relation) than with easy questions (processing of value). Pictures were more emphasised than text with difficult questions. Participants perceived text similarly but pictures quicker with easy question than with difficult question.

Table 3.3: The viewing time on text and on picture and the latency for first fixation on text and on pictures when 5th graders were asked to solve level 1 and level 2/3 questions with goal-free processing, delayed goal-oriented processing and initial goal-oriented processing. The grey areas in the table show the baseline of visualisation, when questions are not displayed.

Grade 5	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Level1-Text	55.21 (25.80)	5.82 (7.66)	48.22 (37.47)
Level2/3-Text	57.35 (23.72)	9.00 (7.67)	50.69(28.36)
Level1-Picture	8.84 (6.69)	9.30 (9.57)	15.08 (1029)
Level2/3-Picture	10.70 (6.84)	16.97 (17.74)	22.10 (18.69)
	b. Time to first fixation (sec)		
Level1-Text	3.54 (3.04)	7.71 (10.10)	4.18 (4.91)
Level2/3-Text	2.35 (2.04)	7.22 (9.79)	4.03 (6.15)
Level1-Picture	5.44 (11.54)	5.18 (10.96)	0.62 (1.75)
Level2/3-Picture	10.20 (17.93)	5.84 (5.47)	1.12 (5.23)

(1) Accumulate fixation duration during text-picture integration

For goal-free and delayed goal-oriented processing, the (2 × 2) ANOVA showed no difference of fixation on text with easy question and with difficult question (see Figure 3.7), $F(1,71) = 2.31$, $p = .13$, $\eta_p{}^2 = .03$. Participants fixated significantly longer on text with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 368.38$, $p < .001$, $\eta_p{}^2 = .84$. Other effect was not significant: *question difficulty* × *task orientation*, $F < 1$, ns.

For picture processing, participants fixated much longer on pictures with difficult question than with easy question, $F(1,71) = 14.16$, $p < .001$, $\eta_p{}^2 = .17$. They fixated significantly longer on pictures when they needed to answer questions than the condi-

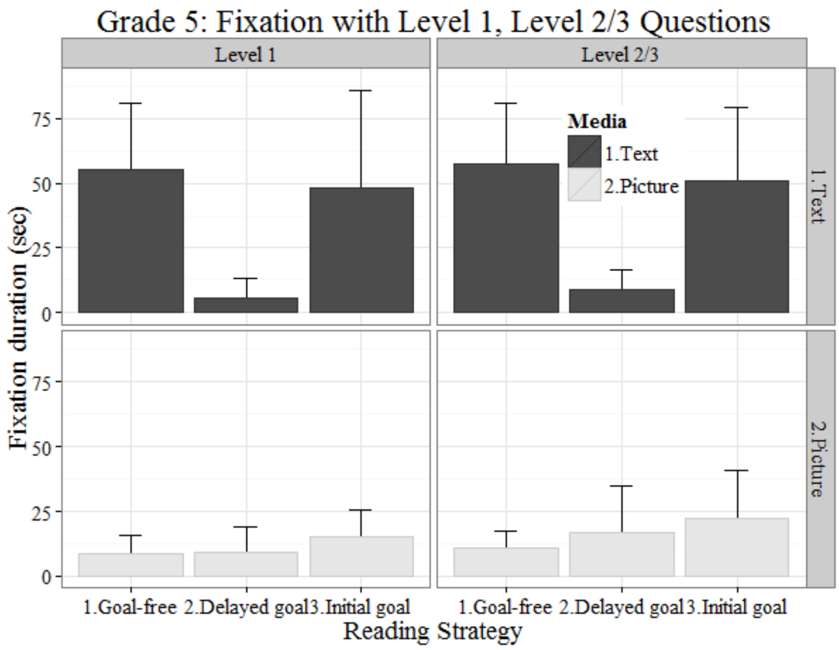
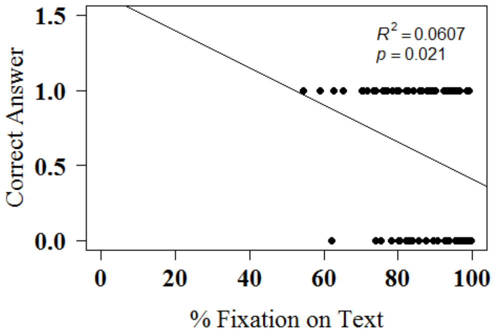


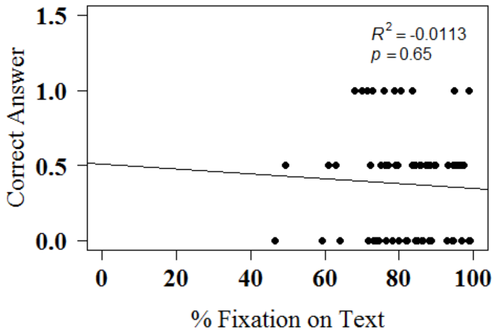
Figure 3.7: Fixation duration on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 5^{th} graders. Goal-free processing is displayed as the baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

tion without question, $F(1,71) = 5.98$, $p = .02$, ${\eta_p}^2 = .08$. A significant interaction was found: *question difficulty* × *task orientation*, $F(1,71) = 4.36$, $p = .04$, ${\eta_p}^2 = .06$.

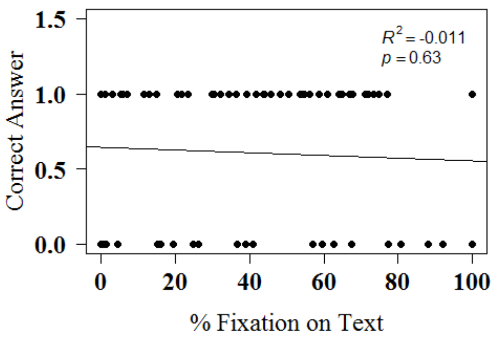
The correlation between correctness and eye tracking indicators (Figure 3.8 (upper) and Figure 3.8 (lower)) was performed in order to check whether higher values in eye tracking indicators indeed suggest the intensive integration of text and pictures. For easy question, the correlation (Spearman Rankorder) illustrated that participants with the right answer had significantly smaller proportion of fixation on text with goal-free processing, $r(72) = -.28$, $p = .02$. There was no robust correlation of accuracy and proportion of fixation on text with delayed goal-oriented processing, $r(72) = -.03$, $p = .83$. For level 2/3 question, there were no robust correlations between accuracy and fixation, $r(72) = -.08$, $p = .51$ for goal-free processing; $r(72) = -.03$, $p = .79$ for delayed goal-oriented processing. As the proportion of fixation on text and pictures is 100%, the results for pictures were not reported to avoid data redundancy.



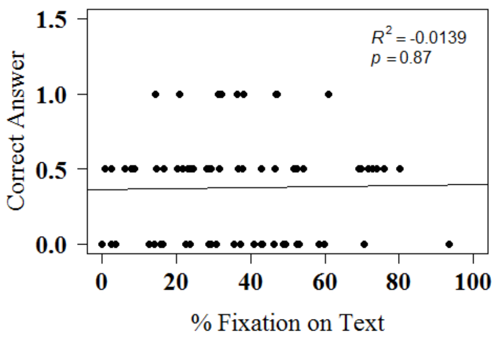
(a) Level 1: Goal-free



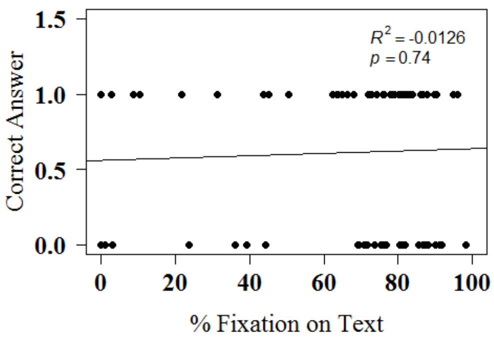
(b) Level 2/3: Goal-free



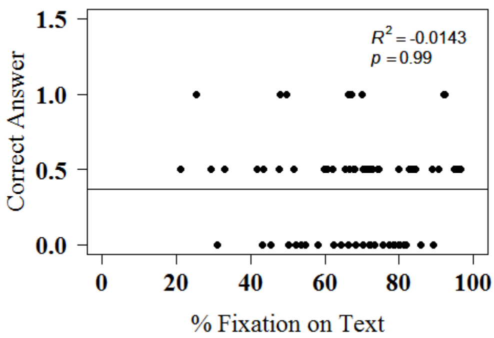
(c) Level 1: Delayed-goal



(d) Level 2/3: Delayed-goal



(e) Level 1: Initial-goal



(f) Level 2/3: Initial-goal

Figure 3.8: Correlations between accuracy and fixation proportion on text among 5th graders with level 1 and level 2/3 questions in goal-free, delayed goal-oriented and initial goal-oriented processing. X-axis refers to the proportion of fixation on text, which is the fixation on text divided by the fixation on text and pictures. Y-axis refers to the number of correct answers. The dots represent single participants and the lines represent the trends of the correlations. R^2 and p value show the variability of the data.

(2) Time to the first fixation during text-picture integration

For text processing, there was no difference in latency of first fixation between easy question and difficult question (see Figure 3.9), $F(1,71) = 1.15$, $p = .29$, $\eta_p{}^2 = .02$. Participants fixated more slowly on pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 25.65$, $p < .001$, $\eta_p{}^2 = .27$. Other effect was not significant: *question difficulty* × *task orientation*, $F < 1$, ns.

For picture processing, participants perceived pictures sooner with easy question than with difficult question, $F(1,71) = 6.82$, $p = .01$, $\eta_p{}^2 = .09$. Other effects were not significant: task orientation, $F(1,71) = 2.30$, $p = .13$, $\eta_p{}^2 = .03$; *question difficulty* × *task orientation*, $F(1,71) = 3.35$, $p = .07$, $\eta_p{}^2 = .05$.

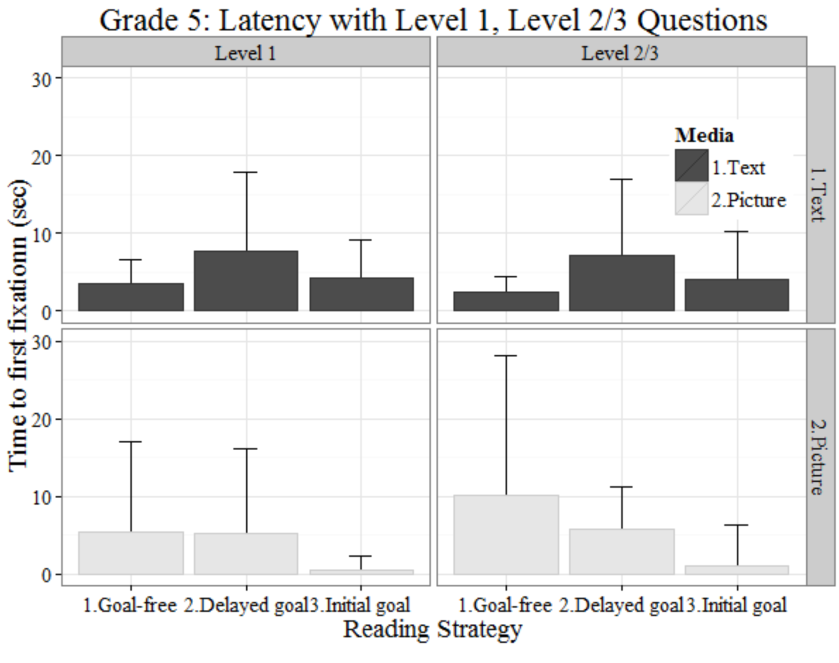


Figure 3.9: Time to the first fixation on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 5th graders. Goal-free processing is displayed as the baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

3.2.2.2 Goal-free processing vs initial goal-oriented processing

The (2 × 2) ANOVA was performed with *question difficulty* (level 1 vs. level 2/3), *task orientation* (goal-free processing vs. initial goal-oriented processing) as within subjects factors. Participants fixated similarly on text but they fixated much longer on pictures with level 1 question (processing of value) than with level 2/3 questions (processing of relations) with initial goal-oriented processing. Besides, the fixation patterns across question levels also showed the robust differences on latency of fixation on text and on pictures with initial goal-oriented processing.

(1) Accumulate fixation duration during text-picture integration

For text processing, participants focused similarly on text with easy question and difficult question (see Figure 3.7), $F < 1$, ns. They fixated shorter on text with goal-free processing than with initial goal-oriented processing, $F(1,71) = 6.33$, $p = .02$, ${\eta_p}^2 = .08$. No interaction was found: *question difficulty* × *task orientation*, $F < 1$, ns.

However, participants fixated significantly longer on pictures with difficult questions than with easy questions, $F(1,71) = 13.69$, $p < .001$, ${\eta_p}^2 = .16$. They also focused more on pictures when they needed to answer the questions than reading without tasks, $F(1,71) = 38.17$, $p < .001$, ${\eta_p}^2 = .35$. A significant effect was detected: *question difficulty* × *task orientation*, $F(1,71) = 3.86$, $p = .05$, ${\eta_p}^2 = .05$.

(2) Time to the first fixation during text-picture integration

No difference was found in latency of first fixation on text with easy question and difficult question (see Figure 3.9), $F(1,71) = 2.06$, $p = .16$, ${\eta_p}^2 = .03$. Participants focused quicker on text with goal-free processing than with initial goal-oriented processing, $F(1,71) = 5.35$, $p = .02$, ${\eta_p}^2 = .07$. No other effect was detected: *question difficulty* × *task orientation*, $F(1,71) = 1.03$, $p = .31$, ${\eta_p}^2 = .02$.

For picture processing, they fixated quicker on pictures with easy question than with difficult question, $F(1,71) = 9.63$, $p = .003$, ${\eta_p}^2 = .02$. They focused significantly slower on pictures with goal-free processing than with initial goal-oriented processing,

$F(1,71) = 18.91$, $p < .001$, ${\eta_p}^2 = .21$. A significant interaction was detected: *question difficulty* × *task orientation*, $F(1,71) = 5.70$, $p = .02$, ${\eta_p}^2 = .07$.

(3) Correlation between correctness and eye tracking indicators

Correlation between accuracy and proportion of fixation duration on text was carried out to check whether fixation indicates integrating text and pictures. As illustrated in Figure 3.8, there were no robust correlations between accuracy and proportion of fixation on text with easy and difficult questions, $r(72) = .01$, $p = .92$ for easy question; $r(72) = .002$, $p = .99$ for difficult question.

3.2.3 Question 3: What are the differences in text processing and picture processing between students from higher and lower school tiers?

According to Hypothesis 3, participants from higher school tier are presumed to perform more accurately than participants from lower school tier. They can process the verbal information more efficiently but the pictorial information more thoroughly than participants from lower school tier. The accuracy results were in line with the hypothesis that participants from higher school tier (Gymnasium) (M = 50%, SD = 22%) performed significantly better than participants from lower school tier (Realschule) (M = 40%, SD = 18%), $F(1,70) = 4.79$, $p = .03$, ${\eta_p}^2 = .06$. In addition, students from higher school tier fixated longer on pictures rather than on text than students from lower school tiers.

To examine the effect of school tier, the 2 (× 2) repeated-measures ANOVAs were conducted with *school tier* (Gymnasium vs. Realschule) as between subject factor and *reading strategy* (goal-free vs. delayed goal-oriented processing) as within subject factor. Only results, that are relevant to school tier, were reported. Correspondingly, eye tracking indicators were compared: (1) goal-free processing (no question yet) vs.

delayed goal-oriented processing (question after material); (2) goal-free processing vs. initial goal-oriented processing (question before material).

3.2.3.1 Goal-free processing vs delayed goal-oriented processing

Higher tier students (Gymnasium) fixated similarly on text but significantly longer on pictures than lower tier students (Realschule). Regardless of school tier (see Table 3.4), all participants had similar latency of first fixation on text and pictures with goal-free processing (no question yet) and with delayed goal-oriented processing (question after material). Besides, higher tier students tended to transfer more frequently between text and pictures than lower tier students with delayed goal-oriented processing. Lower tier students had more transitions between pictures and question than higher tier students.

Table 3.4: Fixation duration, time to first fixation, number of transition on text and on pictures for 5th graders from Gymnasium (Gym) and Realschule (Real).

Grade 5	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Gym-Text	52.76 (15.57)	10.48 (6.73)	53.12 (20.92)
Real-Text	60.52 (24.54)	7.11 (6.30)	46.62 (32.21)
Gym-Picture	10.46 (5.74)	18.73 (14.36)	24.38 (15.48)
Real-Picture	9.69 (5.61)	10.10 (6.49)	15.13 (9.57)
	b. Time to first fixation (sec)		
Gym-Text	2.55 (1.40)	8.52 (7.30)	3.61 (3.80)
Real-Text	2.95 (2.24)	6.25 (7.83)	4.54 (5.11)
Gym-Picture	5.96 (6.31)	5.71 (5.40)	1.41 (4.91)
Real-Picture	11.26 (19.66)	5.53 (5.09)	0.49 (0.84)
	c. Number of fixation		
Gym-Text-Picture	27.47 (20.47)	15.39 (12.29)	30.94 (15.51)
Real-Text-Picture	25.69 (14.50)	8.26 (5.15)	21.91 (18.35)
Gym-Text-Question	NA	12.08 (8.84)	14.19 (10.18)
Real-Text-Question	NA	9.37 (6.89)	10.57 (7.54)
Gym-Picture-Question	NA	27.92 (19.37)	25.28 (15.10)
Real-Picture-Question	NA	21.31 (9.25)	22.37 (12.61)

(1) Accumulated fixation duration during text-picture integration

For text processing, 2 (× 2) ANOVA showed that higher school tier students had similar fixation patterns as lower tier students (see Figure 3.10), $F < 1$, ns. Students from both school tiers fixated much longer on text with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 360.48$, $p < .001$, $\eta_p{}^2 = .84$. A significant interaction was detected: *school tier* × *task orientation*, $F(1,70) = 4.87$, $p = .03$, $\eta_p{}^2 = .07$.

For picture processing, higher tier students fixated much longer on pictures than lower tier students, $F(1,70) = 11.51$, $p = .001$, $\eta_p{}^2 = .14$. They fixated much longer on pictures with delayed goal-oriented processing than with goal-free processing, $F(1,70) = 7.74$, $p = .007$, $\eta_p{}^2 = .10$. A significant effect was revealed: *school tier* × *task orientation*, $F(1,70) = 6.37$, $p = .01$, $\eta_p{}^2 = .08$.

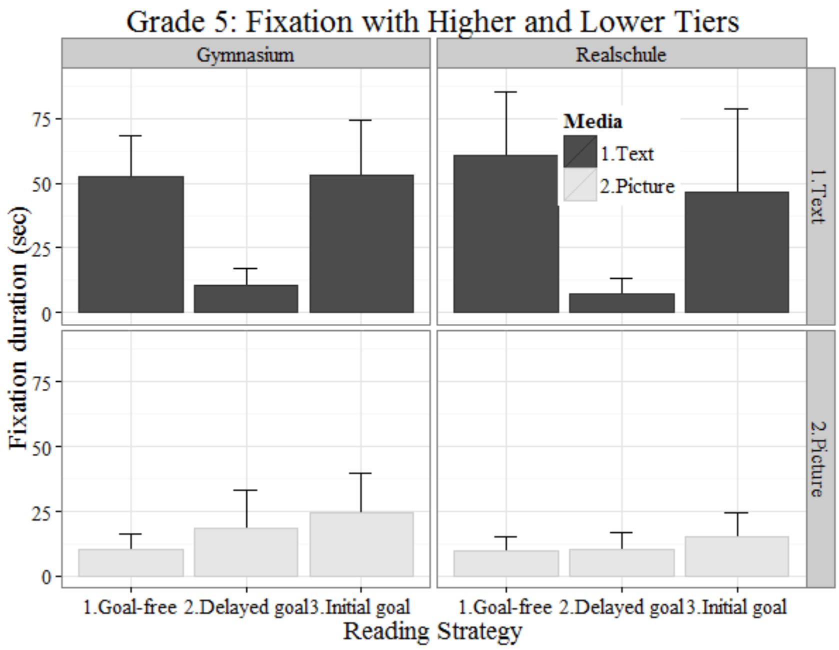


Figure 3.10: Fixation duration on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5. Error bars indicate between subjects standard error of the mean.

(2) Time to the first fixation during text-picture integration

The 2 (× 2) ANOVA on latency of first fixation showed no difference in text processing between higher tier students and lower tier students (see Figure 3.11), $F(1,70) = 1.02$, $p = .32$, $\eta_p{}^2 = .01$. All the students fixated significantly quicker on text with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 25.73$, $p < .001$, $\eta_p{}^2 = .27$. No other effect was detected: *school tier* × *task orientation*, $F(1,70) = 2.12$, $p = .15$, $\eta_p{}^2 = .03$.

Similarly, no difference was revealed on pictures processing between students from two school tiers, $F(1,70) = 1.61$, $p = .21$, $\eta_p{}^2 = .02$. No other effects were found: task orientation, $F(1,70) = 3.43$, $p = .07$, $\eta_p{}^2 = .05$; *school tier* × *task orientation*, $F(1,70) = 2.88$, $p = .09$, $\eta_p{}^2 = .04$.

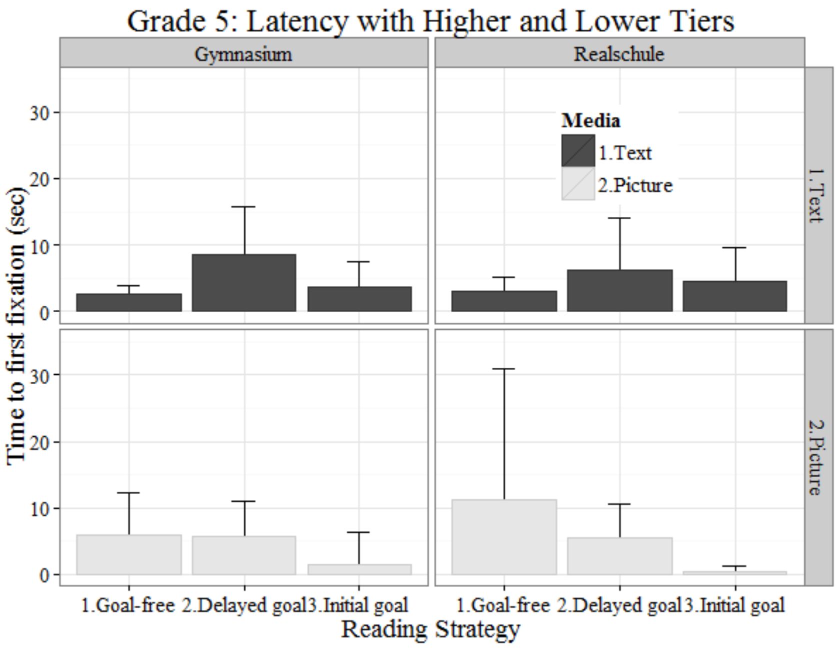


Figure 3.11: Time to the first fixation on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5. Error bars indicate between subjects standard error of the mean.

(3) Number of transitions during text-picture integration

Students from both school tiers had similar transitions between text and pictures (see Figure 3.12), $F(1,70) = 2.92$, $p = .09$, $\eta_p{}^2 = .04$. All the students had more transitions between text and pictures with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 47.22$, $p < .001$, $\eta_p{}^2 = .41$. No interaction was found:

school tier × *task orientation*, $F(1,70) = 1.55$, $p = .22$, $\eta_p{}^2 = .02$.

With delayed goal-oriented processing, higher tier students differed significantly from lower tier students in all the transition patterns, $F(1,70) = 6.34$, $p = .01$, $\eta_p{}^2 = .08$. Higher tier students had 26.4% ($SD = 11.2\%$) transitions between text and pictures, 23.7% ($SD = 13.5\%$) transitions between text and question, 49.9% ($SD = 15.4\%$) transitions between pictures and question. Lower tier students had 19.7% ($SD = 9.9\%$) transitions between text and pictures, 24.2% ($SD = 13.8\%$) transitions between text and question, 56.1% ($SD = 13.6\%$) transitions between pictures and question.

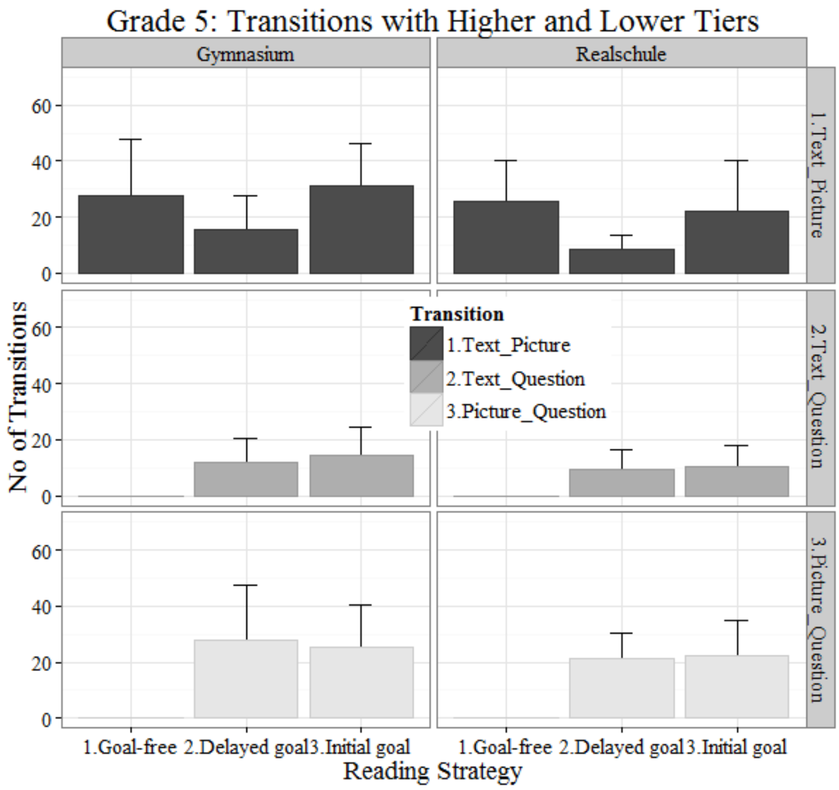


Figure 3.12: Transitions between text, pictures and question with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5. Error bars indicate between subjects standard error of the mean.

3.2.3.2 Goal-free processing vs initial goal-oriented processing

Higher tier students focused similarly on text compared to lower tier students with goal-free processing (no question yet) and initial goal-oriented processing (question before material). Whereas, higher tier students fixated longer on pictures than lower tier students with goal-free and initial goal-oriented processing. Regardless of school tier, there was no difference on delayed time of first fixation on text and on pictures. Besides, participants always had more transitions between text and pictures with initial goal-oriented processing regardless of school tier. Higher tier students seemed to have more transitions between text and pictures than lower tier students with initial goal-oriented processing. Lower tier students transferred their attention more frequently between pictures and question than higher tier students.

(1) Fixation duration during text-picture integration

The 2 (× 2) ANOVA showed that higher tier students processed text in a similar way as lower tier students, $F < 1$, ns. Regardless of school tiers, all participants focused significantly longer on text with goal-free processing than initial goal-oriented processing, $F(1,70) = 7.05$, $p = .01$, ${\eta_p}^2 = .09$. A main interaction was detected: *school tier* × *reading strategy*, $F(1,70) = 7.81$, $p = .007$, ${\eta_p}^2 = .10$.

In contrast, higher tier students fixated significantly longer on pictures than lower tier students, $F(1,70) = 8.62$, $p = .004$, ${\eta_p}^2 = .11$. All participants focused significantly longer on pictures with initial goal-oriented processing than goal-free processing, $F(1,70) = 36.29$, $p < .001$, ${\eta_p}^2 = .34$. A significant interaction was found: *school tier* × *reading strategy*, $F(1,70) = 6.97$, $p = .01$, ${\eta_p}^2 = .09$.

