Inauguraldissertation

zur Erlangung des akademischen Grades eines Doktors

der Philosophie der Universität Mannheim

Intensive vs. Extensive Reading

A Study on the Impact of L2 Reading Interventions in 5th and 6th Grade

von Frauke Fried

Disputation am 10.04.2024

for all students

Table of Content

Tables

Figures

1 Introduction

Students presenting the books from the English library to a teacher colleague at the open house day.

The fact that reading has multiple positive impacts is somewhat obvious and common sense. Research has shown that children who read a lot, develop a positive attitude towards reading and towards their own reading ability (Shapiro & Whitney, 1997). Hence, from an educational perspective, reading cannot be fostered early enough at home and at school. The home environment might be a more significant factor than school programs (Shapiro & Whitney, 1997), yet teachers can help to encourage intrinsic reading motivation that is maybe then carried on at home. For example, after the second language (L2) reading interventions in this study, one student had improved immensely in English and her English teacher asked her what she had changed. She answered that during the reading intervention she had noticed that reading in English was fun and told her father about it. They then searched for English literature at home and sat together every evening, read together and discussed what they had read. This student would probably not have improved that much by only participating in the English reading intervention program, but also might not have received this initial spark without it.

According to research, engagement in leisure reading has a positive influence on children's written and oral expression, general reading ability and creative imagination (Koolstra & van der Voort, 1996). Later on in life, engagement in reading is associated with a higher socioeconomic status (Schutte & Malouff, 2004) and at an older age it can be useful to limit the risk of cognitive deterioration (Verghese et al., 2003). Stanovich (2000) described reading and the positive effects it has on vocabulary growth with the Matthew effect:

> The effect of reading volume on vocabulary growth, combined with the large skill differences in reading volume, could mean that a "rich-get-richer" or cumulative advantage phenomenon is almost inextricably embedded within the developmental course of reading progress. The very children who are reading well and who have good vocabularies will read more, learn more word meanings, and hence read even better. Children with inadequate vocabularies – who read slowly and without enjoyment – read less, and as a result have slower development of vocabulary knowledge, which inhibits further growth in reading ability. (Stanovich, 2000, p. 184)

Therefore, if we want more children on the winning side of this effect, we need to get them engaged in reading. The question is, whether these positive effects of reading also hold for L2 reading, either for children with another first language (L1) and German L2 or when learning further languages at school. If that is the case, extensive reading could be one important source to obtain the necessary amount of language input needed to acquire the vocabulary and grammar of a language. Previous research on the effects of (extensive) L2 reading interventions on L2 language competence have shown various positive relationships between the two that will be discussed further in chapter two, but evidence is still very sparce, leaving L2 researchers with more questions than answers at the moment (Koda, 2005; Nation & Waring, 2020). One aim of this study is to contribute an answer to the question of how L2 reading interventions for early learners of English in 5th and 6th grade of a German secondary school influence their L2 language competence gains.

In this dissertation, I am going to analyze what effects *intensive and extensive English reading interventions* have on student's *English abilities* and on their *goal orientation* and *self-beliefs* concerning the school subject English. Moreover, I am going to investigate to what extent demographic factors, initial L2 competence, L1 reading experience, goal orientation and self-beliefs moderate the effects of reading interventions on L2 English competence gains. I am going to shortly describe the study and the main psychological concepts needed in the next paragraphs to give an overview. They will be presented and discussed in detail in the following chapters.

In this longitudinal, experimental study, students of a German secondary school were randomly assigned to different reading and non-reading intervention classes. The reading classes were taught as intensive reading (reading books and working on corresponding tasks and exercises) and extensive reading classes (reading as much as one can at a low level and with high speed). Additionally, there were English classes that did not involve reading and non-English classes that functioned as control groups. Tests measured the students' English abilities. These included a cloze test, a preposition test, a fictional reading comprehension test and a non-fictional reading comprehension test that were created for the purpose of this study. Additionally, the psychological constructs goal orientation and self-concept were measured using established scales.

Goal orientation theory investigates the motives that people have for pursuing certain goals (Spinath & Schöne, 2003). These can differ in nature and are therefore commonly divided into learning goal orientation, i.e. the goal is to enhance personal competence, and performance goal orientation, i.e. the goal is to show or hide personal competence. Performance goal orientation is commonly divided into performance-approach and performance-avoid goal orientation. The SELLMO scales (Spinath, Stiensmeier-Pelster, Schöne, & Dickhäuser, 2012) used in this experiment additionally measure a fourth goal, i.e. work avoidance - the goal to invest as little work as possible.

The academic *self-concept* belongs to personal self-beliefs and is, for example, defined as the entirety of thoughts about personal abilities in academic performance situations (Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, 2003). The self-concept can be measured as a general self-concept (similar or equal to the concept of self-esteem), a domain specific concept (as most commonly done), or as a task specific concept (similar or equal to self-efficacy). The SESSKO scales (Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, 2012) used in this study measure the domain specific self-concept and were adapted to measure self-concept in the subject English. Self-concept is usually seen as a result of social comparisons (Marsh et al., 2019), but Schöne et al.'s (2012) scales make it possible to differentiate between social, individual and criterial comparisons.

This study will show that these psychological factors deliver an incremental contribution to understanding effects of reading interventions on academic achievement and development of language competence. Up until now, debates on the effect of extensive reading interventions have mostly evolved around possible positive effects of extensive reading on language proficiency aspects like vocabulary growth or reading speed (Nation & Waring, 2020); little to no research has been done concerning the impacts of reading interventions on goals and self-beliefs (see section 2.3, especially section 2.3.5). Additionally, extensive and intensive reading are often juxtaposed in studies, sometimes promoting extensive reading as a better alternative to intensive reading. This study, in contrast, argues that intensive and extensive reading is not an either/or issue. On the basis of cognitive and psychological theories, one can assume that the reading processes underlying these two reading types differ significantly (see section 2.2). Therefore, students might profit differently from these interventions depending on their reading and language experience but also on psychological factors like self-beliefs and goals. This is a decisive aspect that, to my knowledge, has neither been discussed nor shown in the field of extensive reading research so far. The results of this study are a strong appeal for taking students' experience and personality more into account when assessing the impact of didactic measures on academic achievement and implementing these in the foreign language classroom. Especially in the light of equal educational opportunities, this factor should not be ignored.

This thesis is organized as follows: The following chapter 2 provides an overview of reading processes, reading models and differences of L1 and L2 reading. Additionally, concepts and implementations of extensive reading programs are discussed. One aspect of extensive reading is the "fun factor". Motivational (reading) models are therefore discussed, followed by and linked to an overview of the psychological aspects goal orientation and self-concept. Motivation and personality of the reader have only marginally been at the focus of research in the area of extensive reading so far; this is a gap that I am going to address in this study. Differences between intensive and extensive reading are mapped out and hypotheses are developed on this basis. These are summarized at the end of chapter 2.

Chapter 3 provides an overview of the overall study and the research methodology, e.g. participants, design, procedure of the interventions, material and variables. A substantial section of this chapter covers the description of the quality of the reading comprehension and language tests that were developed and piloted for this study. Additionally, the procedure for data analyses is explained.

The results of various analyses are then presented in chapter 4 and discussed in chapter 5 on the basis of the hypotheses developed in chapter 2. Chapter 6 provides suggestions for the implementation of similar interventions at schools on the basis of theoretical implications and the results and experiences from this study. It discusses didactic consequences and suggests questions for further research. A summary and conclusion of the study, results and discussion is presented in chapter 7, followed by the sources, a list of the books of the reading library and the appendix with additional tables and graphs of the statistical analyses.

2 Theoretical Background and Previous Research

"L2 reading is not just someone learning to read in another language; rather, L2 reading is a case of learning to read with languages." (Grabe, 2009, p. 129)

2.1 Reading Behavior and the Necessity of Reading Interventions

The *Medienpädagogischer Forschungsverbund Südwest* conducts regular studies questioning children and teenagers about their media consumption and leisure time behavior. There are different studies available: The miniKIM-study presents results of children age 2-5 and started in 2012, the KIM-study focuses on children aged 6-13 and started in 1999, the JIM-study presents the results of 12-19 year-olds and started in 1998 and the FIM-study started in 2016 and focuses on families[1]. Since the students in this study are 9 to 12 years old, the focus of this section lies on the KIM-study results.

The results of the KIM-study (Medienpädagogischer Forschungsverbund Südwest, 2018) are actually quite alarming, especially if you keep the Matthew effect (Stanovich, 2000) in mind. Figure 2.1 depicts the percentage of children who read books (almost) every day (orange), once or several times a month (dark red), less than once a month (yellow) and never (light red). Depending on the age group, 12-26% of the children questioned never read a book in their leisure time. If we add the number of children who only read a book less than once a week, this group of poor readers increases to about 50%.