(2) Time to the first fixation during text-picture integration

The 2 (× 2) ANOVA on the delayed time of the first fixation showed no difference on text between students from higher and lower school tiers, $F(1,70) = 1.19$, $p = .28$, ${\eta_p}^2 = .02$. Participants from both tiers fixated quicker on text with goal-free processing than with initial goal-oriented processing, $F(1,70) = 6.09$, $p = .02$, ${\eta_p}^2 = .08$. No interaction was found: *school tier* × *reading strategy*, $F < 1$, ns.

Likewise, no difference was found in latency of first fixation on pictures between two school tiers, $F(1,70) = 1.48$, $p = .23$, ${\eta_p}^2 = .02$. However, all participants fixated quicker on pictures with initial goal-oriented processing than with goal-free processing, $F(1,70) = 19.45$, $p < .001$, ${\eta_p}^2 = .22$. No interaction was detected: *school tier* × *reading strategy*, $F(1,70) = 3.20$, $p = .08$, ${\eta_p}^2 = .04$.

(3) Number of transitions during text-picture integration

Higher tier students transferred their attention marginally more frequently between text and pictures than lower tier students, $F(1,70) = 3.52$, $p = .07$, ${\eta_p}^2 = .05$. All participants had similar transition patterns with goal-free processing and initial goal-oriented processing, $F < 1$, ns. No other effect was found: *school tier* × *reading strategy*, $F(1,70) = 1.61$, $p = .21$, ${\eta_p}^2 = .02$.

With initial goal-oriented processing, higher tier students differed in transition patterns from lower tier students, $F(1,70) = 4.49$, $p = .04$, ${\eta_p}^2 = .06$. Students from higher tier had 44.2% (SD = 12.9%) transitions between text and pictures, 20.8% (SD = 12.4%) transitions between text and question, 35.0% (SD = 11.9%) transitions between pictures and question. Students from lower tier had 36.7% (SD = 16.0%) transitions between text and pictures, 20.1% (SD = 11.5%) transitions between text and question, 43.2% (SD = 15.5%) transitions between pictures and question.

3.3 Discussion

The accuracy and eye movement data from 5th graders suggest that text processing differed from picture processing. Text mainly guides general readings; whereas, pictures primarily guide selective processing. Compared to text, pictures have relative advantages in solving difficult questions due to its specific representational format. Good problem solvers are more capable than poorer problem solvers in comprehending pictures rather than text. Adopted eye tracking method, Experiment 1 provids a deeper look at the goal-free processing and goal-oriented processing proposed by Rickards and Denner (1978) in the domain of text and picture comprehension. Goal-free processing (no question yet) yields processing the material in a global coherent way. In contrast, goal-oriented processing (question before material) leads to rather selective processing of the material.

Experiment 1 was conducted to explore the usage of text and pictures can be influenced by task orientation (location of questions). More specifically, it made the comparisons between (1) goal-free (no question yet) vs. delayed goal-oriented processing (question after material) and (2) goal-free vs. initial goal-oriented processing (question before material). Besides, it also examined whether text processing and picture processing can be moderated by question difficulty and by whether learners from higher or lower school tier.

Text is presumed to cause more fixation patterns with goal-free processing than with delayed goal-oriented processing. Pictures are expected to be more intensively used with delayed goal-oriented processing. Participants are assumed to focus mostly on text although their emphasis on pictures increases with initial goal-oriented processing. More emphasis on text is expected with easy question; whereas, larger value of attention on pictures is hypothesized with difficult question. Students from higher tier spend more time on pictures rather than on text compared to students from lower tier.

3.3.1 Text guides general reading; pictures guide selective reading

Consistent with Hypothesis 1, Experiment 1 confirms the differences between text processing and picture processing. Specifically, processing of text and pictures differs with the experimental controlled reading strategies. Based on eye-mind assumption and immediacy assumption, eye tracking indicators were analysed to interpret text processing and picture processing. The results of accuracy and eye movements on task orientation were in line with the Hypothesis 1 with goal-free processing (no question yet), delayed goal-oriented processing (question after material) and initial goal-oriented processing (question before material).

For the goal-free processing (no question yet), text was heavily emphasized rather than pictures on mental model construction. The accumulated fixation duration implies the depth of cognitive processing and the spatial distribution of attention (Hegarty, 1992). The results indicate the main usage of text during the mental model construction in goal-free processing. Differed from the finding in the pilot study (Zhao et al., 2014), text was perceived earlier than pictures in goal-free processing. As time to the first fixation indicates reader's interest (J. H. Goldberg and Kotval, 1999), the results showed that text was considered as the interesting information source to build the mental model when reading was without any task. Opposite to the assumption from Mason et al. (2014), participants' reading routine can be influenced by reading strategy. Furthermore, participants transferred their attention frequently between text and pictures with goal-free processing. The correlation was not significant between accuracy and transitions between text and pictures. Nevertheless, a trend showed (see Figure 3.5) that frequent transitions between text and pictures may lead to better accuracy. In a cognitive level, this may indicate the integration of text and pictures according to the Integrative Model of TPC. In agreement with Hypothesis 1, text is mainly processed to build the mental model in goal-free processing.

For delayed goal-oriented processing (question after material), pictures were intensively used for scaffolding questions after the initial mental model was constructed. Data from fixation duration confirmed Hypothesis 1 that pictures were primarily involved in question answering in delayed goal-oriented processing. In addition, pictures were fixated sooner than text in delayed goal-oriented processing. It can be interpreted as the interest on or the importance of pictures to the reader. Data from transitions showed that the most transitions were between pictures and question, then between text and pictures and the least were between text and pictures. The correlation between accuracy and transitions between text and pictures reveals that the higher value of transitions between text and pictures indeed imply more correct answers. This is consistent with the idea that pictures are mainly used to scaffold question answering. The most transitions were between pictures and question. It indicates that participants tend to focus mainly on pictures to scaffold the question after constructed the initial mental model in goal-free processing. Participants had more transitions between text and question. It implies that participants may evaluate the initial mental model with the question and may update the mental model by visiting the text. Finally, they had the least transitions between text and pictures due to the construction of initial mental model in goal-free processing. The result confirms the hypothesis and reveals unequal processing of text and pictures in the Integrative Model of TPC (Schnotz, 2014). Briefly, it suggests that pictures primarily guide mental model usage to answer questions.

For initial goal-oriented processing, participants focused less on text but more on pictures compared to reading without tasks. Likely, participants build their mental model optionally with text and pictures according to the question. They use pictures as external representations to scaffold the question answering, which causes the higher value of fixation on pictures. Picture had shorter latency of first fixation than text in initial goal-oriented processing. It also supports the idea that pictures are primarily used as a tool to scaffold question answering. The transition data showed the most transitions were between text and pictures, more transitions were between pictures and question and the least transitions were between text and question. Due to the design of materials, participants are required to use information from both text and pictures. Thus, after knowing the question, they focus mainly on text and pictures to construct the mental model, which rather is selective. They mainly use pictures to answer the question, which is in conformity with the hypothesis. The least transitions are between text and question. It can be explained that participants update their mental model by using text. The results are in line with the hypothesis that text is mainly used for mental model construction; whereas, pictures are primarily used for mental model usage.

3.3.2 Pictures are more emphasized than text with difficult question

Furthermore, participants fixated longer on text and on pictures as question difficulty increased. However, attention on pictures were more emphasized than on text when the question was difficult. It seems that participants recognize the difficulty of the question and start to change their emphasis from text into pictures. It is consistent with the assumption that pictures can guide and scaffold question answering. It also corresponds to the suggestions on the superiority of pictures in comprehending the difficult materials (Carney and J. R. Levin, 2002). As defined by Schnotz (2014), depictions (i.e. pictures) have advantages in directly displaying and visualizing the structure of the information. Participants had no difference in the latency of first fixation on text with easy and difficult questions. It implies that participants have similar interest to the text no matter whether the question is easy or difficult. Conversely, participants fixated quicker on pictures with easy questions than with difficult questions. It can be interpreted as participants have more interest in pictures with easy questions than with difficult questions. After they comprehend the question, they are less hesitate to use pictures with easy questions than with difficult questions. It can also imply the advantages of pictures that participants can extract information rather quicker from pictures than from text.

A strong negative correlation is revealed between accuracy and proportion of fixation on text with easy questions in goal-free processing. Participants with higher accuracy distribute less attention to text and more attention to pictures while building their initial mental model. On the one hand, it can be explained that some 5th

graders have problems in text reading as fixation duration may indicate the difficulty of cognitive processing (Rayner and Castelhano, 2007). Those who read efficiently answered more correct answers. On the other hand, it can indicate that using pictures or knowing how to use the pictures may lead to better performance in accuracy. As readers comprehend much longer on pictures, the ability of picture processing may affect the mental model construction. They may have a deep structure in picture comprehension. Those who have less attention on pictures tend to have surface structure. Deep structure of picture comprehension can lead to better retention of information than surface structure (Schnotz and Baadte, 2014). The results support the inequality of text processing and picture processing (McNamara, 2007). In disagreement with the assumption, pictures rather than text are highly emphasised as the difficulty of question increases.

3.3.3 Pictures seem to be the key to distinguish between good vs poorer problem solvers

There is a distinction between higher tier (Gymnasium) and lower tier (Realschule) schools in Germany. Hence, this study considers higher tier students (Gymansium) as good problem solvers; whereas, it considers lower tier students (Realschule) as poorer problem solvers. Higher tier students had more accurate answers than lower tier students. However, higher tier students spent roughly the same amount of time on text as lower tier students. This is consistent with the idea that advanced readers are more skillful in selecting the most important verbal information (Cerdán et al., 2011) and decoding the verbal information (Cataldo and Oakhill, 2000) and making references from graphics (Hegarty and Sims, 1994). Besides, higher tier students had considerably longer fixation on pictures than lower tier students. It indicates that higher tier students have deep processing of pictorial information than lower tier students. In line with the suggestion from Bluth (1973), higher tier students tend to benefit more from pictures than lower tier students. As mentioned by Schnotz, Ludewig, et al. (2014), higher tier students possibly know the strategy to use pictures to scaffold question answering. Most importantly, they seem to be more able to interpret the pictorial information than lower tier students. Furthermore, text and pictures were perceived similarly between higher tier students and lower tier students. They are likely to have similar interest in text and in pictures.

Participants from both school tiers transferred their attention similarly between text and pictures when no question was posted. In contrast, higher tier students had more transitions between text and pictures than lower tier students. Lower tier transferred more frequently between pictures and question. Participants from both school tiers had similar transitions between text and question. According to the Integrative Model of TPC, higher tier students are more able to use both text and pictures to update the mental model after the construction of initial mental model in goal-free processing. As mentioned by Rusted and Coltheart (1979), lower tier students may have more problem in superficial verbal processing, such as semantic recognition. Thus, they resort to pictures to compensate for the inadequacy of verbal comprehension and transfer their attention frequently between pictures and question.

3.3.4 Gender, age and types of illustration issues

Due to the difficulty of data recruitment with young participants in Germany, the gender is unfortunately not balanced in Experiment 1. Research on gender effect in multimedia learning points out that gender may differ in reading scientific materials (Yang, Huang, and Tsai, 2014), efficiency (Pöhnl and Bogner, 2012), interest in design of learning interfaces (Passig and H. Levin, 1999). Although the accuracy may not differ between female and male, they tend to apply different reading strategies. A study from BITE project concerning gender difference finds that females and males differ in choosing reading strategies (Wagner and Schnotz, in press). Male participants mainly rely on spatial ability compared to female participants when answering questions. However, no difference is detected in accuracy of answers.

Due to the limited illustration formats in this study, it is not clear which type of illustrations can facilitate learning the most. Different formats of pictures should be considered because a reader may comprehend bars, lines or graphics differently (Doherty and R. B. Anderson, 2009; Shah and Freedman, 2009; Zacks and Tversky, 1999). Further research should focus more on which specific formats of pictures can better facilitate picture processing.

Furthermore, Experiment 1 focuses mainly on 5^{th} graders in text and picture integration. They just finished elementary school and are still learning vocabulary and the results may be biased to higher graders in secondary school, such as 8^{th} graders. As a consequence, readers from different grades may have different reading patterns with multimedia presentations. It is implied that the effect of task orientation depends also on the reader's age (Seretny and Dean, 1986; Tal et al., 1994; van den Broek et al., 2001). The accuracy increases from lower grade to higher grade (Golden, 1942). It takes readers from lower grade longer time than readers from higher grade to comprehend the materials (Schnotz, 2014). Young readers seem to have difficulty in establishing a coherent representation when questions are posed after the material. Likely, young readers' working memory capacity is occupied with superficial level of processing, such as semantic recognition, syntactic decomposition (Carpenter et al., 1994; Oakhill and Yuill, 1996). Understanding the materials may interfere with integrating the demands from questions, searching for corresponding materials and organising the information. In contrast, older readers are able to automatically process the information and they have enough space in working memory to comprehend the questions (van den Broek et al., 2001). They can possibly benefit from questions during reading because questions can direct their attention to the requested information.

In addition, younger students have different reading patterns on text and on pictures from older students. Younger students are inclined to pay more attention to text than older students (Hannus and Hyönä, 1999; Hochpöchler et al., 2013). Conversely, they fixate less time on pictures than older students. It seems that young students are less able to amend the processing of pictures to the demand of tasks. As the questions become more and more difficult, young students still have similar performance on pictures. In contrast, older students are more sensitive to identify the demands of tasks. They are more skilled at using pictures to fulfil the tasks (Hartman, 2001). Younger readers may not know how to deal with pictures. They may have difficulty in performing the suitable strategy for picture processing (Hasselhorn, 1996). As a consequence, Experiment 2 is designed to address the issues.

3.3.5 Training on multimedia comprehension

The results in regard to task orientation, question difficulty and school tier are particularly interesting not only because it indicates the unequalised processing of text and pictures but also the different functions of text and pictures. General reading is primarily guided by text and selective reading is normally guided by pictures. When question is difficult, pictures are more emphasized than text because pictures can visualize the internal structure of information. Higher tier students differ from lower tier students not in verbal processing but mainly in pictorial processing. However, there is a tendency to a text-oriented in school curriculum in Germany (Kultusministerkonferenz, 2005a). Consequently, school science curriculum may introduce the results from Experiment 1. Teachers may provide more training on text-picture comprehension, owing to the fundamental differences between text processing and picture processing.

Chapter 4

Experiment 2

Posing question before or after the material can activate different reading strategies. Data from 5^{th} graders show that the usage of text is different from the usage of picture with different reading strategy. Text mainly guides general readings; whereas pictures primarily guide selective reading. When the difficulty of the question increases, participants are inclined to fixate much longer on pictures than text. Surprisingly, higher tier students only differ from lower tier students in picture processing rather than text processing. It seems that pictures play crucial roles in obtaining correct answers. However, 5^{th} graders have just finished their elementary school and their reading comprehension is not yet contingent on advanced vocabulary skills (Schnotz et al., 2010). As Perfetti (2010) suggests, vocabulary may affect comprehension when decoding of language is not automatic enough. Thus the influence may be caused by their immature language ability. Does this influence also apply to older students? To answer this question, Experiment 2 is carried out to investigate whether reading strategy, question difficulty and school tier indeed have an impact on advanced readers.

It is assumed that a reader possibly focuses more on text than pictures when s/he is required to understand the materials. In contrast, a reader may fixate longer on pictures than on text when s/he is required to answer the questions. The reader may fixate similarly on text but longer on pictures with difficult questions than with easy questions. Higher tier students probably process verbal information more efficiently and pictorial information more intensively than lower tier students. The results from advanced readers are in conformity with Hypothesis 1 on task orientation. Differed from Hypothesis 2, they fixate longer on text and on pictures as question difficulty increases. However, the increase of attention on pictures is greater than the increase of attention on text. Higher tier students process verbal information similarly as lower tier students. However, higher tier students focused more on pictorial information than lower tier students.

4.1 Method

Based on the hypotheses, Experiment 2 was conducted to examine whether higher graders use text differently from pictures when question is presented before or after the material. Participants were recruited from higher school tier (Gymnasium) and lower school tier (Realschule) in 7^{th} and 8^{th} grade. Their accuracy and eye movements were recorded during reading. Task orientation, question difficulty and school tier showed different degrees of influence on multimedia comprehension.

4.1.1 Participants from Grade 8

Experiment 2 was designed to have the same amount of participants as in experiment 1 (N = 72) based on the power analysis. Seventy-one 8^{th} graders and nine 7^{th} graders (due to the difficulty in recruiting 8^{th} graders) from Germany participated in Experiment 2. Eight participants were excluded due to calibration, motivation and gender problem. Due to the difficulty of balancing the gender in Experiment 2, the number of participants were selected as the compensate gender group as Experiment 1. For instance, data in Experiment 1 were analysed from 44 males and 28 females. In Experiment 2, twenty-eight (38.9%) were males and forty-four (61.1%) were females (M = 14.5 years, SD = 0.7 years). Sixty-three were 8^{th} graders (87.5% of the participants) and nine were 7^{th} graders (12.5% of the participants).

The intelligence tests for 8^{th} graders were used from the same book (Heller and Perleth, 2000) to check their spatial and verbal abilities. Students from Gymnasium had better scores than the average in spatial ability (average T of 51.64; $SD = 7.24$) and verbal ability (average T of 52.61; $SD = 7.48$). Students from Realschule had slightly worse scores than the average in spatial ability (average T of 48.00; $SD = 7.72$) and verbal ability (average T of 46.22; $SD = 5.92$). Students from 7^{th} grade performed similarly in spatial ability (average T of 47.22; $SD = 5.45$) and verbal ability (average T of 43.33; $SD = 4.64$). Their performance on intelligence test was similar to students from 8^{th} grade (spatial ability: average T of 50.19; $SD = 7.88$ and verbal ability: average T of 50.27; $SD = 7.39$), $t(12.64) = -1.49$, $p = .16$. The data from 7^{th} graders and 8^{th} graders were merged into one group 8^{th} graders, considering the similarity between two grades and the small amount of 7^{th} graders.

Except the participants came from higher grade, Experiment 2 used the same design, apparatus, materials, procedure and scoring system as in Experiment 1.

4.2 Results

It is hypothesised that text mainly guides general reading and it facilitates mental model construction. Conversely, pictures mainly guide selective reading and it scaffolds question answering. Text is primarily processed, when the presented question is easy. In contrast, pictures are emphasized when the question is difficult. Besides, students from higher school tier may process text more eficiently than students from lower school tier. Conversely, students from higher school tier are likely to comprehend pictures more intensively than students from lower school tier.

participants from higher and lower school tiers may differ in multimedia comprehension on task orientation and question difficulty. Similar to Experiment 1, the results will be reported in three perspects to answer the research questions: accuracy, eye movements and the correlation between them. The accuracy data showed the average accuracy of solving six tasks with multimedia presentations (text and pictures) among 8^{th} graders from higher (Gymnasium) and lower school tiers (Realschule). Due to the higher proportion of goal-oriented processing in two processing, eye tracking indicators were compared: (1) Goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material); (2) goal-free processing vs. initial goal-oriented processing (question before material). The correlation between accuracy and eye movements suggested the indication of eye movements in cognitive processing.

Data from accuracy and eye movements indicated that text was processed differently from pictures when question was presented before or after the material among 8^{th} graders. In confirmaty with Hypothesis 1, participants mainly focused on text with goal-free processing (no question yet). They mainly emphasized on pictures with delayed goal-oriented processing (question after material). They fixated mainly on text but the fixation on text and pictures was rather balanced with initial goal-oriented processing (question before material). Inconsistent with Hypothesis 2, participants fixated longer on text and on pictures when they answered difficult questions compared to easy questions. However, pictures were more focused than text as the question difficult increased. Incompatable with Hypothesis 3, higher tier students processed pictures more intensively than lower tier students. However, no difference was detected in text processing among students from both school tiers. Besides, the results confirmed that eye tracking indicators can indicate the integration of text and pictures.

4.2.1 Question 1: How do readers deal with text and deal with pictures in different forms of reading?

Participants are presumed to have similar accuracy in post-question condition (reading followed by question) and pre-question condition (reading guided by question). Text is supposed to be mainly focused when only text and pictures are presented. Picture should be emphasised when question is asked. The results showed that participants had similar accurate answers in post-question and pre-question conditions. They mainly emphasized on text when they comprehend the material. Conversely, they mainly fixated on pictures when they were required to solve the question.

The accurate answers refer to the percentage of obtaining correct answers in six materials. Participants from 8^{th} grade had similar performance on accuracy with post-question (M = 68%, SD = 27%) and with pre-question (M = 63%, SD = 25%), $F(1,71) = 1.50$, $p = .23$, $\eta_p^2 = .02$. Yet, text was processed fundamentally different from pictures with different reading strategies. The accumulated number of fixation was highly correlated (r = .94) with the accumulated fixation duration. Therefore, three eye tracking indicators were reported: accumulated fixation duration, time to the first fixation and number of transitions.

To examine whether text processing differs from picture processing regarding task orientation, the repeated-measures analyses of variance (ANOVAs) were carried out on each eye tracking indicator. Two comparisons were undertaken because two processing containing higher proportion of selective processing. (1) Goal-free (no question yet) vs. delayed goal-oriented processing (question after material); (2) goal-free vs. initial goal-oriented processing (question before material).

4.2.1.1 Goal-free processing vs delayed goal-oriented processing

Table 4.1 illustrated robust differences in text processing and in picture processing with goal-free processing (no question yet) and delayed goal-oriented processing (question after material). They emphasized on text when no question was displayed. Conversely, they focused mainly on pictures when the question is shown after their prior experience with the material. They paid rather balanced attention to text and pictures when question is presented first, enabling students to selectively process the question. They perceived text sooner than pictures when no question is displayed. In contrast, they perceived pictures sooner than text when question was displayed after the prior experience with the material. Participants transferred their attention frequently between text and pictures when no question is displayed, whereas they had frequent transitions between pictures and question when question is displayed after the prior experience with the material.

Table 4.1: Fixation duration, time to the first fixation on text and pictures and transitions between text, pictures and question with three reading processing among 8th graders.

Grade 8	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Text	39.89 (9.15)	8.94 (5.31)	37.34 (14.82)
Picture	11.09 (5.66)	15.62 (13.69)	18.98 (9.36)
	b. Time to first fixation (sec)		
Text	2.00 (1.60)	5.55 (3.77)	2.40 (2.99)
Picture	2.73 (3.77)	4.42 (3.58)	1.09 (4.69)
	c. Number of fixation		
Text-Picture	36.97 (16.78)	17.32 (8.90)	39.18 (23.74)
Text-Question	NA	26.94 (12.62)	25.21 (13.67)
Picture-Question	NA	12.68 (6.85)	13.33 (8.03)

(1) Accumulated fixation duration during text-picture integration

The results from ANOVA demonstrated that participants focused much longer on text with goal-free processing than with delayed goal-oriented processing (see Figure 4.1), $F(1,71) = 661.42$, $p < .001$, $\eta_p^2 = .90$. However, they fixated shorter on pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 6.74$, $p = .01$, $\eta_p^2 = .09$. With goal-free processing, they fixated much longer on text than on pictures, $F(1,71) = 451.51$, $p < .001$, $\eta_p^2 = .86$. With delayed goal-oriented processing, they focused much longer on pictures than on text, $F(1,71) = 19.60$, $p < .001$, $\eta_p^2 = .22$.

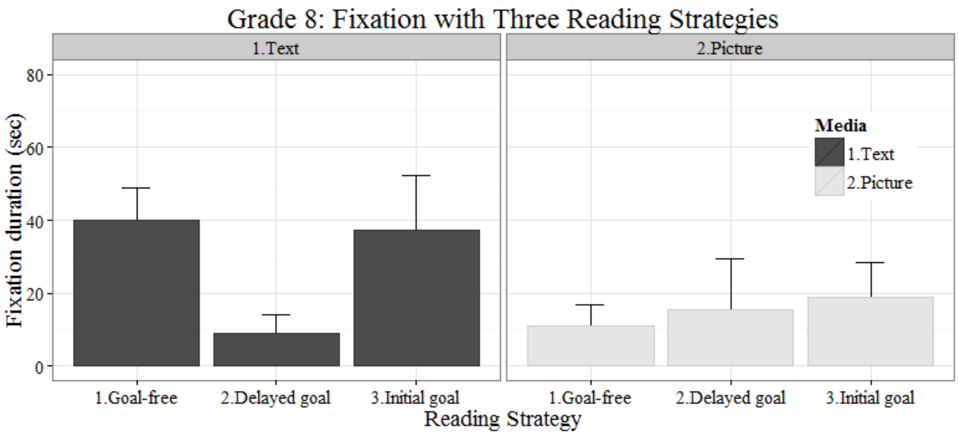


Figure 4.1: Accumulated fixation duration on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 8^{th} graders. Error bars indicate between subjects standard error of the mean.

(2) Time to the first fixation during text-picture integration

Participants viewed text sooner with goal-free processing than with delayed goal-oriented processing (see Figure 4.2), $F(1,71) = 60.15$, $p < .001$, $\eta_p^2 = .46$. They perceived pictures quicker with delayed goal-oriented processing than with goal-free processing, $F(1,71) = 7.97$, $p = .006$, $\eta_p^2 = .10$. No difference was detected on latency of first fixation on text and on pictures with goal-free processing, $F(1,71) = 2.06$, $p = .16$, $\eta_p^2 = .03$, and with delayed goal-oriented processing, $F(1,71) = 3.20$, $p = .08$, $\eta_p^2 = .04$.

(3) Number of transitions during text-picture integration

Participants had frequent transitions between text and pictures with goal-free processing (M = 36.97, SD = 16.78) compared to delayed goal-oriented processing (M = 17.32, SD = 8.9). With goal-free processing, participants had 52.4% (SD = 2.0%) transitions from text to pictures and 47.6% (SD = 2.0%) transitions from pictures to text with goal-free processing. With delayed goal-oriented processing, participants had 23.1% (SD = 11.1%) transitions between text to pictures, 22.7% (SD = 9.5%) transi-

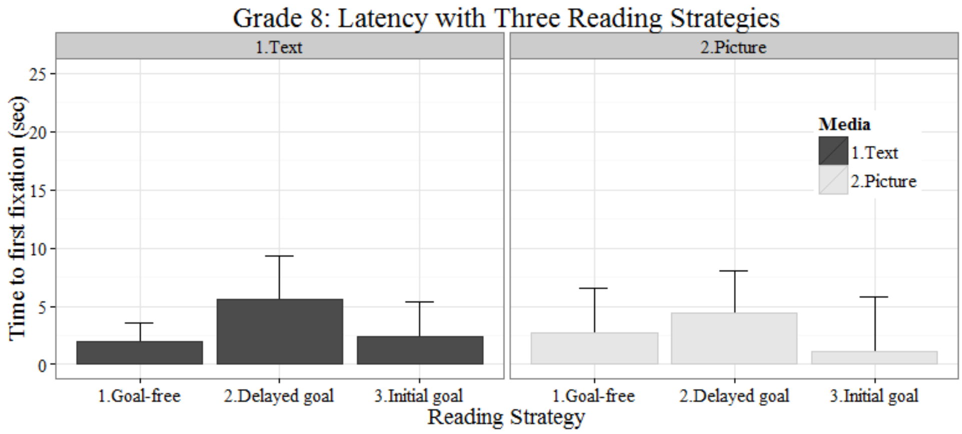


Figure 4.2: Time to the first fixation on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 8th graders. Error bars indicate between subjects standard error of the mean.

tions between text and question, 53.0% (*SD* = 14.8%) transitions between pictures and question.

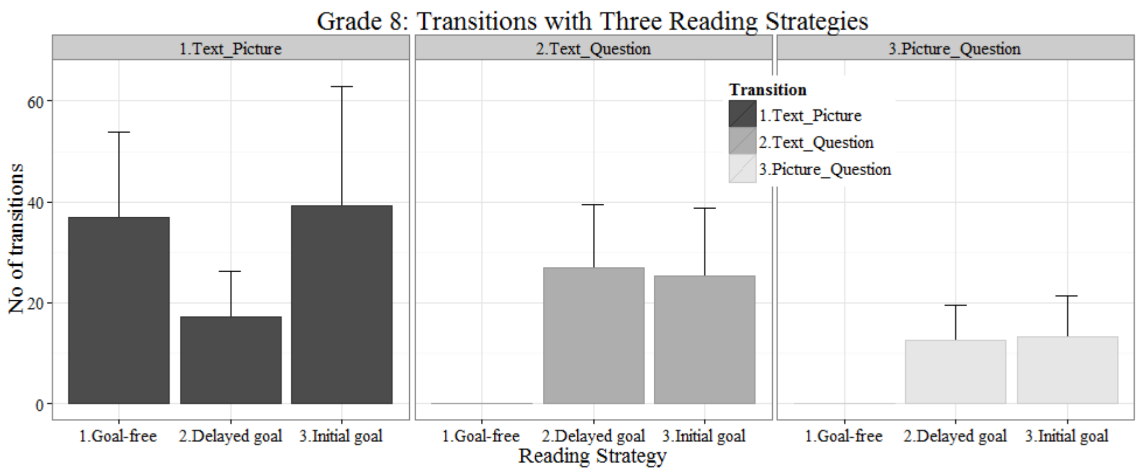
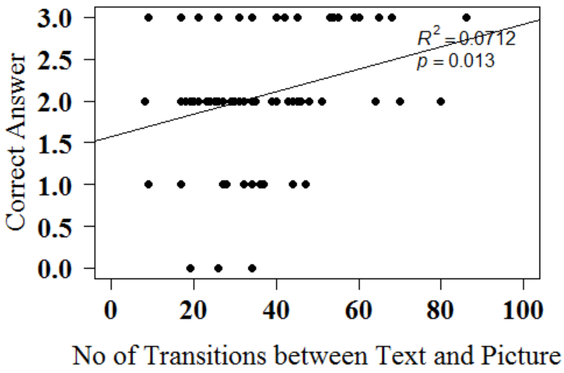
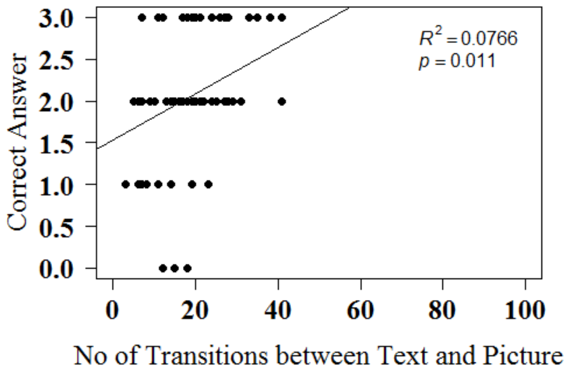


Figure 4.3: Transitions among text, pictures and question in goal-free, delayed goal-oriented and initial goal-oriented processing among 8th graders. No transitions are illustrated between text and question, pictures and question because goal-free processing does not deal with questions. Error bars indicate between subjects standard error of the mean.

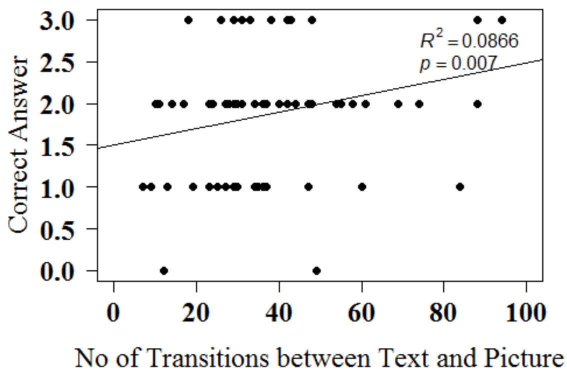
Significant correlations were shown between accuracy and transitions between text and pictures, $r(72) = .27$, $p = .02$ for goal-free processing. There was also significant correlation between accuracy and transitions between text and pictures (see Figure 4.4), $r(72) = .38$, $p = .001$ for delayed goal-oriented processing.



(a) Goal-free processing



(b) Delayed goal-oriented processing



(c) Initial goal-oriented processing

Figure 4.4: Correlations between accuracy and transitions between text and pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among 8th graders. X-axis refers to the number of transitions between text and pictures. Y-axis refers to the number of correct answers. The dots represent single participants and the lines represent the trends of the correlations. R^2 and p value show the variability of the data.

4.2.1.2 Goal-free processing vs initial goal-oriented processing

Participants comprehended text similarly with goal-free processing (no question yet) and with initial goal-oriented processing (question before material). In contrast, pictures was emphasised with initial goal-oriented processing. Participants perceived text quicker with goal-free processing but they viewed pictures sooner with initial goal-oriented processing. No difference was detected on transitions between text and pictures in both processing.

(1) Accumulated fixation duration during text-picture integration

Participants had similar viewing time on text with goal-free processing and initial goal-oriented processing, $F(1,71) = 2.94$, $p = .09$, $\eta_p^2 = .04$. In contrast, they viewed significantly longer on pictures with initial goal-oriented processing than with goal-free processing, $F(1,71) = 47.03$, $p < .001$, $\eta_p^2 = .40$. With initial goal-oriented processing, participants fixated much longer on text than on pictures, $F(1,71) = 158.58$, $p < .001$, $\eta_p^2 = .69$.

(2) Time to the first fixation during text-picture integration

They had similar latency on text with goal-free and initial goal-oriented processing, $F(1,71) = 1.10$, $p = .30$, $\eta_p^2 = .02$. Conversely, they fixated quicker on pictures with initial goal-oriented processing than with goal-free processing, $F(1,71) = 5.44$, $p = .02$, $\eta_p^2 = .07$. With initial goal-oriented processing, they fixated much quicker on pictures than on text, $F(1,71) = 3.93$, $p = .05$, $\eta_p^2 = .05$.

(3) Number of transitions during text-picture integration

Participants had similar numbers of transitions between text and pictures with goal-free ($M = 36.97$, $SD = 16.78$) and initial goal-oriented processing ($M = 39.18$, $SD = 23.74$). Participants had 48.9% ($SD = 13.1\%$) transitions between text and pictures, 17.5% ($SD = 7.8\%$) transitions between text and question, 33.7% ($SD = 11.8\%$) transitions between pictures and question for initial goal-oriented processing. A robust correlation was shown between accuracy and transitions between text and pictures in initial goal-oriented processing (see Figure 4.4), $r(72) = .28$, $p = .02$.

4.2.2 Question 2: Does preference of readers for text vs. pictures depend on question difficulty?

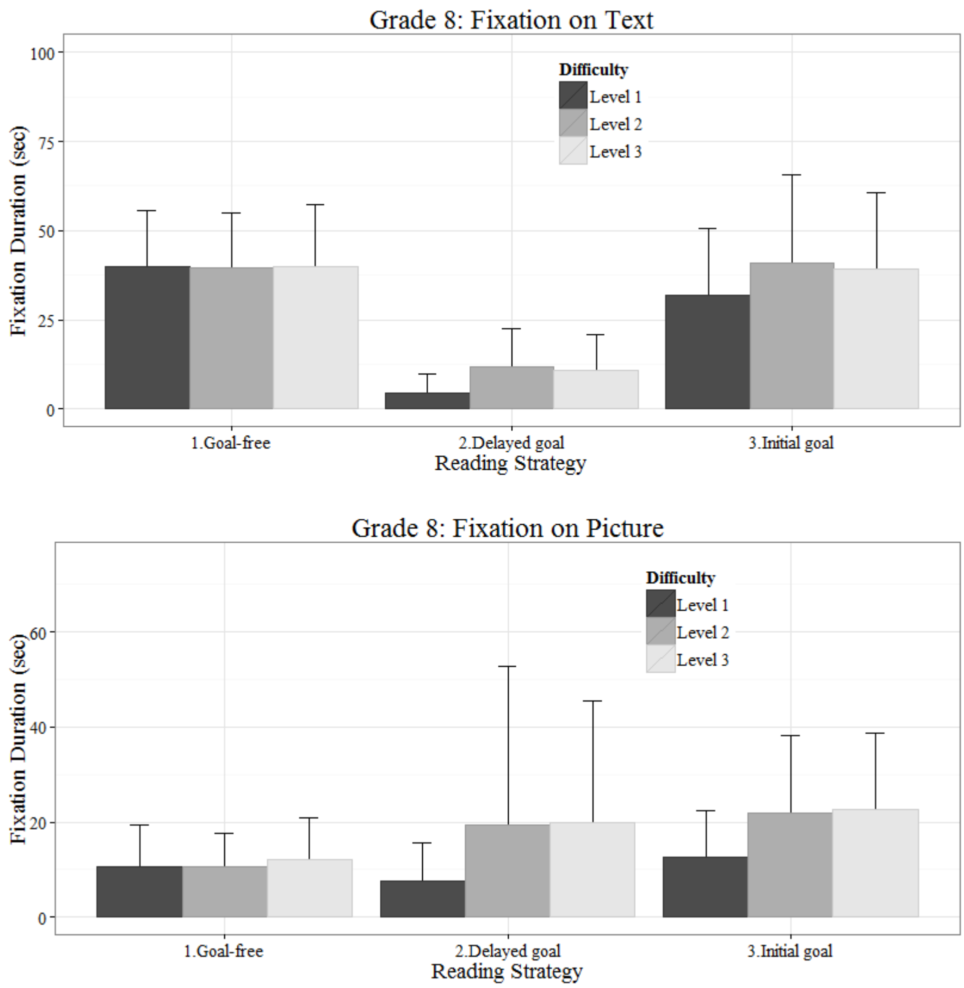


Figure 4.5: Accumulated fixation duration on text (upper graph) and on pictures (lower graph) from 8th graders for solving level 1 (easy), level 2 (medium) and level 3 (difficult) questions with goal-free processing, delayed goal-oriented processing and initial goal-oriented processing. Error bars indicate between subjects standard error of the mean.

According to Hypothesis 2, participants may have higher percentage of accuracy with easy question than with difficult question. For easy question, they possibly focus more on text than on pictures. For difficult question, they may fixate longer on pictures than on text due to its priority in conveying information by its spatial stuctures (Larkin and Simon, 1987). Without any doubt, the participants performed much better with easy questions than with difficult questions. Slightly differed from the hypothesis, they fixated much longer on both text and on pictures with difficult questions than with easy questions. No difference was found on latency of first fixation on text and pictures with easy and difficult questions.

Similar to Experiment 1, level 2 and level 3 were combined into level 2/3 due to the similar design of question difficulty and the similar learning outcome from the participants. Level 1 question checked participant's ability of mapping values, while level 2 and level 3 questions inspected their ability of processing relations. Unsurprisingly, participants had similar scores with level 2 (M = 63%, SD = 38%) and level 3 questions (M = 50%, SD = 38%), $t(71) = 1.85$, $p = .07$. They had also similar fixation patterns with level 2 and level 3 questions (see Figure 4.5), $F < 1$, ns. To make the data easier to interpret, level 2 and 3 questions were combined into level 2/3.

Participants performed significantly better in accuracy with easy question (M = 85%, SD = 31%) than with difficult question (M = 57%, SD = 23%), $F(1,71) = 48.33$, $p < .001$, $\eta_p^2 = .41$. In order to explore the effect of question difficulty in text processing and picture processing, eye tracking indicators were analysed between: (1) goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material); (2) goal-free processing vs. initial goal-oriented processing (question before material). As no question was displayed in goal-free processing, this phase was used as the baseline.

4.2.2.1 Goal-free processing vs delayed goal-oriented processing

To examine the influence of question difficulty in multimedia learning, the (2 x 2) ANOVAs were performed with *question difficulty* (level 1 vs. level 2/3) and *task orientation* (goal-free vs. delayed goal-oriented processing) as within subjects factors.

The results (see Table 4.2) showed that participants fixated much longer on text and on pictures with level 2/3 question (processing of relations) than level 1 question (processing of value). A significant correlation was found with difficult question in delayed goal-oriented processing between accuracy and fixation duration on text. Participants had similar latency of first fixation on text and pictures for easy and difficult questions.

Table 4.2: The viewing time on text and on pictures and the latency for first fixation on text and on pictures when 8th graders were asked to solve level 1 and level 2/3 questions with goal-free processing, delayed goal-oriented processing and initial goal-oriented processing. The grey areas in the table show the baseline of visualisation when questions are not displayed.

Grade 8	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Level1-Text	40.05 (15.56)	4.37 (5.57)	31.88 (18.74)
Level2/3-Text	39.81 (10.99)	11.23 (7.58)	40.07 (17.56)
Level1-Picture	10.64 (8.70)	7.62 (7.97)	12.63 (9.74)
Level2/3-Picture	11.31 (6.02)	19.62 (20.44)	22.13 (15.80)
	b. Time to first fixation (sec)		
Level1-Text	2.07 (1.54)	4.73 (5.06)	2.07 (2.85)
Level2/3-Text	1.97 (2.05)	5.96 (5.38)	2.57 (4.23)
Level1-Picture	2.66 (4.19)	4.22 (5.39)	0.87 (3.00)
Level2/3-Picture	2.76 (5.18)	4.52 (4.77)	1.20 (5.79)

1) Accumulated fixation duration during text-picture integration

The (2 x 2) ANOVAs revealed that participants fixated significantly longer on text with difficult question than with easy question (see Figure 4.6), $F(1,71) = 6.79$, $p = .01$, $\eta_p^2 = .09$. They fixated much longer on text with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 657.54$, $p < .001$, $\eta_p^2 = .90$. A main effect was detected: *question difficulty* × *task orientation*, $F(1,71) = 9.12$, $p = .004$, $\eta_p^2 = .11$.

Participants had similar fixation patterns on pictures as on text. They fixated much longer on pictures as the difficulty increased, $F(1,71) = 17.50$, $p < .001$, $\eta_p^2 = .20$. They had similar fixation patterns on pictures with goal-free and delayed goal-oriented pro-

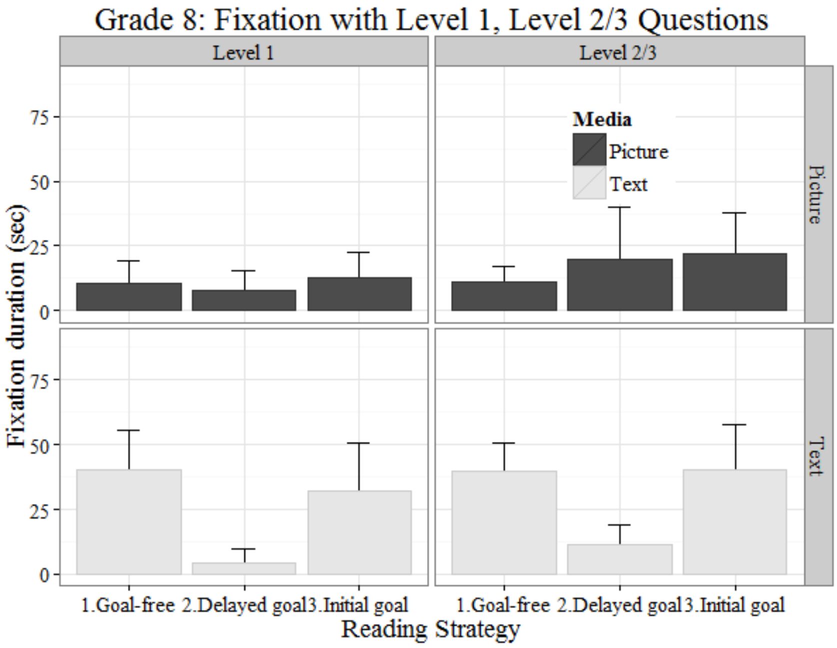
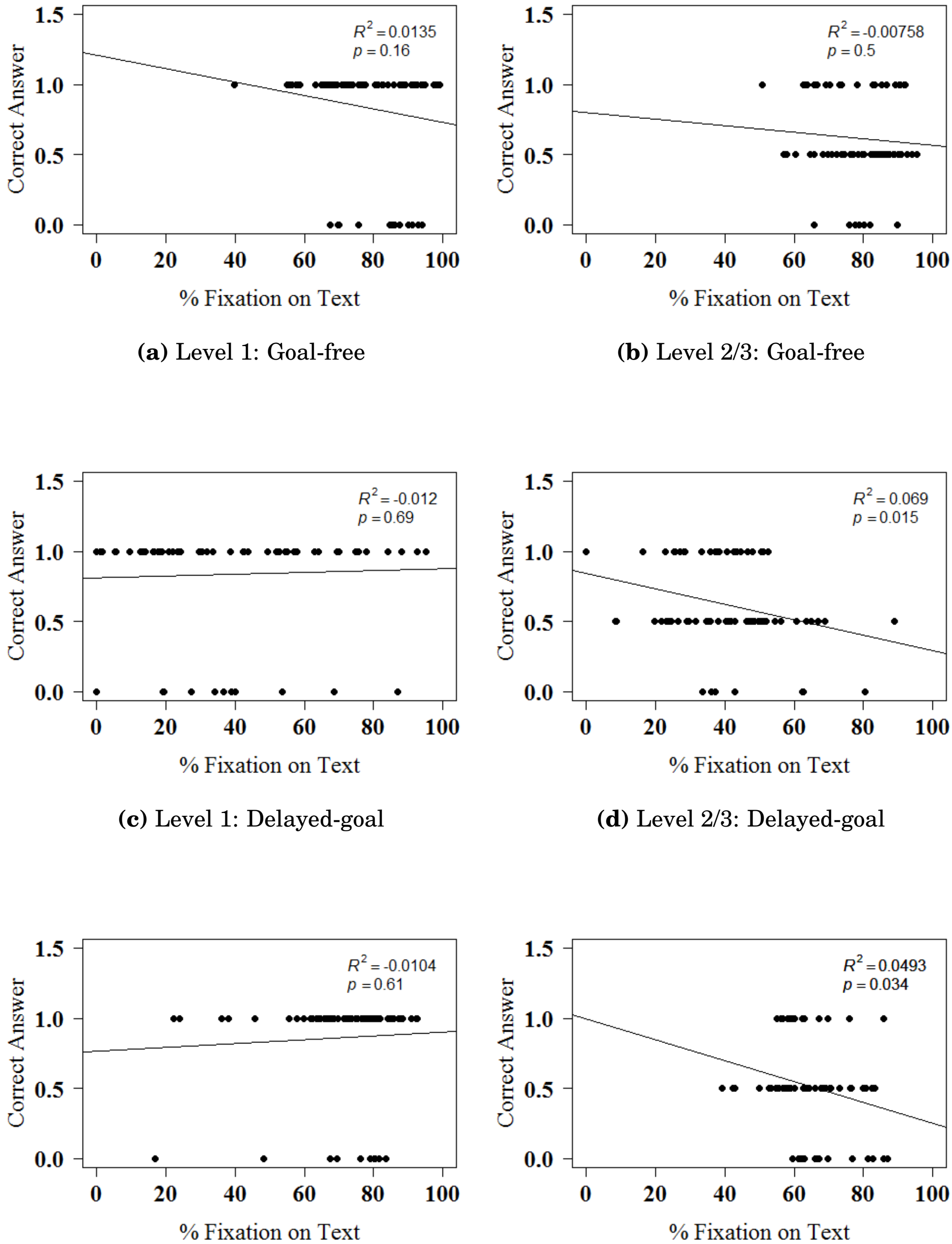


Figure 4.6: Fixation duration on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 8th graders. Goal-free processing is displayed as the baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

cessing, $F(1,71) = 3.16$, $p = .08$, $\eta_p^2 = .04$. A significant effect was detected: *question difficulty* × *task orientation*, $F(1,71) = 18.21$, $p < .004$, $\eta_p^2 = .20$.

The correlation between accuracy and eye indicators were performed on easy and difficult questions. It aims at examining whether eye tracking indicators suggest the cognitive processing. For easy questions, no robust correlations were shown (Figure 4.7) between accuracy and proportion of fixation on text, $r(72) = -.15$, $p = .21$ for goal-free processing; $r(72) = -.04$, $p = .75$ for delayed goal-oriented processing. Results from pictures were not reported to avoid data redundancy because the sum of proportion of fixation on text and pictures was 100%.

For difficult questions, no correlations were detected between accuracy and proportion of fixation on text with goal-free processing, $r(72) = -.07$, $p = .56$. In contrast, participants had a higher accuracy rate if they fixated shorter on text with delayed goal-oriented processing, $r(72) = -.34$, $p = .05$.



(a) Level 1: Goal-free

(b) Level 2/3: Goal-free

(c) Level 1: Delayed-goal

(d) Level 2/3: Delayed-goal

(e) Level 1: Initial-goal

(f) Level 2/3: Initial-goal

Figure 4.7: Correlations between accuracy and fixation proportion on text among 8th graders with level 1 and level 2/3 questions in goal-free, delayed goal-oriented and initial goal-oriented processing. X-axis refers to the proportion of fixation on text, which is the fixation on text divided by the fixation on text and pictures. Y-axis refers to the number of correct answers. The dots represent single participants and the lines represent the trends of the correlations. R^2 and *p* value show the variability of the data.

(2) Time to the first fixation

The (2 x 2) ANOVA on the latency of first fixation disclosed the similar fixation patterns on text with easy question and difficult question (see Figure 4.8), $F(1,71) = 1.32$, $p = .25$, $\eta_p^2 = .02$. They fixated quicker on text with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 64.45$, $p < .001$, $\eta_p^2 = .48$. No other effect was found: *question difficulty* × *task orientation*, $F(1,71) = 2.03$, $p = .16$, $\eta_p^2 = .03$.

For picture processing, no difference was found with easy question and with difficult question, *question difficulty x task orientation*, $F < 1$, ns. They fixated significantly quicker on pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 9.17$, $p = .003$, $\eta_p^2 = .11$. No other effect was detected: *question difficulty* × *task orientation*, $F < 1$, ns.

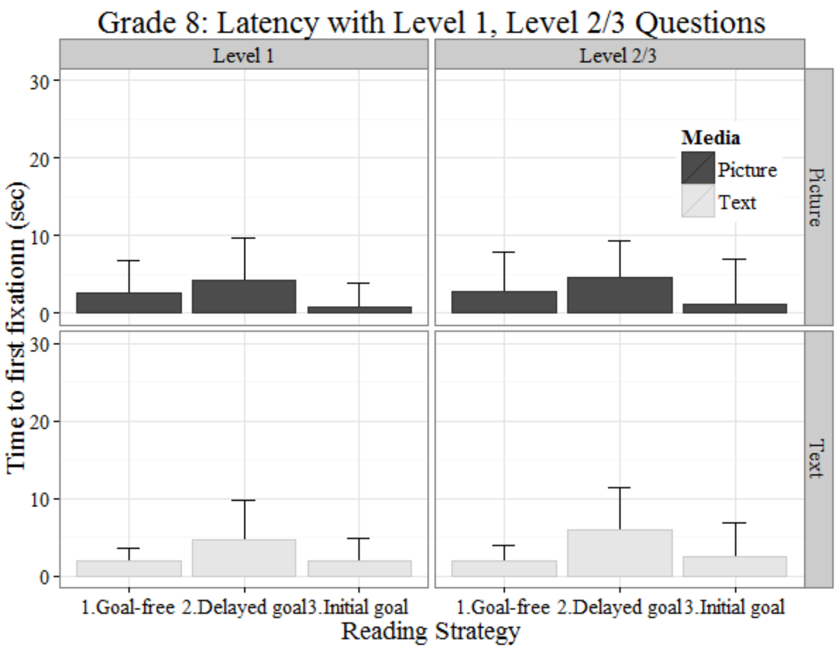


Figure 4.8: Time to the first fixation on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 8^{th} graders. Goal-free processing is displayed as the baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

4.2.2.2 Goal-free processing vs initial goal-oriented processing

Similarly, the (2 × 2) ANOVA was performed with *question difficulty* (level 1 vs. level 2/3), *task orientation* (goal-free processing vs. initial goal-oriented processing) as within subjects factors. Participants had longer fixation time on text and on pictures to solve difficult questions than to solve easy questions with goal-free processing (no question yet) and with initial goal-oriented processing (question before material). They viewed pictures quicker than text with initial goal-oriented processing.

(1) Accumulated fixation duration during text-picture integration

For text processing, the (2 x 2) ANOVA showed that participants fixated marginally more on text with difficult question than with easy question (see Figure 4.6), $F(1,71) = 3.46$, $p = .07$, $\eta_p^2 = .05$. They fixated longer on text with goal-free processing than with initial goal-oriented processing, $F(1,71) = 6.57$, $p = .01$, $\eta_p^2 = .09$. A main effect was found: *question difficulty* × *task orientation*, $F(1,71) = 16.63$, $p < .001$, $\eta_p^2 = .19$.

For picture processing, they fixated significantly longer on pictures with difficult questions than with easy questions, $F(1,71) = 21.19$, $p < .001$, $\eta_p^2 = .23$. They fixated longer on pictures with initial goal-oriented processing than with goal-free processing, $F(1,71) = 30.41$, $p < .001$, $\eta_p^2 = .30$. A main effect was detected: *question difficulty* × *task orientation*, $F(1,71) = 25.63$, $p < .001$, $\eta_p^2 = .27$.

(2) Time to the first fixation during text-picture integration

The (2 x 2) ANOVA on delayed time of first fixation on text showed no difference between question difficulty, $F < 1$, ns; task orientation, $F < 1$, ns; *question difficulty* × *task orientation*, $F < 1$, ns.

Participants fixated quicker on pictures with initial goal-oriented processing than with goal-free processing, $F(1,71) = 7.49$, $p = .008$, $\eta_p^2 = .10$. No other effect was found: task orientation, $F < 1$, ns; *question difficulty* × *task orientation*, $F < 1$, ns.

(3) Correlation between correctness and eye tracking indicators

The correlations between accuracy and fixation duration with level 1 and level 2/3 questions were conducted to explore how well the eye tracking indicators can suggest integrative processing of text and pictures. For easy questions (see Figure 4.7), no robust correlations were found between proportion of fixation on text with initial goal-oriented processing, $r(72) = -.02$, $p = .89$. Results from pictures were not reported to avoid data redundancy. For difficult questions, participants who had more correct answers fixated shorter on text with initial goal-oriented processing, $r(72) = -.29$, $p = .01$.

4.2.3 Question 3: What are the differences in text processing and pictures processing between students from higher and lower school tiers?

According to Hypothesis 3, participants from higher school tier (Gymnasium) are expected to perform better in accuracy than participants from lower school tier (Realschule). Text is fixated longer by lower tier students, whereas pictures are mainly focused by higher tier students. The performance on accuracy was in line with the hypothesis. However, participants from both school tiers had similar fixation patterns on text. Higher tier students seemed to have considerably longer fixation on pictures than lower tier students.

Correspond to Hypothesis 3, participants from higher school tiers ($M = 72\%$, $SD = 17\%$) performed better than participants from lower school tiers ($M = 59\%$, $SD = 19\%$), $F(1,70) = 10.10$, $p = .002$, $\eta_p^2 = .13$. Two repeated-measures ANOVAs were performed in order to explore whether participants from both school tiers differ in multimedia learning. (1) Goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material). (2) Goal-free processing vs. initial goal-oriented processing (question before material).

4.2.3.1 Goal-free processing vs delayed goal-oriented processing

To explore the influence of school factor, 2 (× 2) repeated-measure ANOVAs were performed with between-factor *school tier* (Gymnasium vs. Realschule) and within-factors *reading strategy* (goal-free vs. delayed goal-oriented processing). Data only relevant to school are reported to avoid data redundancy (see Table 4.3). Higher tier students had similar fixation duration on text with lower tier students. However, higher tier students fixated much longer on pictures than lower tier students. There was no difference on latency of first time to text and pictures between higher and lower tier students. Higher tier students had frequent transitions between text and pictures; whereas, lower tier students had frequent transitions between pictures and question.

Table 4.3: Fixation duration, time to first fixation, number of transition on text and on pictures for 8^{th} graders from Gymnasium (Gym) and Realschule (Realschule).