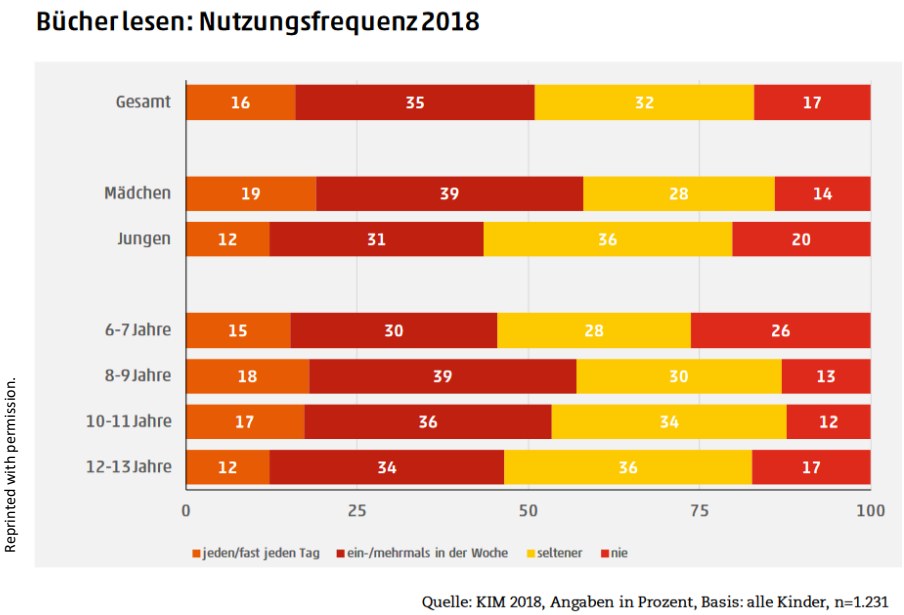


Figure 2.1 Frequency of Book Reading - Children in Germany

[1] https://www.mpfs.de (last accessed 20.02.2021)

One assumption at hand is that television, internet and gaming have replaced regular book reading. Indeed, Koolstra and van der Voort (1996), who conducted a longitudinal study over two years with 1050 Dutch elementary school children, found negative effects of television on children's leisure-time reading. The two factors that enforced the reduction of book reading were a decreasing positive attitude towards books and a decreasing ability of children to concentrate on book reading, probably because they had become more acquainted with fast moving pictures that are easier to consume. Especially when taking the last factor of lacking concentration into account, you would expect this effect to increase even more with the increase of video gaming and internet platforms like Youtube, Instagram and TikTok.

The KIM-study has been run since 1999 and, therefore, can give us insight into the reading development of children over the past 20 years. I summarized the results in figure 2.2: From 1999-2016 children were asked how much they like to read. The number of children who answered that they really like to read was constant at about 15% (green line), about the same number answered that they read daily (dotted green line). Therefore, there always seems to be a group of children who are very engaged readers. First results of the JIM-Study 2020 with older children indicate that in the Covid-19 year 2020 this group neither decreased nor increased (Medienpädagogischer Forschungsverbund Südwest, 2020). Thus, I assume that these are not the children we have to address with reading programs, because they are committed to reading already. On the other end of the spectrum, the group that does not like to read increased from 4% in 1999 to 9% in 2005 (light red line). In 2006 a new category "never read" was introduced (dark red), making up 16-20% of the children from that point on. Together with the "do not like to read" answers, these results have also not changed much in the past 15 years, alternating at about 20%. To reach these reluctant readers with programs at school is probably very difficult, especially when considering socio-economic and cognitive factors. Of special interest to me is the group that rather likes to read (yellow line), where a decreasing trend is visible, maybe due to new technologies. This decreasing trend is parallel to the actual amount of reading ("read a few times a week" – yellow dotted line). These children could be possible targets of reading interventions, because they are probably most likely to be encouraged by reading programs to keep up reading or become more interested in (L2) reading again. These are all results about German L1 reading, I was not able to find statistics on the amount of English L2 reading of children in Germany and can only assume that this is mearly a fraction of their total reading. The students in this study were asked how many English books they possess and how many English books, apart from those in the study, they have read. The students stated to possess an average of 4.63 English books (SD = 9.25, N = 230) and to have read an average of 4.11 (SD = 9.25, N = 296) English books in their life so far.

Figure 2.2 Reading Behavior 1999 - 2018

The quality of reading is, unfortunately, not assessed in the KIM and JIM surveys. The Jim-Study (Medienpädagogischer Forschungsverbund Südwest, 2020) asked children to name their favorite books, which are classics like *Harry Potter*, *Greg's Diary* and *Die drei Fragezeichen*, but it is unclear how much time a day children spend reading these books and how intense their reading engagement is.

Moreover, it is unclear if students who spend time with social media and online media consume mostly pictures and videos, or also read blogs and articles. “Reading” is often named as one activity in contrast to playing on the computer, messaging or spending time on Instagram. Yet, all of these activities involve reading: reading text messages, Instagram stories or descriptions in video games. When interpreting reading as a broader term, reading is something that could have even increased during the Covid-19 pandemic, because, at least at first, many school assignments were in written form and the only way to communicate and have social participation was online. I assume that with “reading” the authors of the surveys above mean “reading books”, but this is not clarified in all surveys.

To become a skilled reader in the sense of enhancing lower level, automatic reading processing skills, the type of text read is probably not as important, as long as it delivers enough rich language input. But when learning to understand and follow a complex narrative structure, an Instagram story might not be as rich in input as a traditional book. There is not a lot of research on the differences of reading strategy use and development depending on text types. Yet, there already seems to be a difference between print and digital reading of the same text (Jong & Bus, 2002; Singer & Alexander, 2017) with better results for print reading. Digitalization has opened new opportunities of electronic reading, but more multi-tasking might also lead to shorter attention spans, which is why Nation and Waring (2020) argue that extensive reading programs with traditional books in schools might be more relevant than ever. The following sections revolve around the question what processes are involved when reading in the L1 and L2. Reading is seen here as offline reading of printed books.

2.2 What is Reading?

Reading is a core skill needed to be successful in our modern society (Grabe, 2009). Due to progressing globalization, digitalization and collaboration across cultures, literacy not only in the native L1, but also in English as a lingua franca and language of trade, business and the internet has been gaining importance over the past century. This calls for training reading and writing abilities of students from an early age on. Yet, as the previous paragraph and latest educational studies (Stanat et al., 2022) have shown, students’ reading skills have decreased. The following sections give an overview of different reading skills and differences in developing L1 and L2 reading skills.

2.2.1 Defining Reading

Looking up the entry *to read* in the Oxford Advanced Learner’s Dictionary leads to the explanation “to look at and understand the meaning of written and printed words or symbols” (Oxford University Press, 2020). This definition just covers a narrow view of literacy, namely the ability to

decode symbols, but literacy, in the sense of functional literacy, also involves the suitable use of this ability in different contexts (Grabe, 2009; Hedgcock & Ferris, 2018; Hudson, 2007). These different levels of demands on the reader and the ability of being able to adjust processing to meet these demands is a key factor of successful reading and are represented in the definition of the International Literacy Association: "Literacy is the ability to identify, understand, interpret, create, compute, and communicate using visual, audible, and digital materials across disciplines and in any context." ("International Literacy Association," 2021)

With literacy defined as "the ability to identify and understand", the decoding process of reading as a phonetic and lexical decoding process and proposition building by constructing meaning out of letters, words, phrases and sentences, is addressed. The following factors "interpret, create, compute, and communicate", already address reading and writing as a knowledge building task. They also summarize the various purposes of reading and writing. Looking at the academic purposes of reading alone shows how different reading purposes could be. Grabe (2009), for example, names multiple purposes: reading to search for information, for a quick understanding, to learn, to integrate information, to evaluate, critique and use information and for general comprehension.

Hudson (2007) lists five features that can be used to classify different reading acts:

1. Reasons: motivation of the reader to engage in reading (e.g. information seeking, pleasure)
   In this study the reason for reading differs between the intensive and extensive reading interventions. Whereas extensive reading, in an idealized version, is supposed to be pure pleasure for the students, intensive reading involves additional tasks and therefore not only reading for pleasure but also to fulfill the task requirements.
2. Media: the physical entity of the text (e.g. books, newspapers, notices, forms, web-content, …).
   In this study the media provided to the students were physical books – some board books, some hard cover, most paper back.
3. Content: material of the text (e.g. fiction, songs, plays, academic articles, t-shirt messages, postcard slogans, poetry, …).
   The content of the books was very diverse. There were fictional (long and short) and non-fictional books, plays, poetry, biographies and comics available.
4. Structure and Form: charts, isolated lines, narrative, …
   The structure and form of the books also varied strongly in this study. There were books with many illustrations and isolated lines of texts and words, text that was divided by headings and subheadings, especially in non-fictional texts, and books with longer narratives.

5. Strategies and Skills: reading techniques – dependent on reader abilities and purpose (e.g. skimming – reading rapidly for an overview, scanning – reading rapidly to find specific information, …).

   While reading, especially to fulfill the tasks in the intensive reading interventions, the students probably engaged scanning to find answers to the given questions and in-depth reading.

The combination possibilities between all of these aspects reveals how different reading processes can be, which reflects in the literacy definition above in the words "across disciplines and in any context". Therefore, a person can and will develop different and multiple literacies and applying these advantageously is seen as mastering different literacy events (Hedgcock & Ferris, 2018).

The following sections will focus on cognitive aspects of reading. Motivational impacts (here self-concept and goal orientation) and social influences (here intensive and extensive reading intervention settings) are discussed in chapter 2.3.

2.2.2 Reading Skills and Processes

The perspective on differs strongly between different sciences and trends (Koda, 2007). Reading can be seen, for example, as an indivisible process. However, a common approach to look at reading is by defining subskills and strategies needed to develop proficient reading skills (Grabe, 2009, 2014; Khalifa & Weir, 2009; Koda, 2005, 2007). From a pedagogical perspective, this view makes it possible to foster student's reading with targeted and specific exercises (Grabe, 2009), to assess certain subskills with attuned tasks or to analyze which subskills certain reading exercises emphasize or presuppose (Weir & Khalifa, 2008) and to investigate L1 and L2 reading differences and transfer of subskills (Koda, 2005, 2007). These different subskills interact in various ways and should therefore not be seen in isolation, but this differentiated view can help to attain a better understanding on processing differences and difficulties.