Grade 8	Goal-free processing *M (SD)*	Delayed goal-oriented processing *M (SD)*	Initial goal-oriented processing *M (SD)*
	a. Accumulated fixation duration (sec)		
Gym-Text	39.36 (9.63)	8.77 (5.24)	36.62 (13.87)
Real-Text	40.41 (8.75)	9.11 (5.45)	38.07 (15.88)
Gym-Picture	12.12 (6.27)	18.90 (17.38)	21.04 (9.24)
Real-Picture	10.06 (4.84)	12.34 (7.49)	16.92 (9.14)
	b. Time to first fixation (sec)		
Gym-Text	2.30 (1.98)	5.50 (3.07)	2.14 (2.41)
Real-Text	1.71 (1.03)	5.60 (4.41)	2.66 (3.49)
Gym-Picture	2.91 (4.14)	3.61 (2.90)	0.90 (2.74)
Real-Picture	2.54 (3.42)	5.23 (4.02)	1.28 (6.08)
	c. Number of fixation		
Gym-Text-Picture	42.50 (18.19)	19.72 (10.34)	42.69 (24.96)
Real-Text-Picture	31.44 (13.33)	14.92 (6.47)	35.67 (22.25)
Gym-Text-Question	NA	12.94 (6.45)	13.44 (8.22)
Real-Text-Question	NA	12.42 (7.31)	13.22 (7.96)
Gym-Picture-Question	NA	28.92 (14.67)	27.72 (13.58)
Real-Picture-Question	NA	24.97 (9.99)	22.69 (13.48)

(1) Accumulated fixation duration during text-picture integration

The 2 (× 2) ANOVA showed that higher tier students viewed text in the same manner as lower tier students (see Figure 4.9), $F < 1$, ns. Students from both tiers fixated significantly longer on text with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 652.90$, $p < .001$, $\eta_p^2 = .90$. No interaction was found: *school tier* × *task orientation*, $F < 1$, ns.

For picture processing, higher tier students viewed much longer on pictures than lower tier students (see Figure 4.9), $F(1,70) = 6.56$, $p = .01$, $\eta_p^2 = .09$. All students fixated longer on pictures with delayed goal-oriented processing than with goal-free processing, $F(1,70) = 6.80$, $p = .01$, $\eta_p^2 = .09$. No interaction was detected: *school tier* × *task orientation*, $F(1,70) = 1.67$, $p = .20$, $\eta_p^2 = .02$.

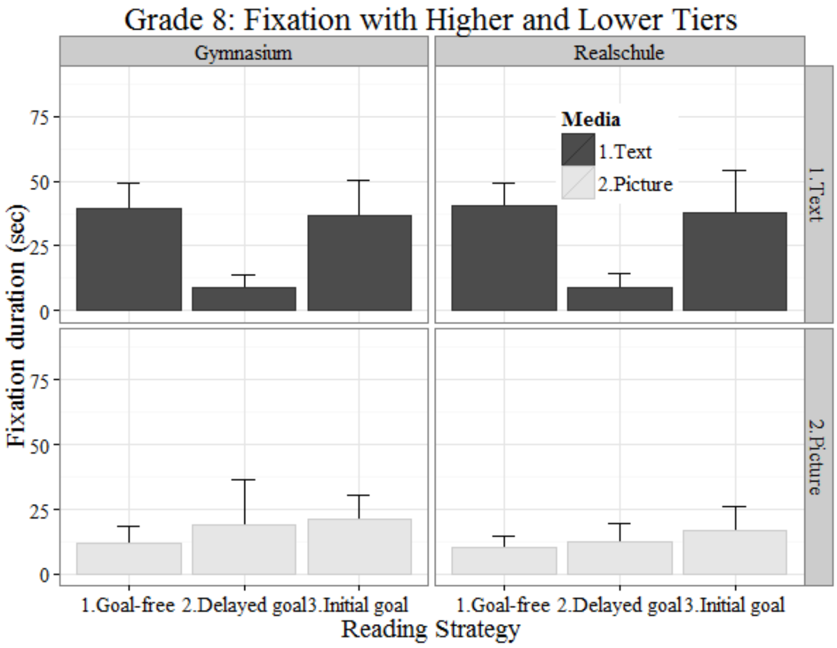


Figure 4.9: Fixation duration on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 8. Error bars indicate between subjects standard error of the mean.

(2) Time to the first fixation during text-picture integration

The 2 (× 2) ANOVA on latency of first fixation showed no difference in text processing between students from both school tiers (see Figure 4.10), $F < 1$, ns. Participants from all school tiers fixated on text quicker with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 59.79$, $p < .001$, $\eta_p^2 = .46$. No other effect was found: *school tier* × *task orientation*, $F < 1$, ns.

Likewise, higher tier students perceived pictures indistinguishably as lower tier students, $F < 1$, ns. They fixated on pictures quicker with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 8.17$, $p = .006$, $\eta_p^2 = .11$. No interaction was found: *school tier* × *task orientation*, $F(1,70) = 2.83$, $p = .10$, $\eta_p^2 = .04$.

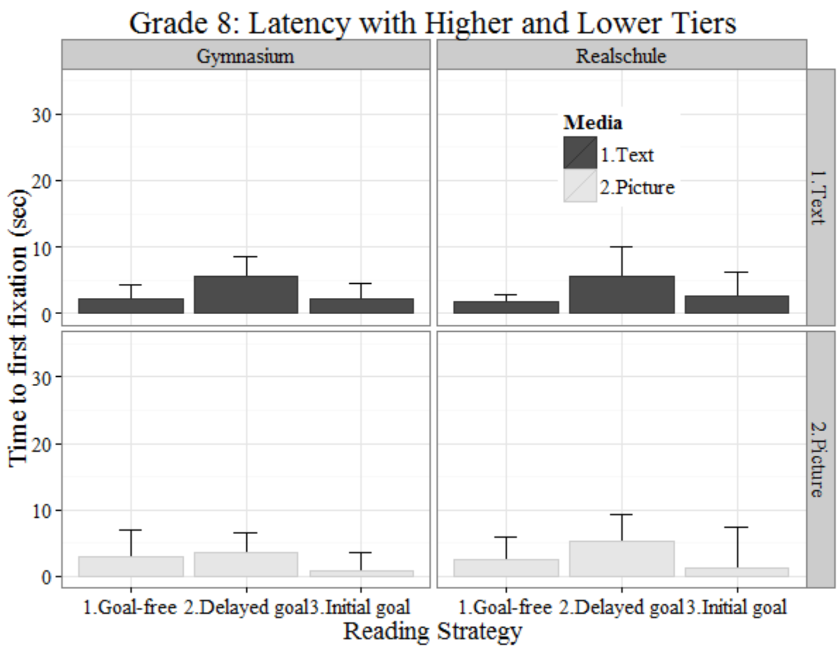


Figure 4.10: Time to the first fixation on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 8. Error bars indicate between subjects standard error of the mean.

(3) Number of transitions during text-picture integration

Higher tier students had significantly more transitions between text and pictures than lower tier students (see Figure 4.11), $F(1,70) = 11.64$, $p = .001$, $\eta_p^2 = .14$. All participants had much more transitions between text and pictures with goal-free processing than with delayed goal-oriented processing, $F(1,70) = 103.75$, $p < .001$, $\eta_p^2 = .60$. No other effect was found: *school tier* × *task orientation*, $F(1,70) = 2.62$, $p = .11$, $\eta_p^2 = .04$.

With delayed goal-oriented processing, higher tier students differed significantly from lower tier students, $F(1,70) = 76.88$, $p < .001$, $\eta_p^2 = .52$. Higher tier students had 31.6% (SD = 8.6%) transitions between text and pictures, 21.7% (SD = 7.6%) transitions between text and question, 46.7% (SD = 10.6%) transitions between pictures and question with delayed goal-oriented processing. Lower tier students had 28.45% (SD = 8.8%) transitions between text and pictures, 23.6% (SD = 11.1%) transitions between text and question, 47.9% (SD = 10.9%) transitions between pictures and question.

4.2.3.2 Goal-free processing vs initial goal-oriented processing

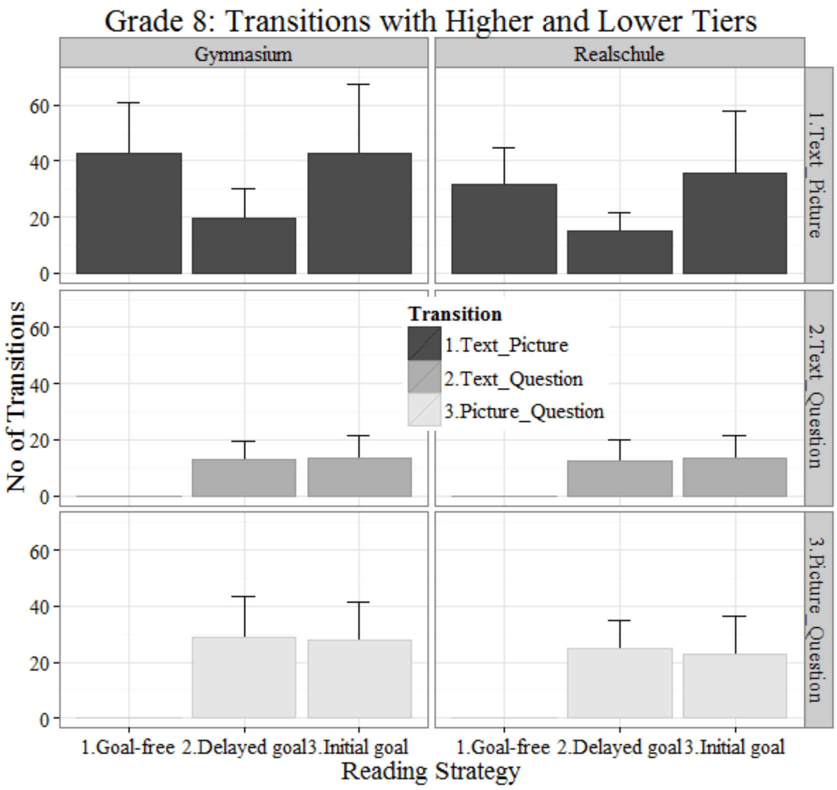


Figure 4.11: Transitions between text, pictures and question with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 8. Error bars indicate between subjects standard error of the mean.

Text was fixated similarly between higher and lower tier students with goal-free (no question yet) and initial goal-oriented processing (question before material). Conversely, pictures were emphasised by higher tier students rather than lower tier students. Participants from both school tiers perceived text and pictures roughly at the same time. Higher tier students had generally more transitions between text and pictures than lower tier students. With initial goal-oriented processing, no difference was revealed between two school tiers.

(1) Fixation duration during text-picture integration

The 2 (× 2) ANOVA on accumulated viewing time on text disclosed no effect of school tier, $F < 1$, ns; task orientation, $F(1,70) = 2.90$, $p = .09$, $\eta_p^2 = .04$; *school tier* × *task orientation*, $F < 1$, ns.

In contrast, higher tier students fixated longer on pictures than lower tier students, $F(1,70) = 5.06$, $p = .03$, $\eta_p^2 = .07$. All the students focused longer on pictures with goal-free than initial goal-oriented processing, $F(1,70) = 46.90$, $p < .001$, $\eta_p^2 = .40$. No effect was detected: *school tier* × *task orientation*, $F(1,70) < 1$, ns.

(2) Time to the first fixation during text-picture integration

The 2 (× 2) ANOVA on latency of first fixation on text displayed no effect of school tier, $F < 1$, ns; task orientation, $F(1,70) = 1.12$, $p = .29$, $\eta_p^2 = .02$; *school tier* × *task orientation*, $F(1,70) = 2.17$, $p = .15$, $\eta_p^2 = .03$.

Higher tier students had similar latency of first fixation on pictures with lower tier students, $F < 1$, ns. All students fixated slower on pictures with goal-free processing than with initial goal-oriented processing, $F(1,70) = 5.38$, $p = .02$, $\eta_p^2 = .07$. No interaction was found: *school tier* × *task orientation*, $F < 1$, ns.

(3) Number of transitions during text-picture integration

Higher tier students had significantly more transitions between text and pictures than lower tier students, $F(1,70) = 5.59$, $p = .02$, $\eta_p^2 = .07$. No other effects were disclosed: task orientation, $F < 1$, ns; *school tier x task orientation*, $F < 1$, ns.

With initial goal-oriented processing, higher tier students had similar transition patterns as lower tier students, $F(1,70) = 1.96$, $p = .17$, $\eta_p^2 = .03$. Higher tier students had 49.8% (SD = 13.5%) transitions between text and pictures, 16.1% (SD = 7.0%) transitions between text and question, 34.1% (SD = 11.8%) transitions between pictures and question. Lower tier students had 48.0% (SD = 12.7%) transitions between text and pictures, 18.8% (SD = 8.5%) transitions between text and question, 33.2% (SD = 11.9%) transitions between pictures and question.

4.3 Discussion

This experiment mainly investigated the differences between text processing and picture processing in posing question before and after the material among 8th graders. As suggested by Rickards and Di Vesta (1974), posing question after the material leads to processing the material in a general and coherent manner. Posing question before the material elicits specifically processing of the details in the material. To examine the effects of task orientation, Experiment 2 compared the accuracy and eye movements between (1) goal-free (no question yet) vs. delayed goal-oriented processing (question after material) and (2) goal-free vs. initial goal-oriented processing (question before material).

Parallel to Experiment 1, Experiment 2 adopted eye tracking method in order to explore whether the factors like question difficulty and school tiers influence the way of text processing and picture processing among 8th graders. Participants are hypothesised to mainly fixate on text when no question is displayed. Conversely, they probably emphasize on pictures when they answer the question after prior experience with the material. Their attention on text and pictures may be more balanced although text is still mainly focused when question is presented first. Participants are expected to mainly use text for easy questions; whereas, they are presumed to primarily use pictures for difficult questions. In addition, higher tier students should put more emphasis on pictures rather than text than lower tier students. The present results are consistent with the prior reports and suggestions that text processing differentiates fundamentally from picture processing (Mayer, 2014a; Schnotz, Mengelkamp, Baadte, and Hauck, 2014). In particular, the present results are in line with the hypotheses concerning the effective factors like task orientation, question difficulty and school tier.

4.3.1 Text guides general reading; pictures guide selective reading

This experiment confirms the assumption that text processing differs fundamentally from picture processing (Schnotz, 2014). Furthermore, it reveals the robust difference between text and pictures in different reading strategies. Eye tracking indicators were analyzed based on eye-mind hypothesis and immediacy hypothesis. The results were consistent with the assumptions on accuracy and eye movements with goal-free (no question yet), delayed goal-oriented processing (question after material) and initial goal-oriented processing (question before material).

For goal-free processing, text was mainly fixated rather than pictures. As fixation duration indicates the depth of processing and distribution of attention (Snowden, Thompson, and Troscianko, 2012), it can be interpreted that participants mainly used text instead of pictures to comprehend the information. It also implies, to some extent, that text processing is not identical to picture processing when no question is shown. The latency of first fixation showed that text was perceived sooner than pictures in goal-free processing. The delayed time to the first fixation on text can be interpreted as the interest of the information (Wedel and Pieters, 2008). It indicates that text is more interesting than pictures when no task is displayed. It seems that participants try to build the initial mental model primarily with the help of text and slightly with pictures. Besides, they had frequent transitions between text and pictures with goal-free processing. There was a significant correlation between correctness and transitions between text and pictures in goal-free processing. It can be interpreted that participants who answered the question correctly transferred their attention frequently between text and pictures. Correspond to eye-mind hypothesis, this significant correlation reveals that number of transitions can indeed reflect participant's cognitive process. Number of transitions can imply the integrative degree of the information based on the Integrative Model of TPC. It can be interpreted that participants had higher accuracy when they attempted to integrate the verbal information and pictorial information at a deeper level. This finding is consistent with the suggestion that the integration of text and pictures contributes enormously to the acquirement of knowledge (Mason, Tornatora, and Pluchino, 2015). In brief, participants tend to comprehend text and pictures differently. Consistent with Johnson and Mayer (2012), they may build their initial mental model by integrating text and pictures. Text rather than pictures primarily guides readers to build the initial mental model.

For delayed goal-oriented processing, participants fixated mainly on pictures rather on text. As fixation duration indicates the depth of processing (Rayner, 1998), It can be interpreted that participants had the deep processing of pictures in delayed goal-oriented processing. Participants appeared to change their emphasis from text to pictures when they needed to answer the question after the construction of initial mental model. It is correspond to the idea that pictures can be used to scaffold question answering (Eitel, Scheiter, Schüler, Nyström, and Holmqvist, 2013). Pictures were also perceived sooner than text in delayed goal-oriented processing. It indicates that participants were more interested in pictures than text. This supports the assumption that pictures' specific function on scaffolding information to enhance mental model construction. With delayed goal-oriented processing, participants had the most transitions between pictures and question, more transitions between text and pictures and the least between text and question. According to the TRACE model (Rouet, 2006), participants build task model after understanding the question. Participants tended to predominately use pictures to scaffold question answering. They also transferred slightly between text and question, which can be interpreted as updating the mental model to meet the criterior of the task model. More importantly, there was a strong correlation between accuracy and transitions between text and pictures in delayed goal-oriented processing. Although eye tracking indicators cannot directly represent what happens in reader's mind, they are useful tools for us to understand how the cognitive process works. Correspond to Vandeberg, Bouwmeester, Bocanegra, and Zwaan (2013), transition, as an eye tracking indicator, indeed represents the integration of text and pictures in mental model construction. In short, pictures can guide and scaffold the usage of mental model.

For initial goal-oriented processing, participants focused mainly on text but also paid a large amount of attention on pictures. It can be explained as they had deep processing with text and pictures after knowing the question. Text was mainly used to construct mental model and pictures were primarily used to answer the question. Participants perceived pictures sooner than text with initial goal-oriented processing. It can be interpreted as their interest in pictures to scaffold question answering. They had the most frequent transitions between text and pictures, more transitions between pictures and question and the least between text and question. They integrated mostly text and pictures because they still needed to build the mental model after comprehending the question. However, this mental model is rather specific as they know the task already before reading the materials. The frequent transitions between pictures and question confirm the advantages of pictures as tools to scaffold question answering and to facilitate mental model construction. The unparalleled transitions among text, pictures and question are in accordance with the fundamental difference between text processing and picture processing (Schnotz and Bannert, 2003). Finally, the robust correlation between accuracy and transitions between text and pictures indicates transitions can indeed reveal the integration of text and pictures. Text is mainly used to construct the mental model, whereas pictures are primarily used to scaffold question answering.

4.3.2 Pictures are more emphasized than text with difficult question

Participants focused on text and on pictures longer as the question difficulty increased. Conversely, they viewed pictures much longer with difficult questions than with easy questions. As fixation indicates the depth of processing and distribution of attention, text was processed at a deeper level with difficult questions than with easy questions. Picture was predominantly used with difficult questions compared to easy questions. It is in agreement with the proposition that pictures can support the comprehension of difficult questions (Schnotz, Ludewig, et al., 2014). Differed from text, pictures can clearly and precisely present the equivalent textual information (Carney and J. R. Levin, 2002; Fletcher and Tobias, 2005). Therefore, pictures can minimize the cognitive load that is required in complicated reasoning tasks (Marcus, Cooper, and Sweller, 1996). In other words, the emphasis on pictures with difficult questions can also due to the superiority of pictorial or depictive representations. It provides clear structure of the content and easy access to the statement. Besides, a strong negative correlation was only found between correctness and eye tracking indicator with goal-free processing when question was at easy level. As suggested from the Integrative Model of TPC (Schnotz, Ludewig, et al., 2014), participants transformed the verbal and pictorial information into a mental model with the goal-free processing. The result showed that participants tended to have higher accuracy with easy question when they emphasized more on pictures during mental model construction. It suggests that pictures can be served to guide the construction of mental model (Gyselinck, Jamet, and Dubois, 2008).

No difference was found in the latency of first fixation on text with easy question and difficult question. In contrast, participants fixated pictures quicker with easy question than with difficult question. It indicates that participants had similar interest on text with easy and difficult questions. Conversely, pictures were more interesting for participants with easy questions than with difficult questions. Apparently, participants realized that pictures can be served as a more visualized tool than text to scaffold question answering (Kieras and Bovair, 1984). They may make a quicker decision to use pictures when the question is easier. Consistent with the hypothesis, text was similarly emphasised regardless of question difficulty, whereas pictures were primarily focused when the question was at a difficult level. In other words, pictures can facilitate mental model construction especially for the difficult questions.

4.3.3 Pictures seem to be the key to distinguish between good vs poorer problem solvers

Corresponding to the hypothesis, higher tier students had higher accuracy than lower tier students. Higher tier students (Gymnasium) viewed text similarly but fixated significantly longer on pictures compared to lower school tier students (Realschule). It indicates that higher school students distributed their attention similarly on text as lower tier students. Conversely, higher school students had deeper processing on pictures than lower tier students. In line with Schnotz, Ludewig, et al. (2014), advanced readers were more skilful in comprehending the pictorial information and in applying the strategy of using pictures to scaffold question answering. No difference was detected in time to the first fixation on text and on pictures. It can be explained that participants from higher and lower school tiers had relatively same interest to text and to pictures.

Transition data showed that higher tier students had more transitions between text and pictures than lower tier students. As transitions indicates the integration of text and pictures (Johnson and Mayer, 2012), it can be interpreted that advanced readers integrated text and pictures more frequently than poorer readers. Consistent with the suggestion proposed by Cerdán et al. (2011), poorer readers have difficulty not only in decoding and encoding the verbal and pictorial information but also in integrating the information from both media. With delayed goal-oriented processing, higher tier students had more frequent transitions between text and pictures than lower tier students. Additionally, higher tier students had better accuracy than lower tier students. This supports the assumption that advanced readers are more capable of integrating text and pictures than poorer readers (Hochpöchler et al., 2013). Higher tier students had more transitions between text and question and pictures and question than lower tier students. As lower tier students had lower accuracy than higher tier students, it suggests that more transitions lead to better locating the useful information to answer the question in text and scaffolding the information in pictures. In addition, it suggests that lower tier students did not construct the mental model good enough to meet the requirements from the task model, which was created from the question. Higher tier students switched their attention frequently among text, pictures and question to update more information from text and pictures so that the question can be answered correctly.

4.3.4 Gender and types of illustration issues

As mentioned in 3.3.4, gender was not balanced in this study due to the difficulty of recruitment. However, gender can influence reading processing of science material (Koedinger and J. R. Anderson, 1998). Besides, it is not clear which type of graphics can, to a large extent, enhance learning (Tversky et al., 2000).

4.3.5 Emphasis on picture comprehension in curriculum

The results from 8th graders show that text processing differs from picture processing in the perspectives of task orientation, question difficulty and school tier. Text is primarily used for mental model construction, which can guide general reading. In contrast, pictures are mainly used for scaffolding question answering, which can guide selective reading. Picture rather than text can be used more frequently as question difficulty increases. Higher and lower tier students had similar performance in text processing; however, higher tier students seem to be more skillful in picture processing than lower tier students. Owing to the focus on text training at school, the gap between higher and lower tiers nearly disappears on text processing. However, the higher tier students distinguish from lower tier students on picture processing. Likely, picture comprehension is the essence of obtaining correct answers. As text comprehension is the focus in science curriculum (Kultusministerkonferenz, 2005b), this study implies the significance and necessity of training on picture comprehension in science education.

Chapter 5

Experiment 1 vs 2

Experiment 1 and Experiment 2 are compared to investigate whether students from 5^{th} grade and 8^{th} grade differ in multimedia learning. According to Hypothesis 4, 8^{th} graders should perform better in accuracy and should be more skillful in comprehending multimedia presentations (text vs. pictures) than 5^{th} graders. Compared to 5^{th} graders, 8^{th} graders may process the verbal information more efficiently and use the pictorial information more intensively. Results on accuracy were consistent with the hypothesis. Inconsistent with the hypothesis, 8^{th} graders fixated considerably shorter on text than 5^{th} graders. However, 8^{th} graders had similar fixation patterns on pictures compared to 5^{th} graders.

5.1 Results

According to the hypotheses, text is assumed to be mainly used for general reading. Pictures are presumed to be mainly used for selective reading. When question is at easy level, readers are expected to focuse mainly on text. When question is at difficult level, they are likely to emphasize mainly on pictures. Furthermore, higher tier students (Gymnasium) may process text more efficiently than lower tier students (Realschule). Conversely, higher tier students may fixate longer on pictures than lower tier students. Finally, higher graders (8^{th} graders) are hypothesized to process text more efficiently than lower graders (5^{th} graders). Conversely, higher graders may emphasized more on pictures than lower graders. Two comparisons were performed as there are two processing containing higher proportion of goal-oriented processing. (1) Goal-free processing (no question yet) vs. delayed goal-oriented processing (question after material). (2) Goal-free processing vs. initial goal-oriented processing (question before material).

The results revealed similar fixation trends between 5^{th} graders and 8^{th} graders in the perspective of task orientation, question difficulty and school tier. In general, text processing differed significantly from picture processing. Eighth graders processed text more efficiently than 5^{th} graders. Conversely, no difference was shown between 5^{th} graders and 8^{th} graders. In agreement with Hypothesis 1, text mainly guided general reading; whereas, pictures mainly guided selective reading. Inconsistent with Hypothesis 2, attention on text and pictures increased when question difficulty increased. The increase of attention on pictures was more than on text when question was difficult. Different from Hypothesis 3, higher tier students did not differ in text processing compared to lower tier students. However, higher tier students tended to more focused on pictures than lower tier students. Higher graders processed text more efficiently than lower graders. Conversely, higher graders had similar fixation patterns as lower graders, which did not accord with Hypothesis 4.

5.1.1 Question 1: How do readers deal with text and deal with pictures in different forms of reading?

Accuracy and eye tracking data from 5^{th} graders and 8^{th} graders were compared in the perspectives of task orientation, question difficulty and school tier. The trends of eye fixation patterns were compared: (1) Goal-free vs. delayed goal-oriented processing; (2) goal-free vs. initial goal-oriented processing. Consistent with the hypothesis, 5^{th} graders (M = 45%, SD = 20%) performed significantly worse than participants from 8^{th} graders (M = 65%, SD = 19%), $F(1,142) = 38.73$, $p < .001$, ${\eta_p}^2 = .21$.

With regard to Hypothesis 1, participants possibly have similar accuracy when question is posed before and after the material. If reading is without any question (i.e. only text and pictures were displayed), participants may emphasize primarily on text. If participants were required to answer questions, they may focus mainly on pictures. The data from both grades were in agreement with the hypothesis on task orientation. Participants had similar accuracy with posing questions after and before the materials. Participants paid attention predominately to text to understand the materials, whereas they focused mainly on pictures to scaffold question answering.

The accuracy data were in conformity with the hypothesis. No difference was found among 5^{th} graders with posing question after the material (M = 45%, SD = 31%) and before the material (M = 45%, SD = 29%), $F < 1$, ns. Similarly, 8^{th} had similar accuracy with placing question after the material (M = 68%, SD = 27%) and before the material (M = 63%, SD = 25%), $F(1,71) = 1.50$, $p = .23$, ${\eta_p}^2 = .02$. The trends of the eye movements will be compared with accumulated fixation duration, time to the first fixation and number of transitions.

5.1.1.1 Goal-free processing vs delayed goal-oriented processing

Fixation data showed similar trend among 5^{th} and 8^{th} graders on text and picture processing with goal-free processing (no question yet) and with delayed goal-oriented processing (question after material). All participants fixated significantly longer on text with goal-free processing than with delayed goal-oriented processing. In contrast, they fixated considerably longer on pictures with delayed goal-oriented processing than with goal-free processing. Eighth graders had shorter fixation on text but similar fixation on pictures than 5^{th} graders. All participants fixated text quicker than pictures with goal-free processing. Conversely, they fixated pictures quicker than text with delayed goal-oriented processing. Eighth graders had shorter fixation on text but similar fixation on pictures compared to 5^{th} graders. Participants from both grades transferred their attention most frequently between text and pictures with goal-free processing.

(1) Accumulated fixation duration during text-picture integration

Figure 5.1 illustrated that 8th graders fixated much quicker on text than 5th graders with goal-free processing. Conversely, 8th graders focused slightly longer on pictures than 5th graders. Fifth graders paid more attention to text and less attention to pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 341.85$, $p < .001$, ${\eta_p}^2 = .83$ for text processing; $F(1,71) = 7.20$, $p = .009$, ${\eta_p}^2 = .09$ for picture processing (see Figure 4.1). Eighth graders had the same trend as 5th graders: $F(1,71) = 661.42$, $p < .001$, ${\eta_p}^2 = .90$ for text processing; $F(1,71) = 6.74$, $p = .01$, ${\eta_p}^2 = .09$ for picture processing.

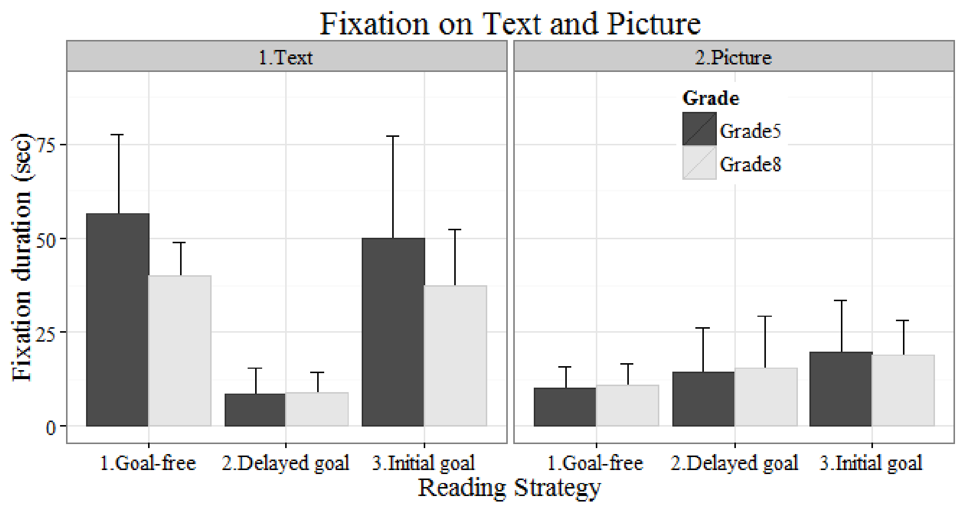


Figure 5.1: Accumulated fixation duration on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders and 8th graders. Error bars indicate between subjects standard error of the mean.

(2) Time to the first fixation during text-picture integration

Eight graders perceived text and pictures sooner than 5th graders with goal-free and delayed goal-oriented processing (see Figure 5.2). Fifth graders fixated significantly quicker on text and marginally slower on pictures with goal-free processing than with delayed goal-oriented processing: $F(1,71) = 25.33$, $p < .001$, ${\eta_p}^2 = .26$ for text processing; $F(1,71) = 3.34$, $p = .07$, $\eta_p{}^2 = .05$ for picture processing. Eighth graders had the same patterns as 5th graders, $F(1,71) = 60.15$, $p < .001$, $\eta_p{}^2 = .46$ for text processing; $F(1,71) = 7.97$, $p = .006$, $\eta_p{}^2 = .10$ for picture processing.

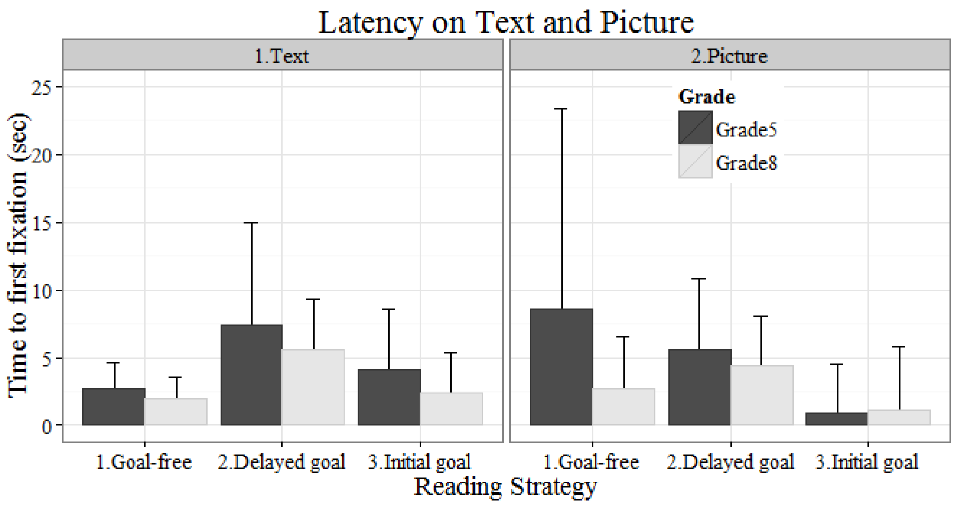


Figure 5.2: Time to the first fixation on text and on pictures in goal-free, delayed goal-oriented and initial goal-oriented processing among 5th graders and 8th graders. Error bars indicate between subjects standard error of the mean.

(3) Number of transitions during text-picture integration

Eighth graders had more frequent transitions among text, pictures and question than 5th graders with goal-free and delayed goal-oriented processing 5.3. Likewise, all participants had significantly more transitions between text and pictures with goal-free processing than with delayed goal-oriented processing, $F(1,71) = 47.98$, $p < .001$, $\eta_p{}^2 = .40$ for 8th graders; $F(1,71) = 101.43$, $p < .001$, $\eta_p{}^2 = .59$ for 8th graders. With delayed goal-oriented processing, they had the most proportion of transitions between pictures and question, 53.0% ($SD = 14.8\%$) for 5th graders; 52.4% ($SD = 2.0\%$) for 8th graders.

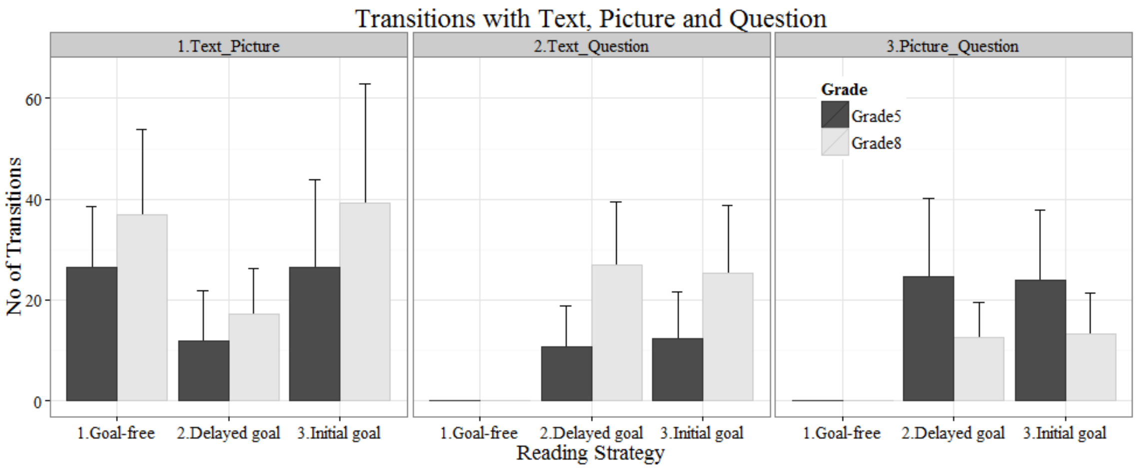


Figure 5.3: Transitions among text, pictures and question in goal-free, delayed goal-oriented and initial goal-oriented processing among 5^{th} graders and 8^{th} graders. No transitions are illustrated between text and question, pictures and question because goal-free processing does not deal with questions. Error bars indicate between subjects standard error of the mean.

5.1.1.2 Goal-free processing vs initial goal-oriented processing

A slight difference was revealed in multimedia comprehension with goal-free (no question yet) and initial goal-oriented processing (question before material). Eighth graders had shorter fixation on text than 5^{th} graders with initial goal-oriented processing. Participants from both grades had similar trends in picture processing with initial-oriented processing. They also fixated quicker on text and pictures with initial goal-oriented processing than with goal-free processing. Eighth graders fixated quicker on text but focused similarly on pictures than 5^{th} graders with initial goal-oriented processing. Participants from both grades had the most transitions between text and pictures with initial goal-oriented processing. Eighth graders had more transitions among text, pictures and question than 5^{th} graders with initial goal-oriented processing.

(1) Fixation duration during text-picture integration

Eighth graders fixated on text quicker than 5^{th} graders with initial goal-oriented processing (see Figure 5.1). Conversely, all participants had similar fixation patterns on pictures with initial goal-oriented processing. Fifth graders viewed longer on text and shorter on pictures with goal-free processing and with initial goal-oriented processing: $F(1,71) = 6.43$, $p = .01$, ${\eta_p}^2 = .08$ for text processing; $F(1,71) = 33.48$, $p < .001$, ${\eta_p}^2 = .32$ for picture processing. However, 8th graders fixated similarly on text but fixated significantly shorter on pictures with goal-free processing than with initial goal-oriented processing: $F(1,71) = 2.94$, $p = .09$, ${\eta_p}^2 = .04$ for text processing; $F(1,71) = 47.03$, $p < .001$, ${\eta_p}^2 = .40$ for picture processing.

(2) Time to the first fixation during text-picture integration

Eighth graders fixated quicker on text than 5th graders with initial goal-oriented processing (see Figure 5.2). In contrast, 8th graders had similar latency to fixate on pictures compared to 5th graders. Fifth graders fixated quicker on text but slower on pictures with goal-free processing than with initial goal-oriented processing: $F(1,71) = 6.15$, $p = .02$, ${\eta_p}^2 = .08$ for text processing, $F(1,70) = 18.86$, $p < .001$, ${\eta_p}^2 = .21$ for picture processing. In contrast, 8th graders had similar latency of first fixation on text but they fixated slower on pictures with goal-free processing than with initial goal-oriented processing: $F(1,71) = 1.10$, $p = .30$, ${\eta_p}^2 = .02$ for text processing, $F(1,70) = 5.44$, $p = .02$, ${\eta_p}^2 = .07$ for picture processing.

(3) Number of transitions during text-picture integration

Eighth graders transferred their attention more frequently among text, pictures and question than 5th graders with initial goal-oriented processing (see Figure 5.3). Participants from both grades had a rather similar transitions between text and pictures with goal-free processing and initial goal-oriented processing, F < 1, ns for both graders. With initial goal-oriented processing, 5th graders had more proportion of transitions between text and pictures, 40.5% (SD = 14.9%) for 5th graders; 48.9% (SD = 13.1%) for 8th graders.

5.1.2 Question 2: Does preference of readers for text vs. pictures depend on question difficulty?

For question difficulty, participants are presumed to have more correct answers with easy questions (level 1: processing of value) than difficult questions (level 2/3: processing of relation). With easy questions, participants are expected to fixate longer on text than on pictures. With difficult questions, they are predicted to focus more on pictures than on text. Participants from both grades show better accuracy with easy questions than with difficult questions. However, 5^{th} graders tend to emphasize pictures rather than text with difficult questions than with easy questions. Eighth graders seem to emphasize both text and pictures with difficult questions than with easy questions. All participants had remarkably better scores with easy questions than with difficult questions, $F(1,71) = 22.74$, $p < .001$, ${\eta_p}^2 = .24$ for 5^{th} graders; $F(1,71) = 48.33$, $p < .001$, ${\eta_p}^2 = .41$ for 8^{th} graders. The trends of eye movements were compared with the accumulated fixation duration and the latency of first fixation.

5.1.2.1 Goal-free processing vs delayed goal-oriented processing

Participants from both grades differed significantly in processing multimedia presentations with easy and difficult questions with goal-free processing (no question yet) and with delayed goal-oriented processing (question after material). Fifth graders fixated similarly on text but significantly longer on pictures with difficult questions than with easy questions. Eighth graders viewed longer on text and pictures with difficult questions than with easy questions. Fifth graders had similar latency of first fixation on text but had shorter latency of first fixation on pictures with easy questions than with difficult questions. Participants from Grade 5 and Grade 8 had similar latency of first fixation on text and on pictures with easy and difficult questions.

(1) Accumulated fixation duration during text-picture integration

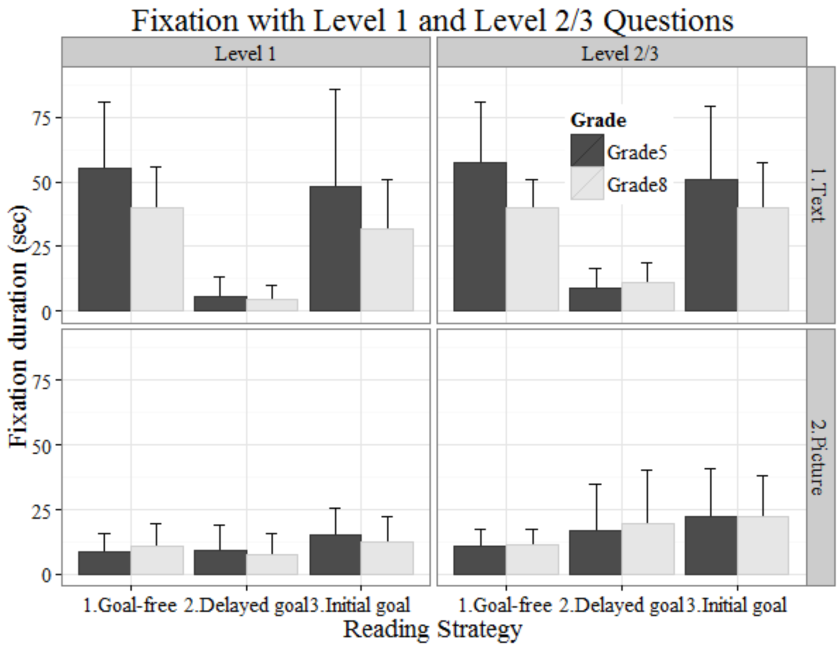


Figure 5.4: Fixation duration on text and on pictures with goal-free and delayed goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 5th graders and 8th graders. Goal-free processing is displayed as baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

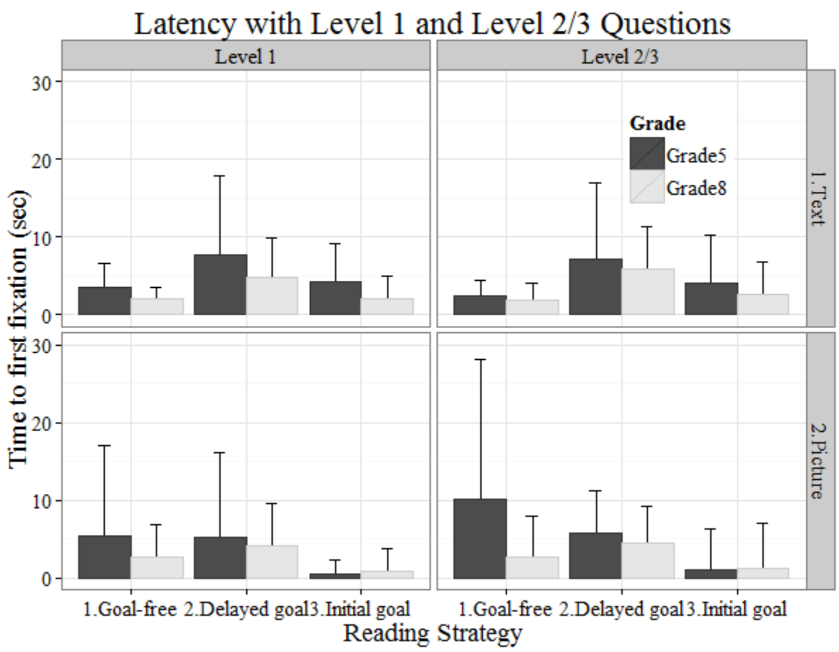


Figure 5.5: Time to the first fixation on text and on pictures with goal-free and delayed goal-oriented processing with level 1 (easy) and level 2/3 (difficult) questions among 5th graders and 8th graders. Goal-free processing is displayed as baseline as no question is shown in this processing. Error bars indicate between subjects standard error of the mean.

With easy question (level 1), 5th graders fixated longer on text and pictures than 8th graders (see Figure 5.4). With difficult question (level 2/3), 8th graders focused longer on text and pictures than 5th graders. Fifth graders had similar fixation on text and significantly longer fixation on pictures with difficult questions than with easy questions, $F(1,71) = 2.31$, $p = .13$, ${\eta_p}^2 = .03$ for text processing; $F(1,71) = 14.16$, $p < .001$, ${\eta_p}^2 = .17$ for picture processing. However, 8th graders fixated significantly longer on text and on pictures with difficult questions than with easy questions, $F(1,71) = 6.79$, $p = .01$, ${\eta_p}^2 = .09$ for text processing; $F(1,71) = 17.50$, $p < .001$, ${\eta_p}^2 = .20$ for picture processing.

(2) Time to the first fixation during text-picture integration

With easy and difficult questions, 8th graders perceived text and pictures sooner than 5th graders (see Figure 5.5). Fifth graders and 8th graders had similar latency of first fixation on text with easy and difficult questions, $F(1,71) = 1.15$, $p = .29$, ${\eta_p}^2 = .02$ for 5th graders; $F(1,71) = 1.32$, $p = .25$, ${\eta_p}^2 = .02$ for 8th graders. 5th graders fixated on pictures quicker with easy questions than with difficult questions, $F(1,71) = 6.82$, $p = .01$, ${\eta_p}^2 = .09$. However, 8th graders fixated similarly on pictures with easy and difficult questions, $F < 1$, ns.

5.1.2.2 Goal-free processing vs initial goal-oriented processing

For question difficulty, 5th graders and 8th graders differed significantly in processing multimedia presentations with goal-free processing (no question yet) and with initial goal-oriented processing (question before material). Fifth graders focused similarly on text but emphasised on pictures with difficult question compared to easy question. Conversely, 8th graders focused longer on text and on pictures with difficult question than with easy question. Lower and 8th graders had similar latency of first fixation on text but had shorter latency of first fixation on pictures with easy question than with difficult question.

(1) Accumulated fixation duration during text-picture integration

With easy question, 5^{th} graders focused longer on text and on pictures than 8^{th} graders (see Figure 5.4). With difficult question, 5^{th} graders fixated longer on text but similarly on pictures compared to 8^{th} graders. Fifth graders viewed similarly on text but significantly longer on pictures with difficult questions than with easy questions, F < 1, ns for text processing; $F(1,71) = 13.69$, $p < .001$, ${\eta_p}^2 = .16$ for picture processing. However, 8^{th} graders fixated marginally longer on text and significantly longer on pictures with difficult questions than with easy questions, $F(1,71) = 3.46$, $p = .07$, ${\eta_p}^2 = .05$ for text processing; $F(1,71) = 21.19$, $p < .001$, ${\eta_p}^2 = .23$ for picture processing.

(2) Time to the first fixation during text-picture integration

Fifth graders had similar latency of first fixation on text but had shorter latency of first fixation on pictures with easy questions than with difficult questions (see Figure 5.5), $F(1,71) = 2.06$, $p = .16$, ${\eta_p}^2 = .03$ for text processing; $F(1,71) = 9.63$, $p = .003$, ${\eta_p}^2 = .02$ for picture processing. Eighth graders had similar latency of first fixation on text with easy and difficulty questions, $F < 1$, ns. They fixated remarkably longer on pictures with difficult questions than with easy questions, $F(1,71) = 7.49$, $p = .008$, ${\eta_p}^2 = .10$. Regardless of question difficulty, 8^{th} graders perceived text and pictures sooner than 5^{th} graders.

5.1.3 Question 3: What are the differences in text processing and picture processing between students from higher and lower school tiers?

Eighth graders are predicted to perform better in accuracy than 5^{th} graders. They are presumed to process text quicker than 5^{th} graders. Besides, they are expected to invest more time on pictures than 5th graders. For both grades, students from higher tier (Gymnasium) had more accurate answers than students from lower tier (Realschule). The eye movements also showed the same pattern for students from both grades. Students from higher tier fixated shorter on text but longer on pictures than students from lower tier.

For 5th graders, students from higher tiers (M = 50%, SD = 22%) had higher percentage of accuracy than lower tier students (M = 40%, SD = 18%), $F(1,70) = 4.79$, $p = .03$, $\eta_p{}^2 = .06$. For 8th graders, students from higher tiers (M = 72%, SD = 17%) also performed better than students from lower tiers (M = 59%, SD = 19%), $F(1,70) = 10.10$, $p = .002$, $\eta_p{}^2 = .13$. The trends of eye movements were compared with the accumulated fixation duration and the latency of first fixation.

5.1.3.1 Goal-free processing vs delayed goal-oriented processing

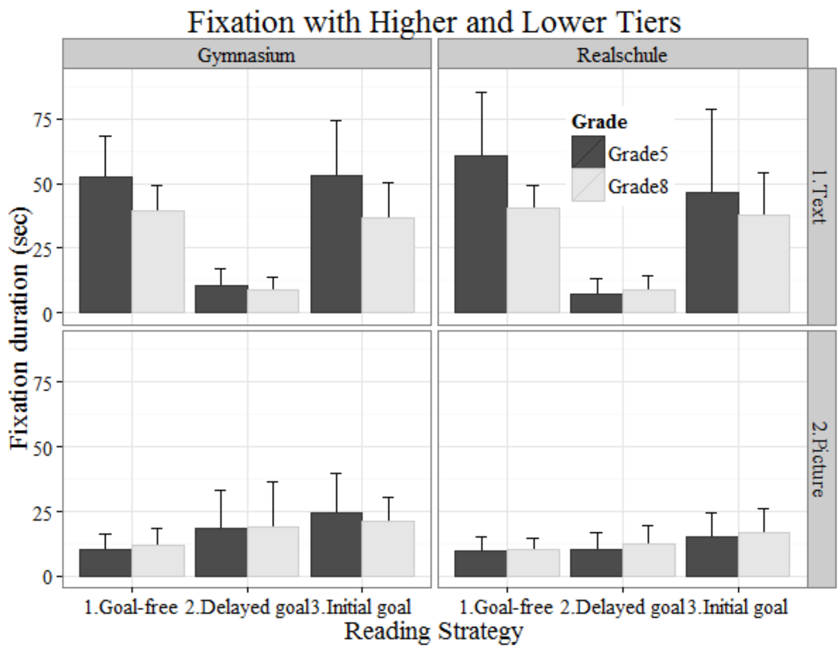


Figure 5.6: Fixation duration on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5 and Grade 8. Error bars indicate between subjects standard error of the mean.

Higher and lower tier students from Grade 5 and Grade 8 differed in text processing and picture processing with goal-free (no question yet) and delayed goal-oriented processing (question after material). Eighth graders focused on text shorter than 5th graders with goal-free processing. No difference was found in text processing between 5th graders and 8th graders with delayed goal-oriented processing. Regardless of grades, higher tier students had similar fixation patterns on text but emphasized pictures compared to lower tier students. Eighth graders perceived text and pictures sooner than 5th graders. No other difference was revealed in latency to fixate on text and pictures between higher and lower tier students. Eighth graders transferred their attention more frequently between text and pictures than 5th graders. Fifth graders from higher tier showed similar transition patterns as lower tier students. Conversely, 8th graders from higher tier had more transitions than lower tier students. (1) Fixa-

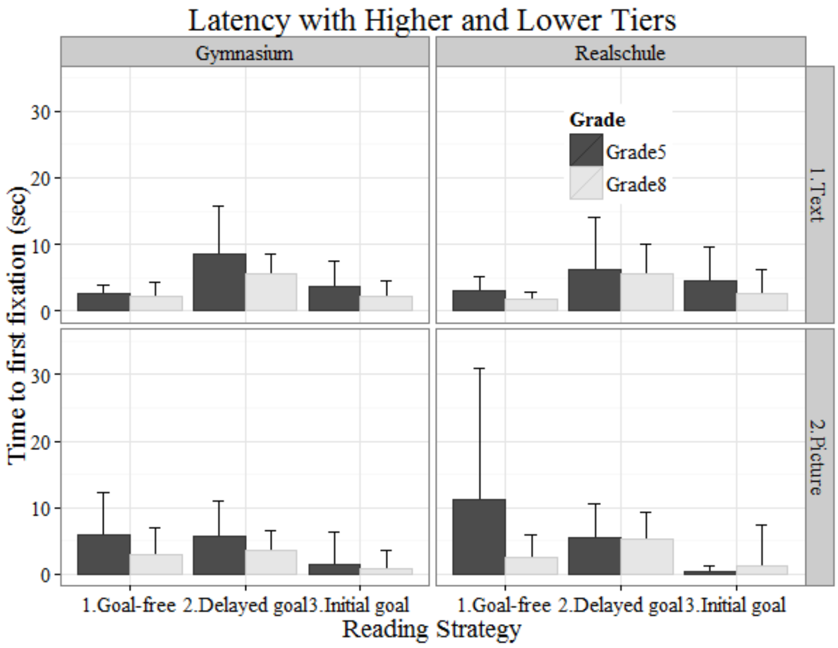


Figure 5.7: Time to the first fixation on text and on pictures with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5 and Grade 8. Error bars indicate between subjects standard error of the mean.

tion duration during text-picture integration

Eighth graders fixated on text shorter than 5th graders with goal-free processing (see Figure 5.6). No difference was found in fixation on text among 5th graders and

8th graders with delayed goal-oriented processing. Participants from both grades had similar fixation patterns on pictures with goal-free and delayed goal-oriented processing. For both grades, higher tier students had similar fixation duration on text as lower tier students, $F < 1$, ns. Higher tier students had significantly longer fixation on pictures than lower tier students, $F(1,70) = 11.51$, $p = .001$, ${\eta_p}^2 = .14$ for 5th graders; $F(1,70) = 6.56$, $p = .01$, ${\eta_p}^2 = .09$ for 8th graders.

(2) Time to the first fixation during text-picture integration

Eighth graders had shorter latency on text and pictures than 5th graders with goal-free and delayed goal-oriented processing (see Figure 5.7). There was also no difference in latency of first fixation on text between lower tier students and higher tier students: $F(1,70) = 1.02$, $p = .32$, ${\eta_p}^2 = .01$ for 5th graders; $F < 1$, ns for 8th graders. No difference was also found in latency of first fixation on pictures between lower tier and higher tier students: $F(1,70) = 1.61$, $p = .21$, ${\eta_p}^2 = .02$ for 5th graders; $F < 1$, ns for 8th graders. (3) Number of transitions during text-picture integration

As displayed in Figure 5.8, 8th graders had more transitions among text, pictures and question than 5th graders with goal-free and delayed goal-oriented processing. For 5th graders, higher tier and lower tier students had similar transitions between text and pictures, $F(1,70) = 2.92$, $p = .09$, ${\eta_p}^2 = .04$. Conversely, higher tier students had more transitions between text and pictures than lower tier students for 8th graders, $F(1,70) = 11.64$, $p = .001$, ${\eta_p}^2 = .14$. For both graders, higher and lower tier students had the most transitions between pictures and question with delayed goal-oriented processing, 56.1% ($SD = 13.6\%$) for 5th graders; 47.9% ($SD = 10.9\%$) for 8th graders.

5.1.3.2 Goal-free processing vs initial goal-oriented processing

Higher and lower tier students from Grade 5 and Grade 8 differed in fixation patterns on text and pictures initial goal-oriented processing (question before material). Eighth graders fixated shorter on text than 5th graders with initial goal-oriented processing.