From the large variation of skills needed to successfully engage in literacy events, Hudson (2007) named eight variables that he sees crucial for a successful reading process: grapheme/phoneme recognition, phonological representation, syntactic structure, background knowledge, processing strategies, text structure understanding, vocabulary and the context of the reading act. Grabe (2009) named features that define reading. For him, reading is a rapid, efficient, comprehending, interactive, strategic, flexible, purposeful, evaluative, learning and linguistic process. When comparing these two approaches Hudson (2007) focused on skills that readers need to acquire and make use of, whereas Grabe (2009) focused on attributes that successful reading can be measured by. For example, readers

with precise and skilled grapheme-phoneme recognition will be able to read faster (speed) (Koda, 2007). When words are recognized highly automatically, this is called a high sight vocabulary (Day & Bamford, 1998). A high sight vocabulary is not sufficient for comprehension yet, the information first needs to be processed, for example with the use of background knowledge, world knowledge, knowledge of the topic, etc. In other words, the information is processed according to and integrated into existing schemas (Day & Bamford, 1998; Grabe, 2009; Koda, 2007).

These various reading processes are commonly divided into higher-level and lower-level processing skills or strategies[2] (Grabe, 2009, 2014; Hudson, 2007). Lower-level processing skills „form a group of skills that have the potential to become strongly automatized, and this automatization of lower-level skills is a requirement for fluent reading" (Grabe, 2009, p. 21). These skills are often carried out simultaneously and include word recognition, syntactic parsing and semantic proposition encoding. "Higher-level processes generally assume that the reader can direct attentional resources to these component skills. […] At the same time, many aspects of these higher-level component abilities can, and often are, carried out in automatic mode except when difficulties arise or when specific goals highlight attentional resource needs." (Grabe, 2009, p. 39)

Even though well-developed lower-level processing skills and high sight vocabulary do not equal high comprehension, students who have highly automatized lower-level processes like grapheme-phoneme recognition will probably show higher comprehension, because they can use their limited processing capacity for high-level processing strategies and reading skills (Day & Bamford, 1998). Therefore, more fluent readers will be able to flexibly interact with the text according to the purpose and come to a better understanding. Extensive reading can most likely deliver the beneficial practice for developing these skills.

Higher- and lower-level processing skills and strategies both need to be taken into account when discussing what reading is and how it can be acquired, because readers cannot just use either low-level processing skills or only high-level processing strategies. On the one hand, without low-level processing skills, reading as making sense out of letters, words and sentences would not take place. Without high-level processing strategies, on the other hand, comprehension would be impaired and the readers would read words and sentences without being able to make any sense out of them, which also would not suit our understanding of reading (Grabe, 2009). However, not all of these skills are used to the

[2] Skills and processes are not always clearly distinguished when talking about reading, see also Hudson (2007); Khalifa and Weir (2009) for further discussion. The word "skills" is often used for automatic processes, whereas "strategies are cognitive processes that are open to conscious reflection but that may be on their way to becoming skills" (Grabe, 2009, p. 221). In the previous sections I refer to lower-level skills, highlighting their automatized and subconscious character and higher-level strategies, highlighting that they are predominantly deliberate, purposeful and goal orientated. This of course does not mean that all higher-level processes are strategies and all lower-level processes are skills, since this obviously dependents to a great extent on factors like proficiency and experience.

same extent, with the same consciousness and at the same speed while reading. Factors that determine processing while reading are for example the proficiency of the reader and the purpose of reading.

2.2.3 Reading Models

The way higher- and lower-level processing strategies and skills interact and lead to understanding is presented and discussed in various reading models. These can be divided into bottom-up, top-down, interactive and interactive compensatory models (Grabe, 2009).

Bottom-up models view reading as a process in which encoding functions in a linear order (Grabe, 2009; Hudson, 2007). Taking the illustration of the interactive compensatory model in figure 2.3 at the end of this section, proposed by Weir and Khalifa (2008), as an example, this would mean that the reading process would move strictly from the bottom box to the top box: A reader would start by decoding visual input into letters and words, then access the mental lexicon and next add syntactic information. Moreover, this bottom-up process would not be influenced by other incoming information, like background knowledge and goals. Yet, these aspects clearly influence the reading process (Grabe, 2009; Hedgcock & Ferris, 2018; Khalifa & Weir, 2009), therefore bottom-up models are insufficient for modeling reading processes.

Top-down models take the opposing view to bottom-up models. Here the reading process is seen as mainly influenced by expectations and goals of the text and is processed accordingly (Grabe, 2009; Hudson, 2007). In figure 2.3 this would mean that the boxes on the left and right strongly influence the reading process in the center, which only functions to verify or falsify the expectations and hypotheses generated about the text. Additionally, syntactic knowledge and inferencing help to identify words and process sentences (Grabe, 2009; Hudson, 2007). This already shows the limitations of pure top-down models: lower-level processes often occur so fast and automatic in skilled readers that they are not influenced by higher-level processes, because this would just delay the reading process. Therefore, strict top-down processing would be too slow and not account for rapid word recognition or wrong processing of syntactic ambiguities.

Interactive models combine bottom-up and top-down processing. They assume that higher-level processes support lower-level processing to heighten reading proficiency (Grabe, 2009). This view also has its flaws, because just like bottom-up models, these models cannot account for very fluent reading, which would not take place as such if readers always integrate context information into their processing or use inferencing to support word-recognition (Grabe, 2009).

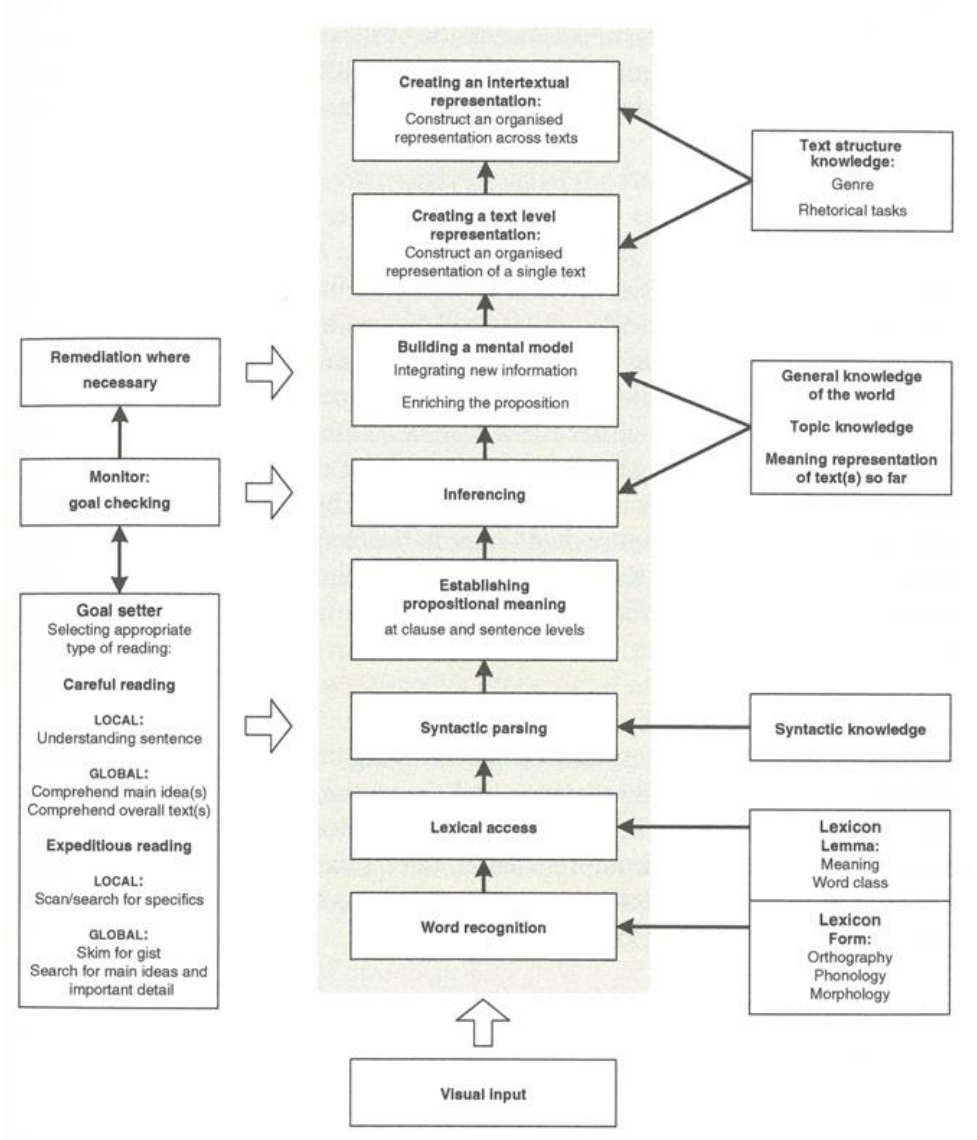
This has led to the development of interactive compensatory models that view the influence of higher-level processing strategies on bottom-up processing as a compensation when the reading

process is impaired – for example due to poor reading skills or high text difficulty (Khalifa & Weir, 2009). The disrupted automatic bottom-up reading process leads readers to draw more strongly on other informational cues to help them in their reading process. This model therefore accounts for the observation that skilled readers do not depend on top-down processing as much as low-skilled readers, because their bottom-up skills are so automatized that top-down processing would be too slow, but also considers the possibilities of the use of higher-level processing strategies for reading support.