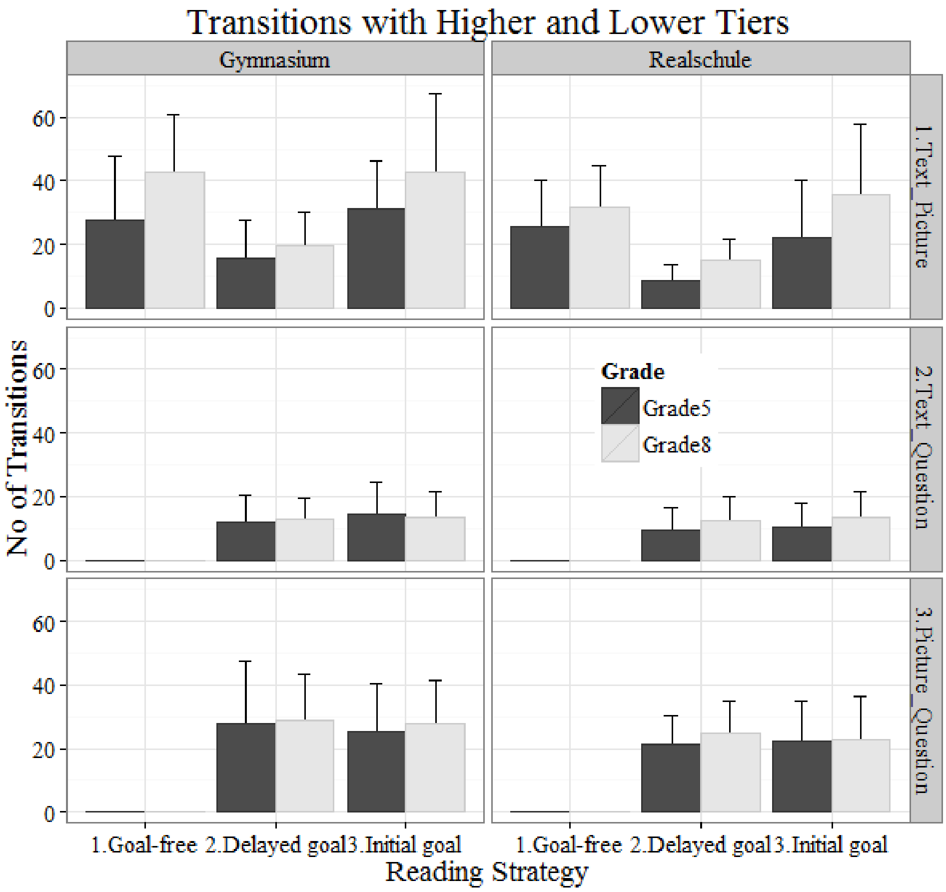


Figure 5.8: Transitions between text, pictures and question with goal-free, delayed goal-oriented and initial goal-oriented processing among students from Gymnasium and Realschule in Grade 5 and Grade 8. Error bars indicate between subjects standard error of the mean.

All participants had similar fixation on pictures. Regardless of grades, higher tier students fixated similarly on text but fixated longer on pictures compared to lower tier students. Eighth graders perceived text sooner but pictures similarly than 5^{th} graders. No difference was detected in latency to fixate on text and pictures between higher and lower tier students. Eighth graders transferred their attention more frequently between text and pictures than 5^{th} graders. For both grades, higher tier students transferred their attention more frequently than lower tier students.

(1) Accumulated fixation duration during text-picture integration

Figure 5.6 illustrated that 8^{th} graders fixated on text shorter than 5^{th} graders. Conversely, 8^{th} graders had similar fixation duration as 5^{th} graders. Participants from both school tiers in both grades fixated similarly on text, $F < 1$, ns for lower grade; $F < 1$, ns for higher grade. In both grades, higher tier students focused significantly longer on pictures than lower tier students, $F(1,70) = 8.62$, $p = .004$, ${\eta_p}^2 = .11$ for 5^{th} graders; $F(1,70) = 5.06$, $p = .03$, ${\eta_p}^2 = .07$ for 8^{th} graders.

(2) Time to the first fixation during text-picture integration

As shown in Figure 5.7, 8^{th} graders perceived text sooner than 5^{th} graders. In contrast, 8^{th} graders fixated on pictures similarly as 5^{th} graders. No difference was detected in delayed time of first fixation on text between students from higher and lower tier students, $F(1,70) = 1.19$, $p = .28$, ${\eta_p}^2 = .02$ for lower grade; $F < 1$, ns for higher grade. Lower and higher tier students also had no difference in latency of first fixation on pictures, $F(1,70) = 1.48$, $p = .23$, ${\eta_p}^2 = .02$ for lower grade; $F < 1$, ns for higher grade.

(3) Number of transitions during text-picture integration

Figure 5.8 showed that 8^{th} graders had more transitions among text, pictures and question than 5^{th} graders. Eighth graders had similar transitions between text and question, pictures and question than 5^{th} graders. Higher tier students transferred their attention more frequently between text and pictures than lower tier students,

$F(1,70) = 3.52$, $p = .07$, ${\eta_p}^2 = .05$ for 5th graders; $F(1,70) = 5.59$, $p = .02$, ${\eta_p}^2 = .07$. With initial goal-oriented processing, all the participants transferred their attention most frequently between text and pictures, 44.2% (SD = 12.9%) for 5th graders; 49.8% (SD = 13.5%) for 8th graders.

5.2 Discussion

Participants from Grade 8 were more efficient in text processing than participants from Grade 5. Conversely, no difference was detected in picture processing between participants from Grade 5 and Grade 8. Experiment 1 and Experiment 2 explored the difference in comprehending text and pictures between 5th graders and 8th graders. As mentioned in 3.3.4, 5th graders are not able to process text as automatic as 8th graders (Oakhill and Yuill, 1996). When 5th graders answered questions, their working memory may interfere with semantic recognition and required information location. Therefore, Experiment 1 vs 2 were compared in task orientation, question difficulty and school tier in order to examine the difference between students from 5th and 8th graders.

It is suggested that combining the data from accuracy and eye tracking together can better show the cognitive integration of text and pictures (Mason et al., 2014). Therefore, the trends of accuracy and eye movements were compared between 5th graders and 8th graders in goal-free (no question yet), delayed goal-oriented (question after material) and initial goal-oriented processing (question before material).

According to Hypothesis 4, participants from Grade 8 should have higher accuracy than participants from Grade 5. Eighth graders may fixate shorter on text than 5th graders. Conversely, 5th graders may invest more time on pictures than 5th graders. In conformity with the assumption on text processing, 8th graders viewed text generally shorter than 5th graders. However, no difference was detected on picture processing between 5^{th} graders and 8^{th} graders.

5.2.1 Improvement on text processing as participants' grade increases

In line with the suggestions in text processing (Seretny and Dean, 1986), older students can also process text and pictures differently from younger students. Participants had more accurate answers as their age increased or as the period of their education increased. Eighth graders processed text more automatically than 5^{th} graders. Conversely, all participants processed pictures rather similarly. However, regardless of grade, all participants emphasized text to build mental model but focused primarily on pictures to scaffold question answering. The trends of eye movements were analysed based on eye-mind hypothesis and immediacy hypothesis in goal-free (no question yet), delayed goal-oriented (question after material) and initial goal-oriented processing (question before material).

Table 5.1: Trends of eye movements between 5^{th} graders (G5) and 8^{th} graders (G8) in goal-free, delayed goal-oriented and initial goal-oriented processing. T, P and Q are short for text, pictures and question.

G5 vs G8	Goal-free processing	Delayed goal-oriented processing	Initial goal-oriented processing
	a. Accumulated fixation duration (sec)		
G5	T>P	T<P	T>P
G8	T>P	T<P	T>P
	b. Time to first fixation (sec)		
G5	T<P	T>P	T>P
G8	T<P	T>P	T>P
	c. Number of transition		
G5	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q
G8	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q

For goal-free processing, 8^{th} graders fixated shorter on text than 5^{th} graders, whereas 8^{th} graders fixated similarly on pictures as 5^{th} graders. As 8^{th} graders had more accurate answers than 5^{th} graders, it can be interpreted that 8^{th} graders comprehend text Rather Than pictures more efficiently than 5th graders. It can be explained by two reasons. On the one hand, according to the Education Standards in Germany (Kultusministerkonferenz, 2005a, 2005b), text comprehension is a focus in teaching curriculum. However, no information was mentioned about picture comprehension or text and picture integration. Unsurprisingly, students improved greatly in text comprehension but not picture comprehension from Grade 5 to Grade 8. On the other hand, text and pictures have different presentation orientations (Tversky, 2001). For example, text is normally linearly and semantic-oriented: it has a beginning and an end. Picture is generally spatial-oriented. It has no beginning and end, which can be observed from any orientation. Eighth graders are more able to automatically process semantic and syntactic information (van den Broek et al., 2001). Their working memory can leave enough space for searching for the information to answer the question. Therefore, 8th graders paid less attention to text but similar attention to pictures compared to 5th graders. Eighth graders had shorter latency on text than 5th graders. It can indicate that 8th graders were more interested in text or more concentrate on text than 5th graders. Eighth graders had more transitions between text and pictures than 5th graders. According to the Integrative Model of Text and Picture Integration (Schnotz, 2014), the integration of text and pictures can result in the construction of mental model. The results from transitions implied that 8th graders had more frequent integration of text and pictures than 5th graders. It can explain why 8th graders performed better than 5th graders.

Furthermore, 5th graders and 8th graders showed similar fixation trends on text and on pictures with goal-free processing. Text was mainly fixated rather than pictures in goal-free processing. The results were consistent with the suggestions (Zhao et al., 2014) that text guides in mental model construction. Besides, this study found that text was perceived sooner than pictures. Although different from the findings in Zhao et al. (2014), this result confirms that text is emphasized in building mental model. There were frequent transitions between text and pictures. It proves the idea that readers also integrate verbal and pictorial information in order to understand the "whole story" (Paivio, 1986). In short, text comprehension was greatly improved from Grade 5 to Grade 8. Participants from both grades used text mainly in mental model construction.

For delayed goal-oriented processing, no difference was found in fixation on text between 5^{th} graders and 8^{th} graders. Conversely, 8^{th} graders fixated slightly longer on pictures than 5^{th} graders. In delayed goal-oriented processing, participants were required to use the information from text and pictures to answer the question. It indicates that 8^{th} graders have the intention to focus more on pictures than 5^{th} graders. Hochpöchler et al. (2013) pointed out that eighth graders may know the strategy to comprehend pictures and may also know where to apply this strategy. Eighth graders had more accurate answers than 5^{th} graders possibly due to the specific scaffolding function of pictures (Eitel et al., 2013).

Data from latency showed that 8^{th} graders always perceived text and pictures sooner than 5^{th} graders. On the one hand, it implies that 8^{th} graders were more interested in text and pictures than 5^{th} graders. Compared to 5^{th} graders, 8^{th} graders were more motivated to search for the required information in order to answer the question. On the other hand, it indicates that 8^{th} graders were more automatic in text and picture processing than 5^{th} graders. Their working memory was not occupied by superficial processing (i.e. semantic recognition). Having understood the question, 8^{th} graders' attention became more focused than in goal-free processing. When they scanned the text and did not find the useful information, they quickly switched their attention to the pictures. This can explain why 8^{th} graders perceived the multimedia more efficiently than 5^{th} graders. This idea is proved by the evidence in the transitions among AOIs. Eighth graders had more transitions among text, pictures and question than 5^{th} graders. Transitions between text and pictures indicate the integration of text and pictures and the mental model construction (Johnson and Mayer, 2012). Transitions between text and question, pictures and question may imply updating of mental model based on the task model (model to answer the question). Frequent transitions can be interpreted that 8^{th} graders diligently updated and re-constructed their mental model. Therefore, 8^{th} graders performed better in accuracy than 5^{th} graders.

Additionally, 5^{th} graders and 8^{th} graders showed similar trends in text and picture processing (see Table 5.1). All participants fixated longer on pictures than on text with delayed goal-oriented processing. They had shorter latency on pictures than on text. Transitions between text and pictures started to decrease in delayed goal-oriented processing. However, transitions between pictures and question increased remarkably. The data support the idea that text processing differs fundamentally from picture processing (Schnotz, Ludewig, et al., 2014). Picture can be easily accessed to scaffold question answering. When question is displayed, pictures are emphasized by all participants.

For initial goal-oriented processing, 5^{th} graders viewed text longer than 8^{th} graders. Conversely, 8^{th} graders viewed pictures similarly compared to 5^{th} graders. Owing to the good performance in question answering for 8^{th} graders, the results indicate that 8^{th} graders comprehended text efficiently. They may have less or even no difficulty in superficial level of comprehension than 5^{th} graders, for instance, semantic recognition. It confirms the finding that students have learned a lot in text comprehension as they grow older (Carpenter et al., 1994). The improvement in text comprehension can also be a result of the school curriculum. Additionally, two reasons can explain why participants did not differ in picture comprehension. The first reason may due to the distinction between text (descriptive representation) and pictures (depictive representation). Text is normally comprehended in a certain sequence, for example, from left to right (Arabic), from up to down (Japanese). Picture represents information rather spatially, which can be accessed from any direction. For example, a world map can be perceived from southern pole or can be comprehended directly from Europe.

Another reason can be explained by the focus of teaching curriculum at school. Text comprehension is clearly stated as an important competence in school curriculum (Kultusministerkonferenz, 2005a). However, no information is mentioned about picture comprehension or text-picture integration. Likely, students are not well-trained in picture comprehension in school because it is not emphasized in school curriculum. The latency of first fixation on text shows that 8^{th} graders perceived text sooner than 5^{th} graders. Conversely, no difference was revealed between 5^{th} graders and 8^{th} graders in latency of first fixation on pictures. It implied that 8^{th} graders were possibly more interested in or concentrated on text than 5^{th} graders. From another perspective, it also indicated that 8^{th} graders were better problem solvers than 5^{th} graders. When the question was displayed, 8^{th} graders wasted less time in searching for the information in text to answer the question than 5^{th} graders. For picture processing, 8^{th} graders did not distinguish from 5^{th} graders, which can be explained by the curriculum. Data from transitions showed that 8^{th} graders had more transitions than 5^{th} graders among text, pictures and question. Schnotz, Ludewig, et al. (2014) demonstrates that the transitions between text and pictures can visualize the cognitive process of mental model construction. Frequent transitions can be interpreted as 8^{th} graders took more effort in mental model construction than 5^{th} graders. Transitions between text and question, pictures and question can be regarded as the process of updating the mental model based on the task model (created by question). Kirby, Cain, and White (2012) highlights that better mental model construction leads to deeper comprehension. Eighth graders updated their mental model more actively and diligently than 5^{th} graders. The results support the idea that 8^{th} graders are better problem solvers than 5^{th} graders.

In general, participants from both grades had similar trends in text and picture processing with initial goal-oriented processing (see Table 5.1). They fixated longer on text than on pictures when question was presented first. Apart from this, they had shorter latency on first fixating on text than on pictures. They transferred their attention most frequently between text and pictures, more frequently between pictures and question and least frequently between text and question. The results are in line with the findings from the pilot study (Zhao et al., 2014), that text is mainly used for mental model construction. Picture is primarily used for question answering. However, this cross-sectional design cannot guarantee whether the age differences are indeed due to the development with age or, alternatively, due to cohort differences. Specifically, the 8th graders may have better training when they were at the same age as the 5th graders.

5.2.2 Emphasis on pictures as question difficulty increases

Fifth graders differed from 8th graders in text and picture processing when they solved easy (level 1) and difficult questions (level 2/3). However, all participants seemed to follow consistent patterns in using text and pictures. With easy questions, 8th graders had more accuracy than 5th graders. Fifth graders fixated longer on text and on pictures than 8th graders with delayed goal-oriented processing and with initial goal-oriented processing. Eighth graders perceived text and pictures sooner than 5th graders. Based on the accuracy data, the fixation patterns can be interpreted as 5th graders had more difficulties in comprehending text and pictures than 8th graders. The variability of reading comprehension may be caused by students' ability in cognitive processing. In line with van den Broek et al. (2001), older students are possibly more capable of processing the information automatically than younger students. Fifth students may still be struggled with superficial level of processing (Hannus and Hyönä, 1999). As a consequence, 5th graders comprehended text quicker than 8th graders. However, the results on picture processing are different from the findings from Hochpöchler et al. (2013). In the current study, 5th graders paid more attention to pictures than 8^{th} graders when they solved easy question. As mentioned in the design (see 3.1.4), easy question only requires processing of elements or values. Fifth graders may have difficulties in semantic recognition or syntactic decomposition, thus they search for the required information rather slow.

In addition, all participants followed a certain trend in focusing on text and on pictures. With goal-free processing, students from Grade 5 and Grade 8 fixated longer on text than on pictures. They had also shorter latency of first fixation on text. With delayed goal-oriented processing, they emphasized mainly on pictures rather than on text. They perceived pictures sooner than text. With initial goal-oriented processing, they fixated longer on text than on pictures. Conversely, they perceived pictures quicker than text. From another perspective, they show evidence that mental model construction is mainly guided by text. Mental model usage or question solving is primarily guided by pictures.

With difficult question, 8^{th} graders had more accurate answers than 5^{th} graders. Eighth graders viewed text and pictures longer than 5^{th} graders with delayed goal-oriented processing. Eighth graders also had less latency on fixating on text and on pictures than 5^{th} graders with this processing. In the design, question was posed in delayed goal-oriented processing. After constructed the task model from the question (Rouet, 2006), participants needed to find the relevant information in text and pictures to answer the question. Eighth graders spent a huge amount of time on text and pictures. It indicated that they looked back and forth to search for the required information. According to the Integrative Model of Text-Picture Integration (Schnotz, 2014), 8^{th} graders may update their mental model based on the criterion from the task model. Data from time to the first fixation implied that 8^{th} graders were more interested in searching the information than 5^{th} graders. One the one hand, eighth graders may be more sensitive to the requirements of the tasks than 5^{th} graders (Hartman, 2001). On the other hand, 8^{th} graders can process the information more automatically than 5^{th} graders (Oakhill and Yuill, 1996). They possibly scan text or pictures quickly to search for the relevant information. In contrast, 8^{th} graders had shorter fixation on text than 5^{th} graders with initial goal-oriented processing. No difference was detected between 5^{th} graders and 8^{th} graders in fixating on pictures with initial goal-oriented processing. Eighth graders perceived text sooner than 5^{th} graders. Conversely, 5^{th} graders perceived pictures quicker than 8^{th} graders. Eighth graders had shorter fixation on text than 5^{th} graders. It implied that they improved text comprehension enormously compared to 5^{th} graders. 8^{th} graders can quickly process the verbal information and obtain rather accurate answers.

However, based on accuracy data, the similar fixation duration on pictures can be interpreted differently for 5^{th} graders and 8^{th} graders. Fifth graders may have difficulty in superficial level of text processing. Thus they may resort to pictures to understand the meaning and to locate the relevant information (Rusted and Coltheart, 1979). Due to the low accuracy, they may also have difficulty in interpreting the information in pictures and applying suitable strategy in using pictures (Hasselhorn, 1996). Conversely, 8^{th} graders may be more sensitive to the difficulty of the task and may conduct deep processing. This idea can be supported by the good performance of accuracy from 8^{th} graders. Besides, indicator of time to the first fixation shows that 8^{th} graders may be more interested than 5^{th} graders in searching for the information in text. Likely, 8^{th} graders were certain what to do with the text. In contrast, the rather opposite data on pictures indicate that 5^{th} graders may depend on pictures to compensate for the deficiency of the verbal channel (Mayer, 2001). Thus students from Grade 8 outperformed students from Grade 8 in accuracy.

Similar to easy questions, participants from both grades had similar fixation trends with difficult questions (see Table 5.2). They focused primarily on text and perceived text quickly with goal-free processing. Picture was mainly emphasized and was viewed soon with delayed goal-oriented processing. In contrast, text was mainly focused but pictures were perceived soon with initial goal-oriented processing. The results support the argument that text processing distinguishes from picture processing (Mayer, 2014b). When using different reading strategies, text can mainly guide readers to construct mental model. Picture can guide readers to scaffold question solving.

Table 5.2: Trends of eye movements between 5th graders (G5) and 8th graders (G8) in goal-free, delayed goal-oriented and initial goal-oriented processing with easy (level 1) and difficult questions (level 2/3). T, P and Q are short for text, pictures and question.

G5 vs G8	Goal-free processing	Delayed goal-oriented processing	Initial goal-oriented processing
	a. Accumulated fixation duration (sec)		
Level1-G5	T>P	T<P	T>P
Level2-G5	T>P	T<P	T>P
Level1-G8	T>P	T<P	T>P
Level2-G8	T>P	T<P	T>P
	b. Time to first fixation (sec)		
Level1-G5	T<P	T>P	T>P
Level2-G5	T<P	T>P	T>P
Level1-G8	T<P	T>P	T>P
Level2-G8	T<P	T>P	T>P

5.2.3 Pictures may distinguish between good vs poorer problem solvers

Regardless of school tier, participants from Grade 8 processed text sooner than participants from Grade 5 (see 5.3). However, higher tier students (Gymnasium) from both grades distinguished from lower tier students (Realschule) only in picture processing. Consistent with the hypothesis, 8th graders had better performance in accuracy than 5th graders. Higher tier students from both grades had higher accuracy than lower tier students from both grades. Higher and lower tier students from Grade 8 processed text rather more efficiently than all 5th graders. No difference was found in picture processing among 5th graders and 8th graders. The results indicate that

Table 5.3: Trends of eye movements between Gymnasium (Gym) and Realschule (Real) students from Grade 5 (G5) and Grade 8 (G8) in goal-free, delayed goal-oriented and initial goal-oriented processing. T, P and Q are short for text, pictures and question.

G5 vs G8	Goal-free processing	Delayed goal-oriented processing	Initial goal-oriented processing
	a. Accumulated fixation duration (sec)		
Gym-G5	T>P	T<P	T>P
Real-G5	T>P	T<P	T>P
Gym-G8	T>P	T<P	T>P
Real-G8	T>P	T<P	T>P
	b. Time to first fixation (sec)		
Gym-G5	T<P	T>P	T>P
Real-G5	T<P	T>P	T>P
Gym-G8	T<P	T>P	T>P
Real-G8	T<P	T>P	T>P
	c. Number of transition		
Gym-G5	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q
Real-G5	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q
Gym-G8	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q
Real-G8	T-P	P-Q>T-P>T-Q	T-P>P-Q>T-Q

8^{th} graders improve in text processing rather than in picture processing compared to 5^{th} graders. The huge improvement in text comprehension is regarded as an effect of Teaching Standards in Germany (Kultusministerkonferenz, 2005b). No apparent improvement in picture comprehension can be considered as the features of pictorial illustrations compared to text. Meanwhile, it can indicate the inadequate training in picture comprehension at school.

Interestingly, higher tier students from both grades had similar fixation on text but longer fixation on picture processing than lower tier students from both grades. The results are contrary to the hypothesis on text processing. The difference between higher and lower tier students did not lie in text processing, although higher tier students performed slightly better in decoding and encoding the verbal information than lower tier students. However, all results indicate that higher tier students invest much longer time on pictures than lower tier students. Based on the accuracy data, picture comprehension may determine the difference between higher and lower students. In line with Schnotz (2014), advanced readers know how to comprehend pictures and know when to apply this strategy compared to poorer readers. Likely, higher tier students outperform lower tier students as a result of the ability in picture processing.

5.2.4 Shortcomings and issues remain to be resolved with future research

Some limitations should be acknowledged when interpreting the results from two experiments. At first, it can be worth integrating eye tracking data with other process measures. As eye tracking indicators can have opposite indications (i.e. longer fixation may indicate deep processing or perceiving difficulty), the combination of multiple data sources can better explain the cognitive processes of readers (Mason et al., 2014). In this respect, integrating other process measures can reduce the disadvantage of eye tracking, for instance, combining eye tracking with think-loud protocol (Kammerer and Gerjets, 2012). Sometimes think-aloud protocols can hinder the cognitive process of multimedia presentation (Dalgarno, Kennedy, and Bennett, 2010). When interview them about their cognitive process after the experiment, the validity and reliability of the data are questionable (Deatrick and Faux, 1989; Ornstein, Gordon, and Larus, 1992). However, think-aloud protocol is still a useful tool to reveal the thinking patterns in integrating text and pictures. The combination of both process measures can minimize the bias of the researchers' interpretation. This method should also be used in the future studies so that researchers can have a better understanding of learning with multimedia presentations.

Besides, prior knowledge can also be helpful to understand the fixation patterns of participants. The current studies did not conduct test on prior knowledge due to the limited time of this project. However, the existed schema or prior knowledge can influence cognitive processes with multimedia presentations. Readers may use their prior knowledge to select strategies to comprehend the meaning of text (R. C. Anderson and Pearson, 2002). When learners learn new information, they actively building new schema (structure of knowledge) or attach the new knowledge to the existing schema (Ausubel, 1968). The existed knowledge can differ in facilitating learning or impeding learning. It is argued that learning by novel readers can be greatly enhanced, whereas learning by experienced readers can be less effective or hindered (Kalyuga, 2014). It can be explained by integrating the perceived information in working memory with the existed knowledge structure in long-term memory. When the prior knowledge is controversial to the presented information, prior knowledge can become obstruction in learning (Lipson, 1982). You can imagine a Chinese speaker who starts to learn German. S/he may struggle from changing from no gender and no case into three genders (i.e. der, die and das) and four cases (i.e. Nominative, Akkusative, Dative and Genitive).

Furthermore, when learners have low prior knowledge, they often have difficulty in comprehending the graphics and making inference with the prior experience (Heiser and Tversky, 2004; Larkin and Simon, 1987). In other words, they are less able to construct a mental model, which can facilitate making inference. When learners have enough prior knowledge, they have less difficulty in extracting relevant information and making inferences to pictures (Hegarty and Sims, 1994; Heiser and Tversky, 2002). However, learners are also reluctant to abandon their prior knowledge only if it can no longer support their experience (R. C. Anderson, 1977). The future study concerning multimedia learning should implement prior knowledge test so that the eye tracking patterns can be better understood.

At last, the study tends to generalize its findings in all the science subjects. In the two experiments, the materials were only chosen from geography and biology. Other science subjects can remarkably differ from them (i.e. mathematics, physics, etc.). In order to obtain a universal cognitive process, other science subjects should be taken into account and should be further explored.

5.2.5 Educational implications for curriculum and teaching practical

Based on the findings in two experiments, pictures are suggested to be the essence in obtaining correct answers. When question is displayed, good problem-solvers tend to use pictures more intensively than poorer problem-solvers. The results also reveal the effects of Teaching Standard (Kultusministerkonferenz, 2005a) that students improve remarkably in text comprehension from Grade 5 to Grade 8. However, picture comprehension has not been enormously enhanced. It highlights the necessity of picture comprehension as school science curriculum did not mention anything about picture comprehension or text-picture comprehension. Apart from this, this curriculum standard in Germany has not been edited for over 10 years. The competences required by the society may change over the decade. For instance, due to the limitation of technology, learning materials were rather text-centred a decade ago. However, presentations with text and pictures become more and more popular in learning now. Having the ability in integrating verbal and pictorial information is a necessity not only at school but also for everyday life. In short, Teaching Standards may be modified by mentioning competences in picture comprehension and integrating text and pictures. Without a doubt, new standards alone will not improve the situation. During teaching practice, teachers should also pay attention to the training on picture comprehension and integration between text and pictures.

5.2.6 Conclusion

Fifth graders and 8^{th} graders show similar fixation patterns in comprehending text and pictures with different reading strategies. By manipulating reading strategies, different functions of text and pictures are discovered. Building mental model is mainly guided by text, whereas using mental model is primarily guided by pictures. Participants from both grades also follow the same trend in using text and pictures with easy and difficult questions. Text is mainly used with easy questions. Conversely, pictures are mainly used with difficult questions. For both graders, higher tier students outperform lower tier students in picture processing rather than text processing. Although participants from both grades follow similar trends in fixating on text and pictures in perspective of task orientation, question difficulty and school tier, 5^{th} graders distinguish from 8^{th} graders in fixating on text rather than on pictures. The educational implications are highlighted that text processing differs from picture processing. Competence on picture comprehension or text-picture comprehension should be added into the Teaching Standards and should be applied to teaching science at school.

Reading strategy caused by task orientation (placing question before or after the material) can constitute different processes in the domain of text comprehension. When question is posed after the material, it activates general coherence-formation processing (Rothkopf and Billington, 1974). In this condition, reading is rather general. When question is posed before the material, it triggers selective task-oriented processing. In this situation, reading is comparatively selective. However, reading strategy is rarely mentioned in the context of multimedia learning. Therefore, it is necessary and meaningful to investigate whether the location of questions can have positive effects on multimedia learning.