The following model (figure 2.3) is an L1 interactive compensatory reading model that the authors described to also function as an L2 model, because it "can be treated as the goal towards which the L2 reader aspires" (Khalifa & Weir, 2009, p. 43). This is an ambitious claim. Admittedly, there seem to be general universal reading processes like making use of phonological processes during encoding or the use of background knowledge or morpho-syntactic information during comprehension. Moreover, there are also other well-known models, like the Psycholinguistic Guessing Game Model (Goodman, 1967), a top-down model, that claims that reading is a universal process. But, especially when looking at details, these processes differ between languages (for a list of universal aspects and discussion, see Grabe, 2009, p. 121-126; Koda, 2005, 2007). Therefore, it is not quite clear if one model can account for reading in all languages. Second, L2 reading differs from L1 reading in so many ways that it is questionable if developing an L1 reading model in the L2 is a realistic and even desirable goal. The following section on L2 reading will elaborate on these differences. On all accounts, the claim that this model can be treated as a goal to strive for in L2 reading is a ventured claim. Nevertheless, this L1 model is useful for identifying subskills of the reading process and discussing their interplay and differences between L1 and L2 processing. Thus – as there are no L2 reading models available - I will refer to Khalifa & Weir's model as a basis for the further discussion.

The boxes word recognition, lexical access, syntactic parsing and establishing propositional meaning in figure 2.3 are usually seen as lower-level processing skills. Inferencing, building a mental model, creating a text level interpretation and creating an intertextual representation involve higher-level processing strategies. I will elaborate on these concerning L1 and L2 differences in the following sections. The goal setter on the left side of the model influences the processes in the center. Therefore, different goals will influence the ways in which a text is processed. This is relevant in this study because students in the intensive and extensive reading interventions were likely to have different goals when reading: In the extensive reading interventions students should read for pleasure and enjoyment. Their goal was to understand the gist of the story. In the intensive reading interventions, students had to engage in vocabulary, grammar and comprehension exercises after reading, meaning a precise understanding of the text was necessary. The differences in reading goals could have an influence on the reading process and thus the skills developed and fostered in the reading interventions. This should then show in differences in the language and reading comprehensions test score gains between the

intensive and extensive reading intervention groups. Where these differences could lie will be deducted and hypothesized in the following sections.



From: Title: Examining Reading: Research and Practice in Assessing Second Language Reading. Authors: Hanan Khalifa, Cyril J. Weir © Cambridge University Press. Reproduced with permission of the Licensor through PLSclear.

Figure 2.3 Khalifa & Weir (2009), p. 43: A Model of Reading.

2.2.4 L2 Reading

Hedgcock and Ferris (2018) stated that "L2 educators have much to learn from expertise of L1 literacy specialists […] [because] L1 and L2 reading development follow a highly comparable path" (p. 45). As already noted in the previous section, this is not so clear and straight forward as some authors claim. Learning to read in an L1 differs from learning to read in an L2 for various reasons (Grabe, 2009, 2014; Hudson, 2007; Koda, 2005, 2007): First, L2 readers often already have L1 reading experience. The benefit from this experience of course differs and is dependent on the orthographic and linguistic distance of L1 to L2 and features of the L2 itself. Second, L2 learners' language competence, when

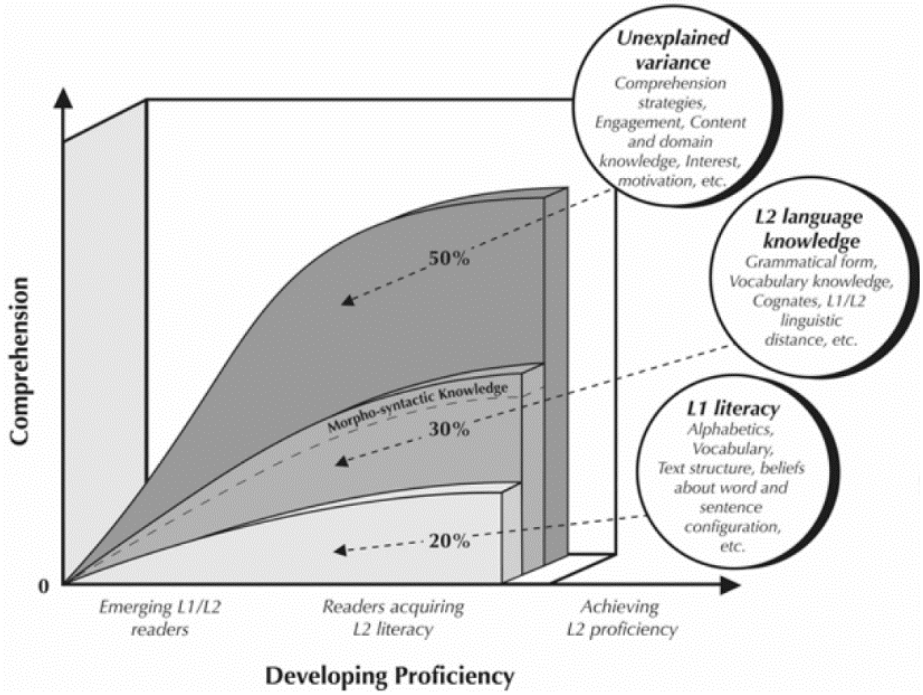
learning to read in the L2, is usually much lower than their L1 language competence that they are usually fluent in when learning to read in their L1. Third, reading in an L2 always has a dual-language involvement, meaning that during processing there is an interaction with the L1 language present. Fourth, assuming that they learn their L2 in an L1 environment, L2 readers usually are exposed to less print than L1 readers. Fifth, L1 and L2 learning can differ significantly concerning the underlying goals and purposes of reading, influencing the reading process. Finally, readers might have different cultural experiences or are acquainted with social conventions when reading in their L1 than it is common for the L2 (Grabe, 2009, 2014; Koda, 2005, 2007). This makes learning to read in the L2 very different from learning to read in the L1.

Which reading processes are relevant for reading and how these interact is already an inexhaustible and disputed field when only looking at L1 reading. Taking into account L2 influences adds even more variables to this already broad field. Hedgcock and Ferris (2018) claimed that "[m]ost L2 learners and teachers can affirm that L2 reading proficiency develops with greater effort and difficulty than does L1 reading" (p. 45). L2 reading is more complex than L1 reading, because two languages are involved (Grabe, 2009; Koda, 2007), but this does not necessarily mean that it is more difficult. Moreover, the question of which, how and why processes are transferred from L1 reading to L2 reading is central, but widely unanswered. Koda (2007) referred to the Connectionist theorem that sees transfer as an "automatic activation of well-established L1 competencies (mapping patterns[3]) triggered by L2 input" (p. 17). Consequently, transfer occurs unintentionally and only for ingrained L1 skills. Additionally, it can be assumed that transfer also occurs if L2 proficiency is low, in other words, there is and always will be some sort of L1 transfer. With increasing L2 input, mapping patterns will be extended and adapted to the L2 language (Koda, 2007). This makes L2 reading a complex, but not necessarily more difficult, process. Moreover, becoming literate in two languages at once does not seem to have negative impacts on students reading and writing, as students find creative strategies for coping with the new language (for more information see Bredel, Fuhrhop, and Noack (2017)).

In what way L1 reading ability and L2 language competence influence L2 reading ability is debated. Two opposing, but also compatible, theories are prominent: the Developmental Interdependence Hypothesis (Cummins, 1991) and the Language Threshold Hypothesis (Clarke, 1980; Cummins, 1979). The Developmental Interdependence Hypothesis, as the name already states, sees a high correlation between L1 and L2 reading ability. The idea is that if L1 reading abilities are well established and automatized, they will transfer to L2 reading. Verification from research for this

[3] Mapping patterns are "a set of relationships between forms and functions […]. Because such relationships do not embody closely matched, one-on-one correspondences, they are seen as correlational, rather than absolute, rules. According to the Connectionist theorem, moreover, the internalization of such relationships occurs through cumulative mapping experience. A pattern of mapping is internalized when its execution is automated." (Koda, 2007, p. 17)

hypothesis is ambiguous (Cummins, 1991; Grabe, 2009; Koda, 2005). Different populations and reading scores used in reading studies could be one reason for inconclusive results, therefore clarification is needed as to which skills transfer and if this is language, age or proficiency dependent (Cummins, 1979; Koda, 2005). The Language Threshold Hypothesis states that students have to overcome a certain, yet undefined, linguistic threshold in their L2 in order to transfer L1 reading skills, making L2 knowledge the dominant factor (Grabe, 2009; Hudson, 2007; Koda, 2005). In fact, L2 knowledge seems to explain 30%-40% of L2 reading competence variances (Koda, 2005), which strongly supports this view. Yet, problems also lie in details: Studies use different measures for L2 language competence and L2 reading and analyze various populations (Koda, 2005). Therefore, the questions which L2 language competences impact which reading skills and how universal these processes are is not fully investigated so far. Grabe (2009) also raised the question if it is even reasonable to talk about a threshold, a certain point, that is overcome, or if this developmental process is rather a continuous one.



From: Bernhardt, E., Progress and Procrastination in Second Language Reading, Annual Review of Applied Linguistics, 25, 140, 2005 © Cambridge University Press, reproduced with permission.

Figure 2.4 Compensatory Model of Reading Proficiency (Bernhardt, 2005, p. 140)

Bernhardt (2005) suggested a model (see figure 2.4) where the impact of L1 experience and L2 knowledge explain about 50% of the variance in reading comprehension - L2 knowledge explains 30% and L1 literacy 20%. The model also suggests that with improving L2 literacy (x-axis), the impact of L2 knowledge and L1 literacy increases (curves), as does comprehension (y-axis) as a rather continuous process. A closer look at the model shows that even though the impact of L1 experience and L2

knowledge increases and, therefore, comprehension increases by proficiency, the proportions of L2 knowledge and L1 literacy that influence comprehension are constant and therefore independent of the developmental stage of the reader. These results conflict with Koda's (2005) reasoning that, on the basis of previous research, “facilitation benefits derived from L1 reading experience seemingly vary among L2 learners at continuous developmental stages” (p. 23) and also with Grabe's (2009) speculations that previous research on the Language Threshold is also “interpretable in terms of a continually changing relationship as L2 proficiency increases” (p. 148). Unfortunately, long-term longitudinal research on L2 reading skill development, which could give insight, is still needed.