With eye-tracking method, this research explores whether readers use text differently from pictures when question is placed before or after the material. To deepen the understanding of task orientation, this study also examins whether question difficulty, school tier and grade factors influence multimedia learning.

It is discovered that the text and pictures are used as different tools when comprehending materials combining text and pictures with different reading strategies. The reader is majorly guided by text when s/he comprehended materials. Conversely, the reader is mainly guided by pictures when s/he answer questions. In other words, text and pictures play fundamentally different roles when s/he builds or uses the mental model. Building the mental model is guided by text; whereas, using the mental model is guided by pictures.

Besides, readers comprehend easy and difficult questions differently with multimedia presentations. Slightly different from Frase's Directed Attention Hypothesis (Frase, 1969a), readers pay similar attention to text with easy and difficult questions. However, they pay much more attention to pictures with difficult questions. It also suggests that the demands of difficult questions can cause more intensive integration of the verbal and pictorial materials than the demands of easy questions. The outcome of the intensive integration can lead to the deep processing. Besides, readers perform much better with easy questions than difficult questions.

Furthermore, readers from higher school tier are more capable to handle pictures than lower school tiers. It is shown that readers from higher tier are prior to readers from lower tier mainly due to comprehending pictures instead of text. Readers from higher school tiers focus less on text and more on pictures than lower tiers. Readers from higher school tiers have more accuracy than readers from lower school tiers.

However, readers from higher grade are more automatic in text processing than readers from lower grade. No difference is detected in picture processing between these two groups. It is also shown that readers from higher grade have a huge improvement of text reading but not picture comprehension. Readers from higher grade have more accurate answers than readers from lower grade.

Finally, it proposes educational implications for designing school science curriculum and improving teaching pedagogy. School science curriculum mainly focuses on verbal reading. This is also proved by this study that readers improve a lot in verbal comprehension from Grade 5 to Grade 8. However, teachers should possibly focus on pictorial comprehension in science curriculum because pictorial comprehension seems to be the gap between poorer readers and advanced readers.

Appendix A

Text-picture materials

The following materials combining text and picture were displayed on Tobii eye tracker (translated from German).

A.1 Banana trade

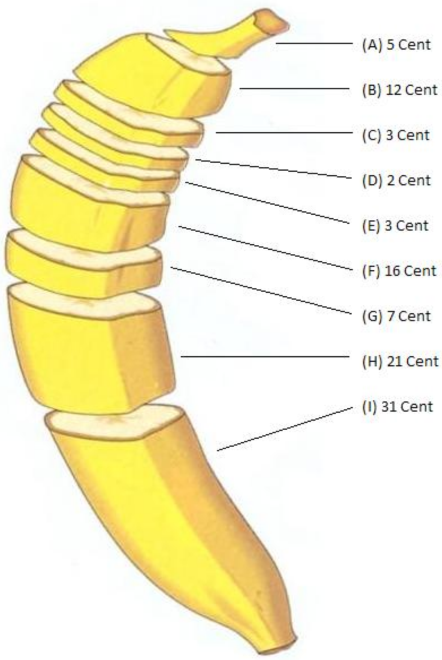


Figure A.1: Material of Banana Trade displayed on Tobii eye tracker.

Many people like to eat bananas. They are planted in countries like Ecuador, Costa Rica or Columbia and then exported to Europe. Undoubtedly, this is related to costs.

The banana that you see in the picture costs just one euro. This euro includes...

(A) salary of the farmers

(B) cost of the fertilizer

(C) cost for transportation to the harbor

(D) profits of the plantation owners

(E) tax for bananas

(F) cost for shipping

(G) profit for the wholesalers

(H) cost for storage

(I) profit of the retailer

Question Level 1

How many cents can a retailer earn from a banana?

- 3 Cent
- 21 Cent
- 31 Cent
- 7 Cent

Question Level 2

For whom do people pay the least amount of money, if they buy a banana?

- profit of the wholesalers
- profit of the plantation owners

- profit of the retailers
- salary of the farmers

Question Level 3

If we compare the farmers, retailers and wholesalers, then...

- farmers earn the most, wholesalers earn less and retailers earn the least from a banana
- retailers earn the most, wholesalers earn less and farmers earn the least from a banana
- retailers earn the most, farmers earn less and wholesalers earn the least from a banana
- wholesalers earn the most, farmers earn less and retailers earn the least from a banana

A.2 Legs of insects

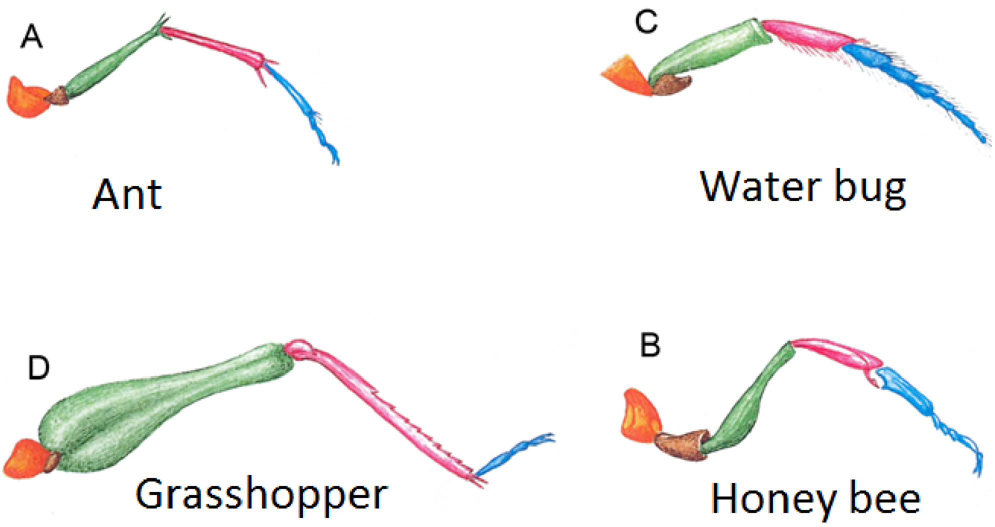


Figure A.2: Material of Legs of Insects displayed on Tobii eye tracker.

The legs of insects presented in figures A to D have the same structure: Hip (orange), leg ring (brown), thigh (green), bar (pink), and foot (blue).

Question Level 1

Which insect has a cleaning leg?

- ant
- honey bee
- water bug
- grasshopper

Question Level 2

Which type of leg has the longest leg ring?

- leg for cleaning
- leg for jumping

- leg for running
- leg for swimming

Question Level 3

Does the leg for cleaning compared to the leg for jumping have...

- a longer bar but a shorter foot
- a thicker thigh but a thinner leg ring
- a longer foot but a shorter bar
- a shorter foot but a longer leg ring

A.3 Auditory

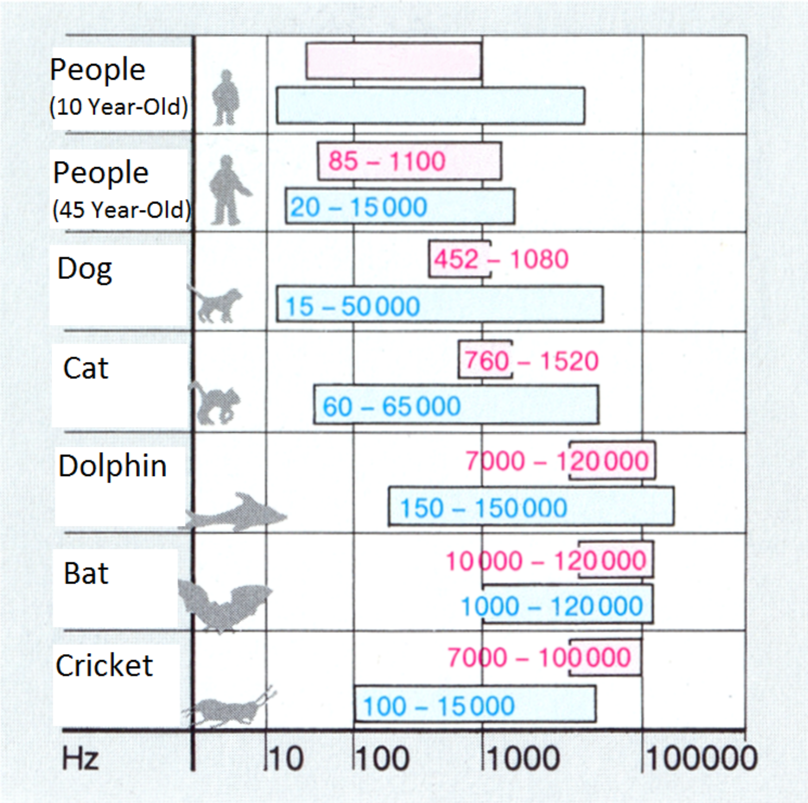


Figure A.3: Material of Auditory displayed on Tobii eye tracker.

Tones and sounds are sound-waves. The faster the sound-vibrations, the higher we perceive the sound/tone. The human ear can differentiate sounds/tones with low vibrations (20) and high vibrations (20 000) per second. The number of vibrations per second is called frequency; its unit is Hertz (Hz). A ten-year-old child is able to hear every sound/tone between the frequencies of 20 and 20000 Hertz. This area is called the hearing range, which is displayed in blue in the picture. Furthermore, a ten-year-old child is able to produce sounds/tones, for example by speaking, which are between 70 and 1000 Hertz. This area is called vocal range, which is displayed in pink in the picture. The illustration shows the hearing and vocal ranges of different species.

Question Level 1

Which of the following species is able to perceive tones/sounds at 120 Hertz?

- dog
- cricket
- cat
- bat

Question Level 2

Which of the following species has a vocal range for producing the lowest tones?

- human being, 45 year-old
- cat
- dog
- cricket

Question Level 3

Which of the following four species is able to hear tones below 100 Hertz as well as produce tones below 1000 Hertz and above 1500 Hertz?

- human being, 10 year-old
- human being, 45 years old
- cricket
- cat

A.4 Pregnancy

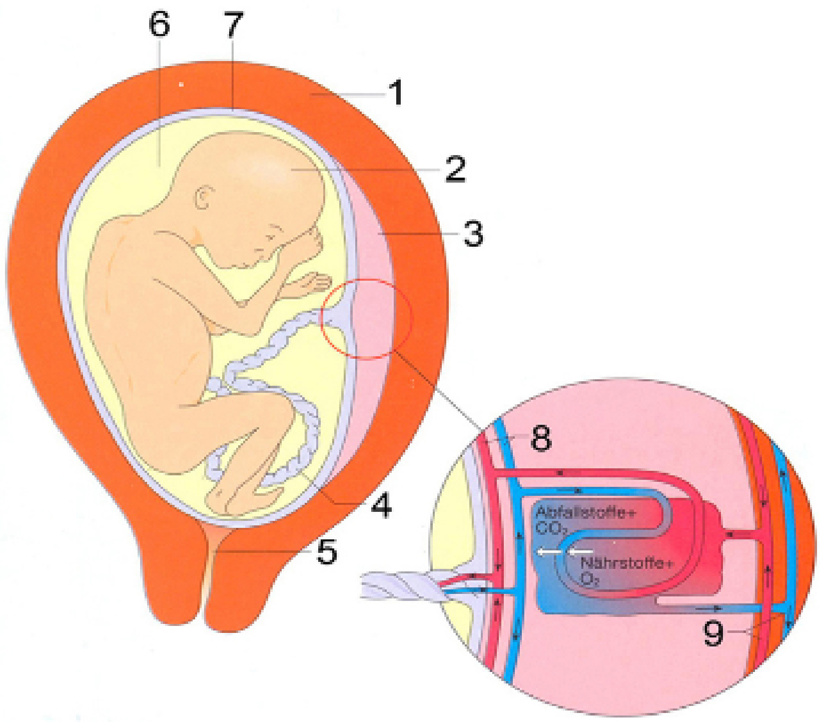


Figure A.4: Material of Pregnancy displayed on Tobii eye tracker.

The child develops in the uterus, or uterine wall (1). From the fourth month of gestation, it is called a fetus (2). The fetus is nourished by the placenta (3). It is here where exchange between the blood vessels of the fetus (8) and the blood vessels of the mother (9) takes place (look at zoomed area). In the blood vessels, nutrients and oxygen (O2) as well as waste products and carbon dioxide (CO_2) are exchanged. The fetus is connected to the mother by the umbilical cord (4). The amniotic fluid protects the child. It could be called a protective-pillow for the fetus because it helps to cushion the fetus from impact. When the mother's water membrane (7) is broken, the delivery process is initiated and the fetus will move through the cervix (5) during the labour.

Question Level 1

What is the name of the pink area?

- placenta

- umbilical cord
- uterine wall
- amniotic fluid

Question Level 2

Which parts do not directly link to each other?

- blood vessels of the mother and blood vessels of the child
- cervix and amniotic fluid
- umbilical cord and water membrane
- placenta and uterine wall

Question Level 3

Which path does the blood of the fetus take after getting nutrition and oxygen (CO_2) from the mother? It flows ...

- back through the placenta and by umbilical cord to the fetus
- back through the placenta and by the blood vessels of the fetus to the water membrane
- back through the placenta and by blood vessels of the fetus to the uterine wall
- back through the placenta and by blood vessels of the mother directly to the fetus

A.5 Map of Europe

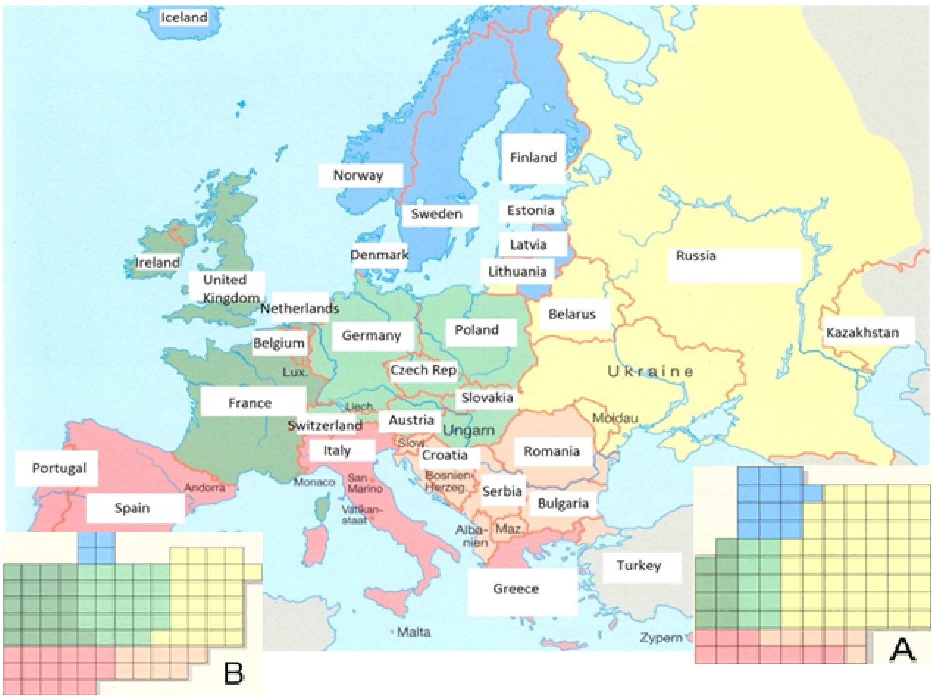


Figure A.5: Material of Europe displayed on Tobii eye tracker.

The big map shows the continent of Europe. Actually, Europe is not an independent continent. Together with Asia, it forms the continent “Eurasia”. In the south, west and north, the border of Europe is clearly defined by the seas. The delineation in the east is more difficult because there are no natural borders. An agreement set the borders at the Ural Mountains and further south, so parts of Russia and Kazakhstan belong to both Europe and Asia.

Europe are divided into different subspaces according to economic and geographic features. These subspaces are:

Northern Europe (blue)

Western Europe (dark green)

Central Europe (light green)

Southern Europe (red)

Eastern Europe (yellow)

In Figure A, one box represents one unit of the European area.

In Figure B, one box represents one unit of the European population.

Question Level 1 [1]

In which subspace are countries located in the area that belongs to both Europe and Asia?

- Northern Europe
- Southern Europe
- Middle Europe
- Eastern Europe

Question Level 2

Which two subspaces have the same number of area units?

- Eastern Europe and Northern Europe
- Southern Europe and South-eastern Europe
- Middle Europe and Southern Europe
- West Europe and Middle Europe

Question Level 3

Which subspace has one more unit of area, but four fewer units of population compared to Western Europe?

- Middle Europe
- Northern Europe

[1] I appreciate the comments from Prof. Dr. Alex Kozulin from Feuerstein Institute, International Institute for the Enhancement of Learning Potential in Jerusalem. Thanks for pointing out that the cream colour in the map is not marked in the text as "Balkan countries". As the experiments have been already finished, the corrected text will be used for the future study.

- Southern Europe
- South-eastern Europe

A.6 Savannah

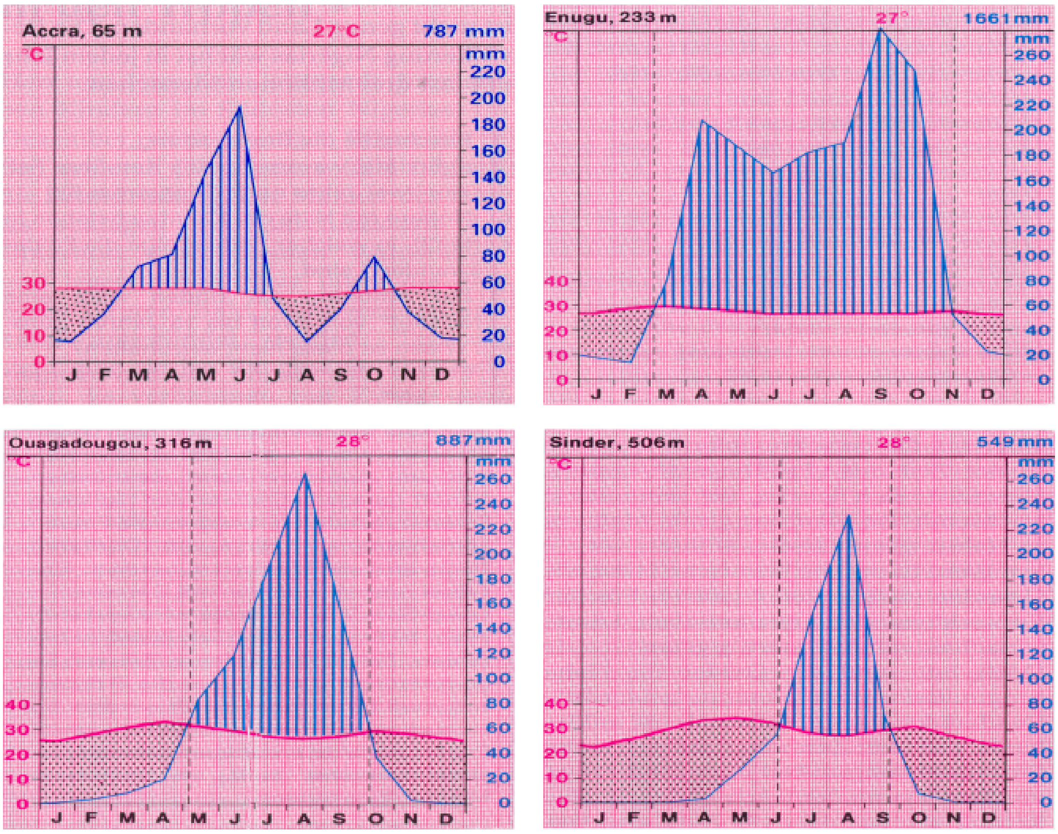


Figure A.6: Material of Savannah displayed on Tobii eye tracker.

Because there are different rainy seasons, the savannah is differentiated into separate types. The city of Enugu is located in the wet savannah; Accra and Ouagadougou are located in the dry savannah and the city of Sinder in the thorn-bush savannah. Depending on the amount of rainfall (e.g. Ouagadougou 887 mm rainfall per year), different food is cultivated and plants exported in each region and city. The months with enough rain for growing the respective plants are shown in the diagrams by the areas with blue stripes above the red lines for temperature. If the line for rainfall (blue) is above the line for temperature (red), then there is more rainfall than evaporation. Each month is represented by a letter in the lower part of the diagrams, for example F = February.

The following plants need different amounts of rainfall per year for ideal growth:

Millet: 180 to 700 mm

Manioc: 500 to 2000 mm

Yams: more than 1500 mm

Peanut: 250 to 700 mm

Cotton: 700 to 1500 mm

Question Level 1

What is the amount of rainfall per year in Accra?

- 887 mm
- 1661 mm
- 787 mm
- 549 mm

Question Level 2

Which plant can grow well in Enugu?

- peanut
- yams
- cotton
- millet

Question Level 3

Which plants can grow well in the most cities, respective of the type of savannah?

- millet
- manioc
- cotton
- peanut

Bibliography

Allen, D. I. (1970). Some effects of advanced organizers and level of question on the learning and retention of written social studies material. *Journal of Educational Psychology*, *61*, 333–339. doi:10.1037/h0029909

Anderson, J. R. & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, pp. 89–132). New York: Academic Press.

Anderson, R. C. (1970). Control of student mediating processes during verbal learning and instruction. *Review of Educational Research*, *40*(3), 349–369. doi:10.2307/1169371

Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In R. C. Anderson, J. Spiro, & W. E. Montague (Eds.), *Schooling and acquisition of knowledge* (pp. 415–431). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, R. C. & Pearson, P. D. (2002). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, B. Rebecca, M. L. Kamil, & P. B. Mosenthal (Eds.), *Handbook of reading research* (pp. 255–291). Mahwah, NJ: Lawrence Erlbaum.

André, T. (1979). Does answering higher-level questions while reading facilitate productive learning? *Review of Educational Research*, *49*(2), 280–318. doi:10.2307/1169962

André, T. & Thieman, A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. *Contemporary Educational Psychology*, *13*(3), 296–307. doi:10.1016/0361-476X(88)90028-8

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune and Stratton.

Ausubel, D. P. (1968). *Educational psychology: a cognitive view*. New York: Holt, Rinehart, and Winston.

Bauer, M. I. & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science*, *6*, 372–378. doi:10.1111/j.1467-9280.1993.tb00584

Beniger, J. R. & Robyn, D. L. (1978). Quantitative graphics in statistics: a brief history. *The American Statistician*, *32*, 1–11. doi:10.1007/BF00055241

Bertin, J. (1967). *Semiologie graphique: les diagrammes-les reseaux-les cartes*. The Hague: Mouton.

Bluth, L. F. (1973). *A comparison of the reading comprehension of good and poor readers in the second grade with and without illustration* (Thesis).

Boyd, W. M. (1973). Repeating questions in prose learning. *Journal of Educational Psychology*, *64*(1), 31–38. doi:10.1037/h0034061

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: a constructive versus interpretive approach. *Cognitive Psychology*, *3*, 193–209. doi:10.1016/0010-0285(72)90003-5

Briggs, L. J. & Reed, H. B. (1943). The curve of retention for substance material. *Journal of Experimental Psychology*, *32*(6), 513–517. doi:10.1037/h0060553

Carney, R. N. & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, *14*(1), 5–26. doi:10.1023/A:1013176309260

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension: evidence from individual differences, aphasia, and aging. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1075–1122). San Diego, CA: Academic Press.

Carswell, C. M. & Wickens, C. D. (1998). *Comparative graphics: history and applications of perceptual integrality theory and the proximity compatibility hypothesis*. Institute of Aviation: University of Illinois at Urbana-Champaign.

Cataldo, M. G. & Oakhill, J. (2000). Why are poor comprehenders inefficient searchers? an investigation into the effects of text representation and spatial memory on the ability to locate information in text. *Journal of Educational Psychology*, *92*(4), 791–99. doi:10.1037/0022-0663.92.4.791

Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011). Selecting information to answer questions: strategic individual differences when searching texts. *Learning and Individual Differences Learning and Individual Differences*, *21*(2), 201–205. doi:10.1016/j.lindif.2010.11.007

Cerdán, R., Vidal-Abarca, E., Martinez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction*, *19*(1), 13–27. doi:10.1016/j.learninstruc.2007.12.003

Chauvet, J. M., Brunel, D. E., & Hillaire, C. (1996). *Dawn of art: the chauvet cave-the oldest known paintings in the world*. New York: H. N. Abrams.

Chiou, C.-C., Tien, L.-C., & Lee, L.-T. (2015). Effects on learning of multimedia animation combined with multidimensional concept maps. *Computers & Education*, *80*, 211–223. doi:10.1016/j.compedu.2014.09.002

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Clark, H. H. & Clark, E. V. (1977). *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Clements, M. A. (2013). *Third international handbook of mathematics education*. New York: Springer.

Cooney, J. B. & Swanson, H. L. (1987). Memory and learning disabilities: an overview. In H. L. Swanson (Ed.), *Memory and learning disabilities: advances in learning and behavioral disabilities* (pp. 1–40). Greenwich, CT: JAI.

Craik, F. I. & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*(6), 671–684. doi:10.3758/BF03210735

Dalgarno, B., Kennedy, G., & Bennett, S. (2010). Can functional brain imaging be used to explore interactivity and cognition in multimedia learning environments? *Journal of Interactive Learning Research*, *21*(3), 317–342.

Daniels, P. T. & Bright, W. (1996). *The world's writing systems*. New York: Oxford University Press.

Deatrick, J. A. & Faux, S. A. (1989). Conducting qualitative studies with children and adolescents. In J. M. Morse (Ed.), *Qualitative nursing research: a contemporary dialogue* (pp. 185–203). Rockville, MD: Aspen.

DeLeeuw, K. E. & Mayer, R. E. (2008). A comparison of three measures of cognitive load: evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223–234. doi:10.1037/0022-0663.100.1.223

Denis, M. (1984). Imagery and prose: a critical review of research on adults and children. *Text Interdisciplinary Journal for the Study of Discourse*, *4*(4), 381–402. doi:10.1515/text.1.1984.4.4.381

Denis, M. & Denhiere, G. (1990). Comprehension and recall of spatial descriptions. *European Bulletin of Cognitive Psychology*, *10*, 115–143. doi:10.3758/BF03200870

Distad, H. W. (1927). A study of the reading performance of pupils under different conditions on different types of materials. *Journal of Educational Psychology*, *18*(4), 247–258. doi:10.1037/h0074065

Doherty, M. E. & Anderson, R. B. (2009). Variation in scatterplot displays. *Behavior Research Methods*, *41*(1), 55–60. doi:10.3758/BRM.41.1.55

Dornisch, M. M. & Sperling, R. A. (2006). Facilitating learning from technology-enhanced text: effects of prompted elaborative interrogation. *The Journal of Educational Research*, *99*(3), 156–165. doi:10.3200/JOER.99.3.156-166

Eitel, A., Scheiter, K., Gerjets, P., & Kuhl, T. (2014). Disfluency meets cognitive load in multimedia learning: does harder-to-read mean better-to-understand? *Applied Cognitive Psychology*, *28*(4), 488–501. doi:10.1002/acp.3004

Eitel, A., Scheiter, K., Schüler, A., Nyström, M., & Holmqvist, K. (2013). How a picture facilitates the process of learning from text: evidence for scaffolding. *Learning and Instruction*, *28*, 48–63. doi:10.1016/j.learninstruc.2013.05.002

English, H. B., Welborn, E. L., & Killian, C. D. (1934). Studies in substance memorization. *Journal of General Psychology*, *11*, 233–260. doi:10.1080/00221309.1934.9917837

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using gpower 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. doi:10.3758/BRM.41.4.1149

Fisher, D. L. (1981). Functional literacy tests: a model of question-answering and an analysis of errors. *Reading Research Quarterly*, *16*(3), 418–448. doi:10.2307/747410

Fletcher, J. D. & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 117–133). New York: Cambridge University Press.