In this study both influences, L1 reading ability and L2 language ability, will be taken into account when analyzing the improvements of students in the pre- and posttests that are used to monitor the students’ language improvement. Since this study tests the students in the intervention classes every six weeks, it does not only assess the influence of L1 reading competence and L2 language competence at one point in time but also its development depending on the intervention type. On the basis of these assumptions, students that develop a higher L2 proficiency through the interventions should have better comprehension results, but also moderation effects of the amount of L1 reading should increase, because students with well facilitated L1 reading skills will be able to profit more strongly from these. Moreover, the English interventions are meant to help students with poor language skills to improve. Therefore, if students with poor language proficiency at T1 improve well during the interventions, they even have chances of outperforming students who already started with very high T1 scores.

.

On the basis of the previous discussion, I hypothesize that

1) Both, L1 reading experience (measured as a self-report of L1 books read per week) and L2 language experience (measured as scores in language tests at T1) will be significantly associated with the competence gain for students in the English language tests and in grade improvement.
    a. For the students in the non-English intervention groups, I expect L1 reading experience and L2 language proficiency both to have a positive impact on the competence gain.
    b. For students that receive additional training, especially for the reading interventions, I do not expect these effects to be as high or significant, or even negative.

The following sections move chronologically through the reading processes as presented in the L1 model of reading by Khalifa and Weir (2009) on p. 14, focusing on the way reading in English L2 differs from learning to read in German L1. The aspiration of theses sections is by no means to give a full overview of skills and processes involved in reading or L1-L2 differences while reading, but to highlight some relevant aspects worth considering for this particular study. For a detailed review of the discussion on the impacts of L1 reading competence and L2 language competence from different periods of time see Cummins (1979) and Koda (2005). For a profound overview of L2 reading see Grabe (2009).

2.2.4.1 Visual Input and Word Recognition

Word recognition is seen as a core skill of learning to read and automatic word recognition is seen as the core asset of a proficient reader (Grabe, 2009). Even though recognizing words and sentences might seem to be a trivial process, word recognition happens in various ways and involves numerous subskills. According to Metsala and Ehri (1998, p. 7) there are five different ways of how words can be read in English:

1. By assembling letters into blends of sounds, referred to as decoding.
2. By pronouncing and blending familiar spelling patterns, a more advanced form of decoding.
3. By retrieving sight words from memory.
4. By analogizing two words already known by sight.
5. By using context cues to predict words.

These different word recognition methods are used independently of reading proficiency and good readers are able to choose appropriately depending on the reading goal or situation (Koda, 2005; Metsala & Ehri, 1998). Grabe (2009) named four subskills that lead to rapid word recognition: secure orthographic processing, quick phonological processing, rapid semantic and syntactic processing and quick lexical access, whereas the last three processes are already part of the next processing steps in the reading model by Khalifa and Weir (2009). These differences in classification already show how complex word recognition is.

Yet, what all processes have in common is that students first have to learn to recognize printed symbols, then need to transform them to oral symbols and attach meaning to them (Thonis, 1980). This is a process that all languages with a writing system share and it can be assumed that once this skill is acquired that it is available in other languages as well (Koda, 2007). According to Thonis (1980), learning to read in English already starts by students having to learn to move their eyes in straight lines

from left to right (directionality)[4] and to recognize word and line boundaries, meaning that readers perceive a cluster of letters as a unit and stay in one line while reading. Additionally, similar letters, like *<b>* and *<d>* or *<p>* have to be distinguished reliably. Developing these skills needs training and practice and can be supported with the use of bigger fonts and space between lines. This also reflects in the reading material used in the interventions of this study: Books for L2 readers are written in a much smaller font with text over the entire page, while books for beginning L1 readers have a much larger font, only a few sentences per page and picture support. Maybe publishers assume that since German and English have very common alphabets and are both languages that are read from left to right, students with a German L1, learning English as L2, already have mastered this step when learning to read in their L2 English by grade 5 and do not need or want any additional support.

The common alphabet and reading conventions yet can also be misleading when it comes to crucial differences between English and German orthography. German nouns for example are capitalized, giving readers an additional orthographic cue about the syntactic structure of the sentence, which is not the case in English. Additionally, English and German have numerous interlingual homographs, some with the same or similar meanings (for example *bus* and *Bus*) and some with different meanings (for example *I will …* and *Ich will … in Engl. I want …*), so called false-friends. Students can, therefore, guess the meaning of unfamiliar words, but are also led off track with others.

Word recognition also involves extracting a word's phonological and morphological information from its orthographic symbols (Koda, 2007). Even though German and English both use alphabetic writing systems and roughly comparable phoneme inventories, especially for the consonants, their grapheme-phoneme correspondence differs. German students might be able to read a word like *theater* and even know its meaning, because it is the same in German, but these words are not pronounced the same (German */teˈaːte/* English */θi(ə)tə˞/*). Consequently, the transfer of lexical and phonological German L1 knowledge can have a positive or a negative impact on English L2 reading. Nevertheless, studies have shown that students with alphabetic reading experience outperform students that only have reading experience in nonalphabetic orthographies (Grabe, 2009), which puts German students learning English at an advantage.

Not only misleading similarities between German L1 and English L2 can prove difficult when reading English words, also the English language itself does not prove to be very consistent when it comes to grapheme-phoneme correspondences. This is described by the Orthographic Depth Hypothesis:

[4] This first concretization of this first step already differs between languages, for example Japanese, where books are read from back to front, top to bottom and right to left. A very confusing experience when trying to read ones' first Manga, but a striking example of how these basic reading skills already fail to transfer here.

> The Orthographic Depth Hypothesis states that orthographies (particularly alphabetic orthographies) have varying degrees of transparency between the phonological segments of the language and the orthographic symbols intended to represent the phonological segments. An orthography that closely represents the phonology of the language in a clear one-to-one relationship is a shallow (or transparent) orthography (e.g., Finnish, Serbo-Croatian). An orthography that does not closely represent the phonology in a clear one-to-one relationship is a deep (or opaque) orthography (e.g. English). (Grabe, 2009, p. 114)

German has a fairly shallow orthography, which can make it more difficult for students to cope with a language with a deep orthography like English. Students will experience and need to learn that letters and graphemes are not consistently pronounced the same throughout different words and that, from the opposite perspective, one sound can be spelled in various ways (Grabe, 2009). An enlightening example is a sentence I commonly cite in class: *He believed Caesar could see people seizing the seas* (e.g. "Interesting facts about the English language," 2020). This one sentence actually contains seven different spellings of one sound: */hi*ː *bɪˈliːvd ˈsiːzə kʊd si*ː *ˈpiːpl ˈsiːzɪŋ ðə siːz/*. Even if students manage to learn the pronunciation of a word, this does not leave them in a safe place, because affixation can quickly change pronunciation (Lieber, 2022) or a change in word stress can make the difference between a noun and a verb, e.g. *present* as /ˈprez(ə)nt/ and /prɪˈzent/. In consequence, phonological information is commonly available and retrieved only after the word is identified - whereas this is faster and rather simultaneous when reading languages with shallow orthographies (Koda, 2005).

Koda (2007) described this difference between shallow and deep orthographies by using the Psycholinguistic Grain Size Theory – the deeper the orthography of a language, the larger the grain sizes used for decoding need to be. Due to English having a deep orthography, decoding at a phonemic level would just take too long and be very frustrating; decoding needs to take place at a syllable or morpheme level (Koda, 2007). Taking these difficulties into account, students will not have the chance to quickly become confident readers by just studying and practicing letter-sound correspondences, as they might have been successful with when learning to read in their German L1. Grabe (2009) therefore suggested that students can experience a motivating and efficient progress in language learning by developing a high sight vocabulary in English - of course parallel to pronunciation practices, help for recognizing syllable structures or morphological awareness tasks. Extensive reading can be very useful in helping students to learn a fundamental high sight vocabulary, that helps them to quickly recognize highly frequent words and become proficient readers.

Yet, since oral language and literacy develop at the same time for many L2 learners, i.e. for the students in this study, it is not quite clear if this simultaneous learning might impair recognition development (Koda, 2005). Quick and secure phonological decoding has shown to be a crucial predictor for reading success, especially in the initial reading stage (Koda, 2005, 2007). Becoming

proficient is, therefore, even more difficult for L2 students in this respect, because they have not yet had extensive oral English input. Students who had good English lessons at primary school[5] might be at an advantage to students who just started learning English in grade 5, because they have had previous oral language experience.

Taking these factors into account, I hypothesize that

2) Independently of the intensive or extensive reading intervention, the students' competence gains in the pre- and posttests is positively related to the number of years of English they have had at primary school.

2.2.4.2 Lexical Access

After a word is recognized, may it be by phonological decoding or by sight vocabulary, semantic and syntactic information becomes available to the reader (lexical access). During this recognition process, entries in the mental lexicon and the incoming information are constantly matched. This also leads to activation of words with similar phonological or orthographic features, until a perfect match is found (Grabe, 2009). A word can be deciphered and therefore read from a phonological and orthographic point of view (referred to as the *form*) and if not already available, readers can form a lexical entry. This does not mean that they have to have a semantic component for this word yet. The meaning and syntactic category of the word is commonly referred to as *lemma*. Some authors use word recognition and lexical access as synonyms, others distinguish between the terms with lexical access only being complete with access to a semantic concept of the word read (Grabe, 2009). The latter is also the perception of word recognition and lexical access in the model used here (Khalifa & Weir, 2009; Weir & Khalifa, 2008), as shown by the arrows in figure 2.3: the form entry of the lexicon influences word recognition and the lemma entry affects lexical access. Since this differentiation is especially relevant for beginning L2 readers – students will probably encounter words, they do not have a concept for – I will use the term word recognition when talking about phonological encoding and lexical access when talking about successful access to the meaning of a word.

There are high correlations between vocabulary knowledge and reading comprehension (Koda, 2007). A common number, on the basis of previous research, is that 98% of the vocabulary of a text need to be known for comprehension, known as the vocabulary threshold (Koda, 2005, 2007; Nation & Waring, 2020). Even though this number is commonly cited, I cannot relate to it from my teaching

[5] English was taught in German primary schools from grade 1 on, but this changed to grade 3 in 2017. The teachers in primary schools are advised to only introduce students to oral language, e.g. singing, rhymes and simple phrases, but not to print. The English teachers in grade 5 report that students come to the secondary school (grade 5) with very different experiences and competencies.

experience and believe that the vocabulary threshold strongly depends on the type of book being read and the background knowledge of the reader. A book with text matching illustrations, in the simplest and most extreme form, a board book with one word of an animal per page and a picture of the matching animal, can be comprehended even if the student knows 0% of the vocabulary. Also, students who already know a story in their L1 or are very interested in a certain scientific topic, comprehend more when reading that story in English or reading a book about that topic in English. In this study students are offered a variety of texts that make it possible to infer from pictures or use knowledge about the stories to increase their comprehension.

2.2.4.3 Syntactic Parsing and Establishing Propositional Meaning

When moving from the word to the clause or sentence level of a text, readers make use of numerous pieces of syntactic information while processing, like word order, tense and modality (Grabe, 2009). While reading, readers combine the semantic information of the individual words (lexical access) and incoming morpho-syntactic information to form units of meaning and to decode the propositions of sentences. Grabe (2009) noted that "[t]he number of proposition units occurring in sets of sentences predicts how long it takes to process different sentences even when the number of words and clauses are kept equivalent"[6] (p. 31) and that "syntactic awareness has a strong relationship with reading comprehension" (p. 200). Therefore, syntactic parsing has a strong impact on processing fluency and accuracy.

The importance of syntactic parsing for understanding can be seen in the following example, a sentence taken from *The fair-haired Samurai* (Hunt & Brychta, 2015, p. 4); one of the books in the English library of this study.

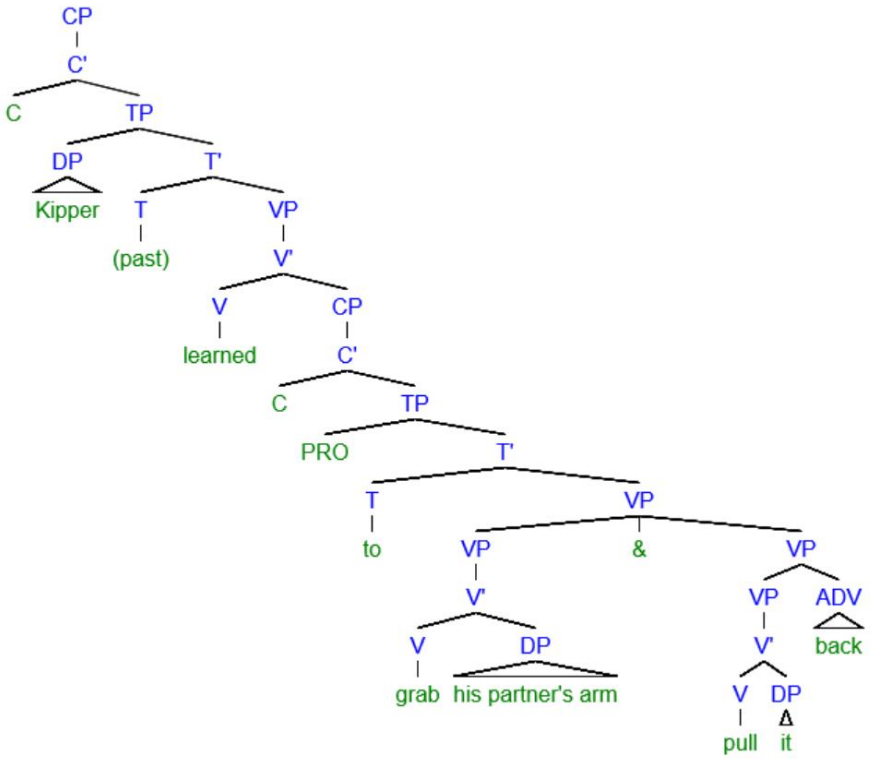
(1) Kipper – learn – grab – partner – arm – pull – back.
(2) Kipper – learned – grab – partner's – arm – pull – back.

(1) lists all content words of the sentence. By "simply" accessing their meaning, there are multiple possibilities of what action this sentence could describe. But the words in (1) are already in the correct order and, other than in German, English has a very strict word order. Therefore, this syntactic knowledge, if available, can be used to understand that *Kipper* is the subject of the sentence and *learn* the predicate. Due to the syntactic component of the entry *learn* in the mental lexicon the reader might know that it is a transitive verb *learn something* - in this case *to grab*. So far, we do not know the semantic role of *partner* and how the sentence continues because *partner* could be part of the

[6] This would imply that common difficulty measures like the number of headwords or words per sentence measures are not informative enough to determine text difficulty.

verb phrase (e.g. *Kipper learned to grab the partner or the partner's arm …*) or the subject of a new main clause (e.g. *… the partner's arm pulls him back* or *… the partner's arm is pulled to his back*). In (2) morphological information is added. Combined with syntactic knowledge, the genitive case of *partner* tells us that something happens with the arm of the partner and that *partner's arm* forms a determiner phrase (DP) – either as part of the object of the first main clause or as subject of a second main clause. Morphological information of *pull* signals that the latter cannot be the case, because *pull* is not inflected for subject-verb agreement or passive voice. Therefore, the DP merges with the verb (V) *grab* to form a verb phrase (VP). Moreover, we now know that the action happened in the past, because *learn+ed* signals past tense. If we now add function words to the sentence, in this case the infinitive particle *to*, the determiner *his*, the pronoun *it* and the conjunction *and*, we get:

(3) Kipper learned to grab his partner's arm and pull it back.[7]



In this example, the function words would not have helped to disambiguate the syntactic role of *partner's arm* even if they had been known from the beginning on. Therefore, when parsing this sentence, the different possibilities of integrating the DP *partner's arm* in the sentence is not clear until

[7] The bottom part of this tree has already been explained in the text: the determiner phase (DP) merges with the verb (V) and forms a verb phrase (VP). This first VP is then conjoined with a second VP and forms a larger VP. This VP is the complement to the T node which contains information on tense and agreement, in this example the infinitival particle *to*. The subject in this clause is filled with an implicit subject (PRO) that refers back to *Kipper* and completes the tense phrase (TP). The subordinate clause is completed by the functional projection CP (complementizer phrase), which carries force features and might host subordinating conjunctions. To simplify the tree structure, features are not explicitly mentioned in this example. The CP functions as a complement to the main verb *learn* with which it merges to a VP. This projects into a finite TP and eventually again a CP.

arriving at *pull*. Also, the meaning of the word *back* needs to be disambiguated by using syntactic information. The syntactic structure of the sentence above tells us that *back* is not a noun in this case (meaning a part of something that is opposite to the front) but an adverb that tells us something about where the arm is pulled.

This example illustrates that syntactic information is also essential for establishing meaning. With an average reading speed of about 250 words per minute (Renner, 2010), this example also shows the complex interaction of lexical, morphological and grammatical processes working together to establish a propositional meaning of a sentence in only 2.64 seconds.

The students in this study are at the very beginning of their L2 learning experience, thus, they do not have a lot of grammatical knowledge yet. Syntactic parsing differs between languages (Koda, 2007), also between German and English. It is important that students are able to encounter and internalize basic and simple sentence structures in texts. For example, to understand that English follows a strict subject-predicate-object order, which is different from their German L1, where word order is more flexible. Additionally, it is difficult for students to identify and understand verbs of the sentences, especially if they are inflected for tenses and aspects that the students do not know yet. Inflection for past tense or progressive aspect could already prove to be a problem here. However, texts can also introduce new grammatical forms and enhance grammatical understanding, giving students the opportunity to learn these implicitly.

Finding challenging but not overwhelming texts for beginning L2 readers of English is quite difficult from a syntactic point of view since syntactic difficulty levels of texts are usually not reported. Students in 5th grade, for example, had not learned irregular forms of past tense verbs yet, but frequently encountered the word “said”, because it was commonly used for direct speech in books aimed at L1 beginning readers like the Lady Bird Series and Biff, Chip and Kipper. After asking about it, they knew what it meant and also used it in texts they produced themselves, even though not always correctly, as the following two examples show:

Figure 2.5 Answer in the Portfolio of Michl, Grade 5

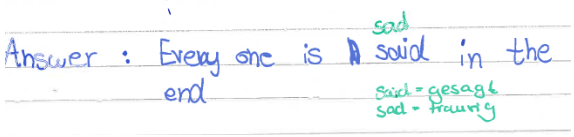


Figure 2.6 Answer in the Portfolio of Ayliss, Grade 5

Due to the individual differences between different L1 and L2 languages and a lack of research, it is unclear to what extend syntactic similarities and differences between languages influence or

enhance L2 reading. What is known, is that syntactic knowledge seems to significantly contribute to reading performance in the L2 (Koda, 2007). Therefore, integrating tasks raising linguistic awareness could be an important asset of an intensive reading intervention.