Frase, L. T. (1967). Learning from prose material: length of passage, knowledge of results, and position of questions. *Journal of Educational Psychology*, *58*(5), 266–272. doi:10.1037/h0025028

Frase, L. T. (1968a). Effect of question location, pacing, and mode upon retention of prose material. *Journal of Educational Psychology*, *59*(4), 244–249. doi:10.1037/h0025947

Frase, L. T. (1968b). Questions as aids to reading: some research and theory. *American Educational Research Journal*, *5*(3), 319–332. doi:10.1007/BF01079934

Frase, L. T. (1969a). Cybernetic control of memory while reading connected discourse. *Journal of Educational Psychology*, *60*(1), 49–55. doi:10.1037/h0026636

Frase, L. T. (1969b). Structural analysis of the knowledge that results from thinking about text. *Journal of Educational Psychology Monograph Supplement*, *60*(6, Pt.2), 1–16. doi:10.1037/h0021520

Frase, L. T. (1970a). Boundary conditions for mathemagenic behaviors. *Review of Educational Research*, *40*(3), 337–349. doi:10.1037/h0076934

Frase, L. T. (1970b). Influence of sentence order and amount of higher level text processing upon reproductive and productive memory. *American Educational Research Journal*, *7*(3), 307–319. doi:10.3102/00028312007003307

Frase, L. T. (1972). Maintenance and control in the acquisition of knowledge from written materials. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. New York: Wiley.

Friendly, M. (2005). Milestones in the history of data visualization: a case study in statistical historiography. In C. Weihs & W. Gaul (Eds.), *Classification: the ubiquitous challenge* (pp. 34–52). New York: Springer.

Friendly, M. (2008). The golden age of statistical graphics. *Statistical Science*, *23*(4), 502–535. doi:10.1214/08-STS268

Funkhouser, H. G. (1937). Historical development of the graphical representation of statistical data. *Osiris*, *3*, 269–405. doi:10.1080/01621459.1938.10502343

Gaschler, R., Marewski, J. N., & Frensch, P. A. (2015). Once and for all-how people change strategy to ignore irrelevant information in visual tasks. *Quarterly Journal of Experimental Psychology*, *68*(3), 543–567. doi:10.1080717470218.2014.961933

Gelb, I. J. (1963). *A study of writing* (2nd ed.). Chicago: University of Chicago Press.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83.

Godau, C., Wirth, M., Hansen, S., Haider, H., & Gaschler, R. (2014). From marbles to numbers—estimation influences looking patterns on arithmetic problems. *Psychology*, *5*(2), 127–133. doi:10.4236/psych.2014.52020

Goldberg, F. (1974). Effects of imagery on learning incidental material in the classroom. *Journal of Educational Psychology*, *66*, 233–237. doi:10.1037/h0036279

Goldberg, J. H. & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, *24*(6), 631–645. doi:10.1016/S0169-8141(98)00068-7

Golden, M. L. (1942). Reading guided by questions versus careful reading followed by questions. *Journal of Educational Psychology*, *33*(6), 463–468. doi:10.1037/h0055622

Goldman, S. R. & Durán, R. P. (1988). Answering questions from oceanography texts: learner, task, and text characteristics. *Discourse Processes*, *11*(4), 373–412. doi:10.1080/01638538809544710

Goodman, N. (1978). *Languages of art: an approach to a theory of symbols*. New York: Bobbs-Merrill.

Goulmas, F. (1989). *The writing systems of the world*. Oxford: Basil Blackwell.

Graesser, A. C., Gordon, S. E., & Brainerd, L. E. (1992). Quest: a model of question answering. *Computers and Mathematics with Applications*, *23*(6-9), 733–745. doi:10.1080/01638539009544760

Guercin, F. (2001). Can children process complex information from different media? In J. F. Rouet (Ed.), *Multimedia learning cognitive and instructional issues* (pp. 59–64). Oxford, UK: Elsevier Science Ltd.

Guthrie, J. T. & Kirsch, I. S. (1987). Distinctions between reading comprehension and locating information in text. *Journal of Educational Psychology*, *79*(3), 220–27. doi:10.1037/0022-0663.79.3.220

Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology*, *22*(3), 353–374. doi:10.1002/acp.1411

Hamilton, R. J. (1985). A framework for the evaluation of the effectiveness of adjunct questions and objectives. *Review of Educational Research*, *55*(1), 47–85. doi:10.3102/00346543055001047

Hannon, B. & Daneman, M. (2004). Shallow semantic processing of text: an individual-differences account. *Discourse Processes*, *37*(3), 187–204. doi:10.1207/s15326950dp3703_1

Hannus, M. & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, *24*(2), 95–123. doi:10.1006/ceps.1998.0987

Hartman, H. J. (2001). *Metacognition in learning and instruction: theory, research, and practice*. Dordrecht, Boston: Kluwer Academic Publishers.

Hasselhorn, M. (1996). *Kategoriales organisieren bei kindern: zur entwicklung einer gedächtnisstrategie*. Göttingen, Seattle: Hogrefe.

Heckler, S. & Childers, T. (1992). The role of expectancy and relevancy in memory for verbal and visual information: what is incongruency? *Journal of Consumer Research*, *18*(4), 475–492. doi:10.1086/209275

Hegarty, M. (1992). Mental animation: inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1084–1102. doi:10.1037/0278-7393.18.5.1084

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, *32*(6), 717–742. doi:10.1006/jmla.1993.1036

Hegarty, M. & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition*, *22*, 411–430. doi:10.3758/BF03200867

Heiser, J. & Tversky, B. (2002). Descriptions and depictions in acquiring complex systems. In *The proceeding of the 24th annual meeting of the cognitive science society*.

Heiser, J. & Tversky, B. (2004). Characterizing diagrams produced by individuals and dyads. In T. Barkowsky (Ed.), *Spatial cognition: reasoning, action, interaction* (pp. 214–223). Berlin: Springer-Verlag.

Heiser, J. & Tversky, B. (2005). Characterizing diagrams produced by individuals and dyads. *Lecture Notes in Computer Science*, (3343), 214–226. doi:10.1007/978-3-540-32255-9_13

Heller, K. & Perleth, C. (2000). *Kft 4-12+r. kognitiver fähigkeitstest für 4. bis 12. klassen, revision*. Göttingen: Beltz Test GmbH.

Hershberger, W. A. & Terry, D. F. (1965). Delay of self-testing in three types of programed text. *Journal of Educational Psychology*, *56*(1), 22–30. doi:10.1037/h0021951

Hirtle, S. C. & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory and Cognition*, *13*(3), 208–217. doi:10.3758/bf03197683

Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2013). Dynamics of mental model construction from text and graphics. *European Journal of psychology*, *28*(4), 1105–1126. doi:10.1007/s10212-012-0156-z

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hunkins, F. P. (1968). The effects of analysis and evaluation questions on various levels of achievement. In *Meeting of american educational research association*.

Jacob, R. J. K. & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 573–605). Amsterdam: Elsevier.

Johnson, C. I. & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*(2), 178–191. doi:10.1037/a0026923

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354. doi:10.1037/0033-295X.87.4.329

Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning (2nd ed.).* (pp. 576–597). Cambridge handbooks in psychology. New York, US: Cambridge University Press.

Kammerer, Y. & Gerjets, P. (2012). Effects of search interface and internet-specific epistemic beliefs on source evaluations during web search for medical information: an eye-tracking study. *Behaviour and Information Technology*, *31*(1), 83–97. doi:10.1080/0144929X.2011.599040

Kieras, D. & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*, 255–273. doi:10.1207/s15516709cog0803_3

Kirby, J. R., Cain, K., & White, B. (2012). Deeper learning in reading comprehension. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning:*

dispositions, instruction, and learning processes (pp. 315–338). New York, US: Cambridge University Press.

Koedinger, K. R. & Anderson, J. R. (1998). Should we be worried? what the research says about gender differences in access, use, attitudes, and achievement with computers. *Educational Technology*, *4*(38), 56–61. doi:10.2190/K4U2-PWQG-RE8L-UV90

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of internal representation? evidence from studies of information retrieval times. *Memory and cognition*, *4*(3), 291–297. doi:10.1037/0096-1523.4.1.47

Kress, G. R. & Leeuwen, V. T. (1996). *Reading images: the grammar of visual design*. London, New York: Routledge.

Kultusministerkonferenz. (2005a). Bildungsstandards im fach deutsch für den primarbereich (jargansstufe 4). Standard. Wolters Kluwer Deutschland GmbH.

Kultusministerkonferenz. (2005b). Bildungsstandards im fach mathematik für den hauptschulabschluss (jahrgangsstufe 9). Standard. Wolters Kluwer Deutschland GmbH.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *COGS Cognitive Science*, *11*(1), 65–100. doi:10.1111/j.1551-6708.1987.tb00863.x

Lee, V. R. (2009). *Getting the picture: a mixed-methods inquiry into how visual representations are interpreted by students, incorporated within textbooks, and integrated into middle-school science classrooms* (Thesis).

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. Cambridge, Massachusetts: The MIT Press.

Levie, W. H. & Lentz, R. (1982). Effects of text illustrations: a review of research. *Educational Communication and Technology: A Journal of Theory, Research, and Development*, *30*(4), 195–232. doi:10.1007/BF02765184

Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In H. A. Houghton & D. M. Willows (Eds.), *The psychology of illustration vol.2: basic research* (pp. 51–85). New York: Springer-Verlag.

Levin, J. R. & Mayer, R. E. (1993). Understanding illustrations in text. In B. K. Britton, A. Woodward, & M. Brinkley (Eds.), *Learning from textbooks* (pp. 95–113). Hillsdale, NJ: Erlbaum.

Lipson, M. Y. (1982). Learning new information from text: the role of prior knowledge and reading ability. *Journal of Reading Behavior*, *14*(3), 243–261. doi:10.1080/02773949009390869

Lockhart, R. S. & Craik, F. I. (1990). Levels of processing: a retrospective commentary on a framework for memory research. *Canadian Journal of PsychologyRevue canadienne de psychologie*, *44*(1), 87–112. doi:10.1037/h0084237

Mani, K. & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, *10*(2), 181–187. doi:10.3758/BF03209220

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*(1), 49–63. doi:10.1037/0022-0663.88.1.49

Mason, L., Pluchino, P., & Ariasi, N. (2014). Reading information about a scientific phenomenon on webpages varying for reliability: an eye-movement analysis. *Educational Technology Research and Development*, *62*(6), 663–685. doi:10.1007/s11423-014-9356-3

Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? evidence from eye-movement patterns. *Computers & Education*, *60*(1), 95–109. doi:10.1007/s11145-015-9552-5

Mason, L., Tornatora, M. C., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: an eye-movement study. *Reading and Writing*, (3). doi:10.1007/s11145-015-9552-5

Mastropieri, M. A. & Scruggs, T. E. (1989). Constructing more meaningful relationships: mnemonic instruction for special populations. *Educational Psychology Review*, *1*(2), 83–111. doi:10.1007/BF01326638

Mayer, R. E. (1975). Forward transfer of different reading strategies evoked by testlike events in mathematics text. *Journal of Educational Psychology*, *67*(2), 165–169. doi:10.1037/h0077011

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, *59*(1), 43–64. doi:10.1021/ja00084a056

Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist*, *32*(1), 1–19. doi:10.1207/s15326985ep3201_1

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

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 31–49). Cambridge, New York: Cambridge University Press.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, N.J.: Pearson Merrill Prentice Hall.

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

Mayer, R. E. (2014a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd, pp. 43–71). New York: Cambridge University Press.

Mayer, R. E. (2014b). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 1–26). Cambridge, U.K.: Cambridge University Press.

Mayer, R. E. & Anderson, R. B. (1992). The instructive animation: helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*(4), 444–452. doi:10.1037/0022-0663.84.4.444

Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*(4), 638–643. doi:10.1037/0022-0663.91.4.638

McConkie, G. W., Rayner, K., & Wilson, S. J. (1973). Experimental manipulation of reading strategies. *Journal of Educational Psychology*, *65*(1), 1–8. doi:10.1037/h0034822

McNamara, D. S. (2007). *Reading comprehension strategies: theories, interventions, and technologies*. New York: Lawrence Erlbaum Associates.

Merriënboer, J. J. G. v. & Kirschner, P. A. (2007). *Ten steps to complex learning: a systematic approach to four-component instructional design*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Mosenthal, P. B. & Kirsch, I. S. (1991). Toward an explanatory model of document literacy. *Discourse Processes*, *14*(2), 147–80. doi:10.1080/01638539109544780

Oakhill, J. & Yuill, N. (1996). Higher order factors in comprehension disability: processes and remediation. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: processes and intervention* (pp. 69–92). Mahwah, NJ: Erlbaum.

Ornstein, P. A., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personally experienced event: implications for testimony. *Applied Cognitive Psychology*, *6*(1), 49–60. doi:10.1002/acp.2350060103

Ozgungor, S. & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology*, *96*(3), 437–443. doi:10.1037/0022-0663.96.3.437

Paivio, A. (1986). *Mental representations: a dual coding approach*. New York, Oxford [Oxfordshire]: Oxford University Press , Clarendon Press.

Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, *44*, 267–278. doi:10.1002/acp.3100

Parkhurst, D. J. & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16*, 125–154. doi:10.1049/iet-cvi.2009.0067

Parkhurst, P. E. (1974). *Assessing the effectiveness of self-paced visualized instruction: a multifactor analysis on five different educational tasks* (Thesis).

Passig, D. & Levin, H. (1999). Gender interest differences with multimedia learning interfaces. *Computers in Human Behavior*, *15*(2), 173–183. doi:10.1016/S0747-5632(99)00016-3

Peeck, J. (1974). Retention of pictorial and verbal content of a text with illustrations. *Journal of Educational Psychology*, *66*(6), 880–888. doi:10.1037/h0021531

Perfetti, C. (2010). Decoding, vocabulary, and comprehension: the golden triangle of reding skill. In M. G. McKeown & L. Kucan (Eds.), *Bringing reading research to life*. New York: Guilford Press.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Lawrence Erlbaum.

Playfair, W. (1786). *The commercial and political atlas*. London: Corry.

Pöhnl, S. & Bogner, F. X. (2012). Learning with computer-based multimedia: gender effects on efficiency. *Journal of Educational Computing Research*, *47*(4), 387–407. doi:10.2190/EC.47.4.c

R Core Team. (2015). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from http://www.R-project.org/

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. doi:10.1037/0033-2909.124.3.372

Rayner, K. & Castelhano, M. (2007). Eye movements. *Scholarpedia*, *2*(10), 3649. doi:10.4249/scholarpedia.3649

Reid, D. J., Briggs, N., & Beveridge, M. (1983). The effect of picture upon the readability of a school science topic. *British Journal of Educational Psychology*, *53*(3), 327–335. doi:10.1111/j.2044-8279.1983.tb02565.x

Rickards, J. P. (1979). Adjunct postquestions in text: a critical review of methods and processes. *Review of Educational Research*, *49*(2), 181–196. doi:10.3102/00346543049002181

Rickards, J. P. & Denner, P. (1978). Inserted questions as aids to reading text. *Instructional Science*, *7*(3), 313–346. doi:10.1007/BF00120936

Rickards, J. P. & Di Vesta, F. J. (1974). Type and frequency of questions in processing textual material. *Journal of Educational Psychology*, *66*(3), 354–362. doi:10.1037/h0036349

Robertson, S. P., Black, J. B., & Lehnert, W. G. (1985). Misleading question effects as evidence for integrated question understanding and memory search. In A. C. Graesser & J. B. Black (Eds.), *The psychology of questions* (pp. 191–218). Mahwah, NJ: Erlbaum.

Rothkopf, E. Z. (1963). Some conjectures about inspection behavior in learning from written sentences and the response mode problem in programed self-instruction. *Journal of Programed Instruction*, *2*(4), 31–47. doi:10.1037/h0024117

Rothkopf, E. Z. (1964). Learning and the educational process, selected papers from the research conference on learning and educational process, held at stanford university, june 22-july 31, 1964. In J. D. Krumboltz (Ed.), *Research conference on learning the educational process* (pp. 193–221). Chicago: Rand McNally.

Rothkopf, E. Z. (1966). Learning from written instructive materials: an exploration of the control of inspection behavior by test-like events. *Journal American Educational Research Journal*, *3*(4). doi:10.3102/00028312003004241

Rothkopf, E. Z. & Billington, M. J. (1974). Indirect review and priming through questions. *Journal of Educational Psychology*, *66*(5), 669–679. doi:10.1037/h0037488

Rothkopf, E. Z. & Bisbicos, E. E. (1967). Selective facilitative effects of interspersed questions on learning from written materials. *Journal of Educational Psychology*, *58*(1), 56–61. doi:dx.doi.org/10.1037/h0024117

Rothkopf, E. Z. & Coke, E. U. (1963). Repetition interval and rehearsal method in learning equivalences from written sentences. *Journal of Verbal Learning and Berbal Behavior*, *2*, 406–16. doi:10.1016/S0022-5371(63)80041-9

Rouet, J. F. (2001). Designing multimedia systems for learning: some lessons and further issues. In J. F. Rouet, J. J. Levonen, & A. Biardeau (Eds.), *Multimedia learning cognitive and instructional issues*. Oxford, UK: Elsevier Science Ltd.

Rouet, J. F. (2006). Question answering and document search. In J. F. Rouet (Ed.), *The skills of document use: from text comprehension to web-based learning* (pp. 93–121). Mahwah, NJ: Erlbaum.

Rouet, J. F. & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19–52). Charlotte, NC, US: IAP Information Age Publishing.

Rouinfar, A., Agra, E., Larson, A. M., Rebello, N. S., & Loschky, L. C. (2014). Linking attentional processes and conceptual problem solving: visual cues facilitate the automaticity of extracting relevant information from diagrams. *Frontiers in Psychology*, *5*, 1094. doi:10.3389/fpsyg.2014.01094

Rusted, J. & Coltheart, M. (1979). Facilitation of children's prose recall by the presence of pictures. *Memory & Cognition*, *7*(5), 354–359. doi:10.3758/BF03196939

Schmandt-Besserat, D. (1992). *Before writing: from counting to cuneiform*. Austin: University of Texas Press.

Schneider, T. (2007). Does the effect of social origins on educational participation change over the life course in germany? social inequalities in entering the academic school type and dropping out. *Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften*, *127*(1), 21–31.

Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2nd, pp. 72–103). Cambridge, U.K.: Cambridge University Press.

Schnotz, W. & Baadte, C. (2014). Surface and deep structures in graphics comprehension. *Memory & Cognition*. doi:10.3758/s13421-014-0490-2

Schnotz, W. & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*(2003), 141–156. doi:10.1016/S0959-4752(02)00017-8

Schnotz, W. & Kulhavy, R. W. (1994). *Comprehension of graphics*. Amsterdam, New York: North-Holland.

Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. R. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology*, *106*(4), 974–989. doi:10.1037/a0037054

Schnotz, W., Mengelkamp, C., Baadte, C., & Hauck, G. (2014). Focus of attention and choice of text modality in multimedia learning. *European Journal of Psychology of Education*, *29*(3), 483–501. doi:10.1007/s10212-013-0209-y

Schnotz, W., Ullrich, M., Hochpöchler, U., Horz, H., McElvany, N., Schroeder, S., & Baumert, J. (2011). What makes text-picture integration difficult? a structural and procedural analysis of textbook requirements. *Ricerche di Psicologia*, *1*, 103–135. doi:10.3758/s13428-014-0528-1

Schnotz, W., Ullrich, M., Horz, H., McElvany, N., Schröder, S., Ullrich, M., ... Richter, T. (2010). Das bite-projekt: integrative verarbeitung von bildern und texten in der sekundarstufe. projekt bite. *Zeitschrift für Pädagogik*, *56*, 143–153. doi:10.3389/fnins.2014.00385

Schüler, A., Arndt, J., & Scheiter, K. (2015). Processing multimedia material: does integration of text and pictures result in a single or two interconnected mental representations? *Learning and Instruction*, *35*, 62–72. doi:10.1016/j.learninstruc.2014.09.005

Seretny, M. L. & Dean, R. S. (1986). Interspersed postpassage questions and reading comprehension achievement. *Journal of Educational Psychology*, *78*(3), 228–229. doi:10.1037/0022-0663.78

Shah, P. & Freedman, E. G. (2009). Bar and line graph comprehension: an interaction of top-down and bottom-up processes. *Topics in Cognitive Science*, *2009*, 1–19. doi:10.1111/j.1756-8765.2009.01066.x

Shavelson, R. J., Berliner, D. C., Ravitch, M. M., & Loeding, D. (1974). Effects of position and type of question on learning from prose material: interaction of treatments with individual differences. *Journal of Educational Psychology*, *66*(1), 40–48. doi:10.1037/h0035813

Slavin, R. E. (1997). *Educational psychology: theory and practice*. Boston: Allyn and Bacon.

Small, J. P. (1997). *Wax tablets of the mind: cognitive studies of memory and literacy in classical antiquity*. New York: Routledge.

Smith, K. U. & Smith, M. F. (1966). *Cybernetic principles of learning and educational design*. New York: Holt, Rinehart, and Winston.

Snowden, R. J., Thompson, P., & Troscianko, T. (2012). *Basic vision: an introduction to visual perception* (2nd). Oxford, New York: Oxford University Press.

Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: using if-then plans to facilitate cognitive processing. *Learning and Instruction*, *35*, 1–15. doi:10.1016/j.learninstruc.2014.09.002

Stolurow, L. M. (1961). *Teaching by machine*. Michigan: University of Michigan Library.

Sweller, J., Ayres, P. L., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Tal, N. F., Siegel, L. S., & Maraun, M. (1994). The role of question type and reading ability in reading comprehension. *Reading and Writing*, *6*(4), 387–402. doi:10.1007/BF01028850

Taylor, H. A. & Tversky, B. (1992). Descriptions and depictions of environments. *Memory and Cognition*, *20*, 483–496. doi:10.3758/BF03199581

Tenenberg, M. (1969). *Effects of different study instructions on learning from written materials*. University of California.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Tilling, L. (1975). Early experimental graphs. *British Journal of Historical Science*, *8*. doi:S00070874000014229

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire: Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison Wesley.

Tulving, E. & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review Psychology*, *21*, 437–84. doi:10.1146/annurev.ps.26.020175.001451

Tversky, B. (1991). Distortions in memory for visual displays. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 61–75). London: New York : Taylor & Francis.

Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), *Understanding images*. New York: Springer-Verlag.

Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79–112). Cambridge, Mass.: MIT Press.

Tversky, B. & Lee, P. (1999). Pictorial and verbal tools for conveying routes. In C. Freksa & D. Mark (Eds.), *Spatial information theory. cognitive and computational foundations of geographic information science* (Chap. 4, pp. 51–64). Lecture Notes in Computer Science. Springer Berlin Heidelberg. doi:10.1007/3-540-48384-5_4

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, *57*(4), 247–262. doi:10.1006/ijhc.2002.1017

Tversky, B., Zacks, J., Lee, P., & Heiser, J. (2000). Lines, blobs, crosses and arrows: diagrammatic communication with schematic figures. *Lecture Notes in Computer Science*, (1889), 221–230. doi:10.1007/3-540-44590-0_21

van den Broek, P., Tzeng, Y., Risden, K., Trabasso, T., & Basche, P. (2001). Inferential questioning: effects on comprehension of narrative texts as a function of grade and timing. *Journal of Educational Psychology*, *93*(3), 521–529. doi:10.1037/0022-0663.93.3.521

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

van Gog, T. & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, *2*(20), 95–99. doi:10.1016/j.learninstruc.2009.02.009

van Oostendorp, H. & Goldman, S. (1999). *The construction of mental representations during reading*. Mahwah, NJ: Erlbaum.

Vandeberg, L., Bouwmeester, S., Bocanegra, B. R., & Zwaan, R. A. (2013). Detecting cognitive interactions through eye movement transitions. *Journal of Memory and Language*, *69*(3), 445–460. doi:10.1016/j.jml.2013.05.006

Vidal-Abarca, E. & Sanjose, V. (1998, June). Levels of comprehension of scientific prose: the role of text variables. *Learning and Instruction*, *8*(3), 215–233. doi:10.1016/S0959-4752(97)00020-0

Wagner, I. & Schnotz, W. (in press). Entwicklung der text-bild-integrationskompetenz im 4.schuljahr an grundschulen. *Unterrichtswissenschaft*.

Wainer, H. (1992). *Understanding graphs and tables*. Princeton, N.J.: Educational Testing Service.

Wainer, H. (2005). *Graphic discovery: a trout in the milk and other visual adventures*. Princeton, NJ: Princeton University Press.

Wainer, H. & Thissen, D. (1981). Graphical data analysis. *Annual Review of Psychology*, *32*, 191–241. doi:10.1146/annurev.ps.32.020181

Washbourne, J. N. (1929). The use of questions in social science material. *Journal of Educational Psychology*, *20*, 321–354. doi:10.1037/h0075982

Wass, S., Forssman, L., & Leppänen, J. (2014). Robustness and precision: how data quality may influence key dependent variables in infant eyetracker analyses. *Infancy*, *19*(5), 427–460. doi:10.1111/infa.12055

Watts, G. & Anderson, R. (1971). Effects of three types of inserted questions on learning from prose. *Journal of Educational Psychology*, *62*(5), 387–394. doi:10.1037/h0031633

Wedel, M. & Pieters, R. (2008). A review of eye-tracking research in marketing. *Review of Marketing Research*, *4*, 123–147. doi:10.1108/S1548-6435(2008)0000004009

Winn, W. (1987). Charts, graphs, and diagrams in eduational materials. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illusions* (Vol. 1, pp. 152–198). New York: Springer-Verlag.

Winn, W. (1989). The role of graphics in training documents: toward an explanatory theory of how they communicate. *IEEE Trans. Profess. Commun. IEEE Transactions on Professional Communication*, *32*(4), 300–309. doi:10.1109/47.44544

Wixson, K. K. (1983). Postreading question-answer interactions and children's learning from text. *Journal of Educational Psychology*, *75*(3), 413–23. doi:10.2307/1169962

Yang, F.-Y., Huang, R.-T., & Tsai, I. J. (2014). The effects of epistemic beliefs in science and gender difference on university students' science-text reading: an eye-tracking study. *International Journal of Science and Mathematics Education*, 1–26. doi:10.1007/s10763-014-9578-1

Yates, F. A. (1969). *The art of memory*. New York: Penguin.

Zacks, J. & Tversky, B. (1999). Bars and lines: a study of graphic communication. *Memory and Cognition*, *27*(6), 1073–1079. doi:10.3758/BF03201236

Zhao, F., Schnotz, W., Wagner, I., & Gaschler, R. (2014). Eye tracking indicators of reading approaches in text-picture comprehension. *Frontline Learning Research*, *2*(4), 46–66. doi:10.14786/flr.v2i4.98

Zwaan, R. & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–85. doi:10.1037/0033-2909.123.2.162

Declaration

I declare that, to the best of my knowledge, the research described herein is original except where the work of others is indicated and acknowledged, and that the thesis has not, in whole or in part, been submitted for any other degree at this or any other university.

____

Fang Zhao

Landau in der Pfalz

22 July 2015

Curriculum Vitae

Personal Details

Name: Fang Zhao

Contact: zhao@uni-landau.de

Education

2012.11-2015.11: PhD in Psychology, *with a concentration in eye tracking*, PhD thesis: “Reading Strategy in Multimedia Learning”, University of Koblenz-Landau, Germany

2014.10-2014.12: Visiting scholar at Prof. Dr. Mary Hegarty’s laboratory, in Psychology and Brian Science, University of California, Santa Barbara, USA

2010.09-2012.01: Master of Arts in Translation Studies, *with a concentration in translation technology and theory*, Master thesis: “Chinese Students’ Needs and Translation Programmes in the UK”, Durham University, UK

2006.09-2010.06: Bachelor of Arts in English language, *with a concentration in linguistics*, Bachelor thesis: “The Enlightment from Kraschen’s Input Theory on Foreign Language Teaching Approaches”, Henan University, China

2003.09-2006.06: Anyang No.1 High School, *with a concentration in science*, China

2000.09-2003.06: Huanbei Middle School, China

1994.09-2000.06: Yong’an Jie Primary School, China