For filling in the cloze tests in this study, students needed to make use of syntactic cues to establish a propositional meaning of the text, even with words and information missing. Prepositions were tested as they need combined semantic as well as syntactic knowledge which will only be acquired via enough input and not via explicit teaching. On the one hand, students in the intensive reading interventions could have higher competence gains in these tests, because they had a chance to develop more language awareness through the additional tasks. On the other hand, students in the extensive reading intervention read more in total and therefore had more language input and thus possibilities to develop explicit grammar knowledge.

2.2.4.4 Inferencing

Inferencing involves higher-level processing strategies. At this stage, we assume that the reader is able to recognize the printed words, access their entry in the mental lexicon and add additional morpho-syntactic information to establish a propositional meaning of the units in the sentence. By inferencing the "reader can go beyond explicitly stated ideas as links between ideas in a passage are often left implicit" (Khalifa & Weir, 2009, p. 50). Since information has to be added to achieve a coherent understanding of the text, this is a creative process, but thus it is also prone to misinterpretations.

Inferencing already takes place when retrieving the correct entry from the lexicon for (ambiguous) words. For example, when looking at the title of the book *The Fair Haired Samurai* (Hunt & Brychta, 2015). *Fair* and *Samurai* are words that the students already know from their German L1 and *hair* is a word they learn early and probably already know. If the student is able to make use of morpho-syntactic knowledge – again highlighting the importance of morphological knowledge in this example – understanding that *haired* must have something to do with *hair*, it becomes clear that *fair* cannot be meant in the sense of being impartial. Syntactic knowledge shows that *fair haired* modifies *Samurai.* This does not help disambiguate, because the Samurai could be fair in the sense of impartial and students could think, "well, I don't care what *haired* means, but the story seems to be about an impartial Samurai", or they could try to find more cues. One could be world knowledge and the picture on the cover. Students might know that Samurais are Japanese warriors that usually have black hair, but there is also a boy with blond hair on the cover wearing an outfit of a Samurai. Thus, *fair haired* could be a description of the boy's hair, meaning that it is a synonym to *light*.

This example shows that knowledge that is suitable in the L1 language might not fit or be misleading in the L2 (Khalifa & Weir, 2009). Moreover, it shows that inferencing is also an important process for incidental vocabulary learning (Koda, 2007), because word information from the text is used not only to disambiguate information (Khalifa & Weir, 2009) but also to guess the meaning of unfamiliar words.[8]

2.2.4.5 Building a Mental Model, a Text Level and an Intertextual Representation

Whereas literature on L1 and L2 reading is quite consistent about naming and describing lower level reading skills, higher-level processing strategies are described and named differently by various authors. Grabe (2009) named and described a two-process model of comprehension that is made up of a text and a situation model and an additional set of skills and strategies like goal setting, inferencing, making use of background knowledge and comprehension monitoring as higher-level processes. These are very similar to the last four steps of the reading model by Khalifa and Weir (2009), but structured differently. I will use the two-process model to describe these last steps in the reading process, because it shows very well how these higher-level processing skills interact.

The text model of reader comprehension describes processes in which the reader connects propositions to form relationships, identifies main ideas or important information and activates existing information to integrate the information into the existing mental network. The text model therefore can be described as an understanding of the text itself. In order to create a text model, the reader must disambiguate information, make use of general background knowledge, cope with left out information, suppress irrelevant information, activate knowledge about the genre, etc. (Grabe, 2009). Some texts are always structured in a similar fashion. This information can help readers to keep an overview of the text while reading and to identify its main points (Koda, 2007).

The situation model adds the component of interpretation to the text model and accounts for why different readers perceive the same text in different ways. Readers build on previous experience with books, genres, narrations, topic knowledge, attitudes etc. before, while and after reading (Grabe, 2009). Grabe listed eight factors that influence the construction of a situation model that I will use to show how these higher-level processing strategies could be relevant in the L2 reading of this study:

1. Reader purpose: In a broader sense, the students in this study all engaged in English reading because they were in a class where they had to. To not bias results, they were randomly chosen to be in the intervention classes and therefore they are probably not

[8] When ordering and preparing the books for the English library and keeping them in a shelve at my office, the secretary of our department approached me one day and happily said "I read some of your books this morning and I hope it was okay. And I actually learned a few new words. It was nice and easy because I could judge from the pictures."

(initially) intrinsically motivated. They also read the books to read and not to extract information needed for something else. As already described in section 2.2.1, the reading of intensive and extensive reading differs in this study, due to the additional tasks of the intensive reading interventions.

2. Task expectation: Students probably came to the reading lessons with different expectations. Some students might have expected to have fun reading the stories, some might have had the expectation to improve their language skills, some might have expected to have the worst time in their life, because they hate English. These emotional states and expectations are likely to influence the students' reading process.
3. Genre activation: Students already have experience with different genres from their L1, e.g., concerning comics, plays, fairy tales or fact books. They know, for example, that you usually do not read a fact book from beginning to end, but flip through it and read interesting captions or pages. This would not work for a fairy tale. Yet, it is important to keep in mind that genres are also shaped by social conventions and expectations and can therefore differ between cultures and languages (Hudson, 2007). In the case of the genres used in this study and German genre expectations of the students, I do not expect this to be a problem.
4. Similar story instances: Students in this study have experienced similar stories for example when reading board books, because they have already read those or they were read to them in their L1. Another example are the *Peppa Pig* stories of the Lady Bird Series. They were very popular during the interventions and students explained that they had seen the TV series and liked them.
5. General background knowledge resources: I have already given some examples in previous sections how background knowledge can help but also interfere with students reading comprehension.
6. Evaluation of the importance of information, its enjoyment value, its interest value: This point is similar to point 1 and 2 and is likely to differ between the intensive and extensive reading interventions, but also between different students. I will discuss the impact of value as an important asset of motivation in section 2.3.4.8.
7. Attitudes (and inferences) towards writer, story, genre, episode: Attitude as a central variable influencing the reading process will also be discussed further in section 2.3.4.8
8. Inferences needed for interpretation: This point can be seen as a superordinate of points 3, 4 and 5. Students might prefer books where they are able to make inferences based on their experiences or knowledge.

This list shows that various factors influence the situation model of reading but that it is also difficult to analyze these in isolation, because they are interdependent. Background knowledge or similar story instances can, for example, influence the attitude towards reading or the expected enjoyment value and consequently task expectation. Additionally, this list shows how many individual factors readers bring along to their reading process and that we can expect to have high individual differences.

The text model and situation model both simultaneously contribute to understanding during the reading process. Depending on the reading ability, the purpose of reading or the range of background knowledge, one model might have a stronger influence than the other (Grabe, 2009). The students in this study have different L2 language competencies and different background knowledge. For example, students with a high background knowledge but low language competence might overcompensate their poor language skills with background knowledge, leading to misinterpretations of the text. This might be the case for students reading the *Simpsons Comics* because the language is very difficult but they can get a gist of the story by looking at the pictures and combining those with previous experiences watching Simpsons episodes on TV. It is obvious that this reading behavior will neither lead them to a deeper understanding of the complex jokes, parodies and allusions of the comics, nor improve their language competence, but they have an illusion of understanding. Students in the extensive reading interventions can maintain this illusion, whereas students in the intensive reading intervention are tested and disillusioned by the additional language and comprehension tasks they have to engage in. Therefore, students who are not that conscientious about their reading process might perform better in an intensive reading intervention environment where they are under closer control.

Yet, background knowledge, previous story instances and expectations of the text genre might also help in monitoring the reading process and comprehension, as the following example shows: One student came to me very disappointed but also very confused about a fairy tale she had read, because the story was significantly different to the way it is told in Germany. This student had expectations about the text and the story (situation model) which were not in line with her perception of the text (text model) which made her insecure if she had understood the text correctly and asked about it. This is an example of how the situation model can positively interact with the text model to enhance a reading process and in this case also intercultural competence.

The extensive and intensive reading interventions in this study differed concerning the purpose of reading and therefore might have also led to different emphasis on the situation and the text model within the reader. In the extensive reading interventions students were supposed to read for pure enjoyment, understanding the main ideas, which could emphasize the use of the situation model. In

the intensive reading interventions, students had to engage in tasks afterwards like answering comprehension questions, or retelling a similar story, where they had to adopt the narrative style of the story. These tasks involved developing a close understanding of the texts and its discourse structure, therefore emphasizing a text model representation of the text. Additionally, even if the students have background knowledge or experience that helped them cope with the text, they will still have to develop a thorough representation of the text itself to fulfill the tasks.

Taking differences between intensive and extensive reading into account when it comes to building a text model or situation model representation of the text and the additional tasks in the intensive reading intervention drawing awareness to word acquisition and syntactic structures, I hypothesize that:

3) Students participating in the extensive reading intervention will perform better in the fictional and non-fictional reading comprehension test than students in the intensive reading intervention, who will perform better in the cloze test and the preposition test.
4) In the extensive reading intervention students will prefer reading books where they can draw on background knowledge to understand them, whereas this will not show as strongly for intensive reading.

The differences in hypothesis three between the test types derive from the assumption that non-fictional and fictional reading comprehensions are tasks that involve making use of higher-level processing skills and prepositions and cloze tests involve being able to use semantic and syntactic cues to establish a propositional meaning of the text.

2.3 Extensive Reading and Intensive Reading

2.3.1 What is Intensive Reading?

Hedgcock and Ferris (2018) named four characteristics that are typical for intensive reading courses:

- The texts to be studied are selected by the teacher (perhaps with input from students).
- All students read the same text at the same time and complete in-class or out-of-class exercises and assessments designed or assigned by the teacher.
- The teacher highlights specific linguistic features and content dimensions of the text, introducing and reinforcing selected reading strategies through whole-class instruction activities.
- Assessment of student's comprehension, reading development, and reading efficiency is facilitated by the fact that all students work simultaneously with the same text (p. 173).

These aspects only partially apply to the intensive reading courses in this study, because they were designed to be as comparable as possible to the extensive reading courses. Therefore, the material that the students read was not selected by the teacher, the students chose from all the books that were available for the extensive reading intervention. Moreover, the students also did not read the same texts at the same time. However, the other characteristics apply to the intensive reading part of this study: The students completed assignments on the texts in class. These individually designed assignments for each book highlighted specific linguistic features, narrative and content of the books. The students' comprehension and reading development was also facilitated, even though the students did not work simultaneously: the students and teacher kept track of the individual reading amount with the portfolios and library program (see section 3.2.2.2 for more details). Moreover, the students received feedback on their portfolios from lesson to lesson and were given individual support in the lessons. This last aspect is a central reason that Hedgcock and Ferris (2018) name as an argument for intensive reading: that the reading material and activities need to be synchronized in an intensive reading classroom is reasoned by Hedgcock and Ferris (2018) with the argument that it is otherwise too difficult for teachers to assess the reading activities, "as the teacher likely will not be able to read everything that individual students read" (p. 175). Yet, this is exactly what I did in my intensive reading courses, since I wanted students to have the same choices as students in the extensive reading interventions.

2.3.2 What is Extensive Reading?

> Extensive reading, in contrast to intensive reading, generally involves rapid reading of large quantities of material or longer readings (e.g. whole books) for general understanding, with the focus generally on the meaning of what is being read than on the language. Extensive reading is intended to get the reader to focus on reading for the sake of reading (for information or entertainment), and less on reading for the sake of mastery of a particular linguistic structure of even a particular reading strategy or skill. Thus, it can involve a wide variety of reading skills or strategies. (Carrell & Carson, 1997, pp. 49–50)

Even though this definition from Carrell and Carson is almost 25 years old, it is very suitable for the purpose of this study, because it defines extensive reading in contrast to intensive reading. Moreover, the points named are still included in current discussions about the nature of extensive reading, but have been narrowed down or elaborated on.

The first quality of extensive reading named in the definition is high reading speed. This point has been put into more concrete terms in the past years: a reading speed of about 200 words per minute is named as an aim. This leads to an additional aspect of extensive reading that is not

mentioned in the definition above, but logically follows from this point: Reading should be individual and silent in order to read at a personal fast pace (Nation & Waring, 2020).

The second aspect mentioned is that readers read large amounts and longer texts. These texts should be at the right level, which means that only a maximum of 2% of the vocabulary should be unknown (Nation & Waring, 2020). Additionally, the texts should be extensive in the amount, level and topics, so that students can find the perfect books for themselves (Nation & Waring, 2020).

Third, the focus of this reading should be on general understanding. This is a point seen differently by authors. Day and Bamford (1998), for example, also name this point in their top ten principles of extensive reading. Nation and Waring (2020) do not mention this aspect, in contrast, they have a strong focus on comprehension. Therefore, it is very important to them that students read at the right level, where they know almost all the vocabulary. This also has consequences on practical didactic implications, for example, the use of dictionaries, which will be discussed in the next section.

The last point of the definition is that students read to read and not to improve their language competence. Improvement of language competence, naturally, could be a consequence of extensive reading and is the reason why it is promoted by researchers and teachers, but it should not be the main aim of engaging in reading. Nation and Waring (2020) do not include reading for enjoyment in their definition of extensive reading: Extensive reading "involve[s] each learner independently and silently reading a lot of material which is at the right level for them" (p. 3). The reasoning of the authors is that enjoyment and motivation is not directly influenceable, but that hopefully a suitable fit of the material would lead to fulfillment and enjoyment when reading. Yet, the idea of enjoyment is essential to the idea of extensive reading programs. Enjoyment is supposed to give students the opportunity to experience texts not only as a medium they have to study with, but something entertaining (Day & Bamford, 2002). A positive attitude towards reading can be crucial for a positive language development. I illustrated this idea in the following diagram:

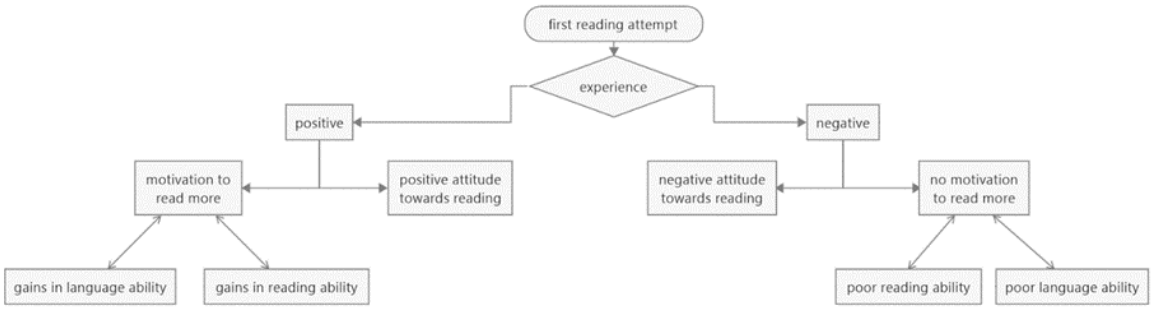


Figure 2.7 The Vicious or Virtuous Cycle when Reading

If students make positive experiences when reading, they are motivated to read more; if they make negative experiences, they are not. According to Koda (2005) this experience already starts with word recognition:

> [P]oor decoders become increasingly frustrated with reading. With sustained frustration, reading becomes anything but rewarding, further discouraging voluntary reading. In absence of adequate reading practice, word-recognition skills remain underdeveloped, and poor comprehension continues. (Koda, 2005, p. 30)

Extensive reading has been shown to have positive effects on students reading and language abilities (Day & Bamford, 2002; Nation & Waring, 2020) which again has a positive effect on their academic success. Therefore, integrating motivating extensive reading programs with manageable reading material into classrooms could enhance students' positive attitude towards reading and their academic success.

2.3.3 Effects of L2 Extensive and Intensive Reading on L2 Language Competence – Previous Research

Although L2 extensive reading studies and papers always point to the multiple proven positive effects of extensive reading in previous research, numerous studies on the impact of extensive reading can only be found in L1 reading research. L2 reading is relatively unexplored (Grabe, 2014; Koda, 2007). Some of these L2 studies also do not fulfill simple requirements of solid empirical research (see Nation and Waring (2020) for some interesting examples). Moreover, very different forms of courses are labeled as extensive reading. Different views on essentials of extensive reading courses (see section 2.3.4) make it difficult to draw a line on what is still acceptable to call an extensive reading course and what is not, and it further complicates comparing and replicating studies. Moreover, many studies are often small scale with only one class of participants and without a control group.

As described in section 2.2, L2 reading depends on various factors, including what L1 the students speak/read, when and how they acquire their L2 and if they already have L2 knowledge. Studies with alphabetic orthographies, preferably English as L2 or another language with a deep orthography, where the learners have already developed L1 reading skills, but have not developed L2 language skills yet, are surprisingly sparce. Of numerous and promising recent studies available, almost all studies did not fit these criteria, because they assessed either Asian or Arabian students learning English, or did not name the L1 and L2 of the students in their participant descriptions at all. I will therefore begin by summarizing meta-analyses on the impact of extensive reading programs on L2 development for an overview and then discuss some studies that deliver a detailed contribution to the questions addressed in this study, even if the basic conditions are not exactly comparable.

Many authors point to the Book Flood Studies (Elley, 2000; Elley & Mangubhai, 1981, 1983), when summarizing studies and positive results on extensive reading (Bernhardt, 2005; Grabe, 2009; Nation & Waring, 2020). Elley's studies are in fact impressive, because classes were equipped with books and teachers with training in various countries, the design always used control groups and the interventions were often evaluated over many years, making it possible to report longitudinal effects of extensive reading on different language skills. The students in the Book Flood programs made significant progress in reading comprehension, listening comprehension and smaller but also significant progress in writing and speaking. These gains usually increased in the second year (Elley, 2000). However, the Book Flood interventions took place in developing countries, thus the preconditions in these studies differ from this study and many other studies in the field: First, Elley implemented these reading programs in very poor schools in developing countries. They were, therefore, not assets of a well running EFL program, but an attempt to increase educational quality in English. Second, his initial aim was not to enhance reading abilities of students who are already on their way to learning English in a (well established) school curriculum, but to enhance students' overall language abilities, that did not stand a chance of improvement in poorly equipped classes with low skilled teachers. Third, "most students came from virtually bookless homes" (Elley, 2000, pp. 243–244), therefore, they were very inexperienced readers. Fourth, the studies were conducted in countries like South Africa or the Solomon Islands, that have a colonial history and English might even be (one of) the official language(s). English plays a larger role or different role in these countries, being more prominent in everyday situations, and more important for educational and job-related success, but also as the language of the oppressors. All in all, the Book Flood studies deliver important and inspiring results on the impact of extensive reading programs in developing countries, but need to be questioned in terms of transferability to other contexts and countries. Unfortunately, this issue is not addressed by literature and previous research, for example, by adding type of L1 as a variable in meta-analyses.

Krashen (2007) conducted a small meta-analysis with 9 studies from Taiwan, Japan, the Philippines, Japan and Yemen. He found overall positive effects for adolescents' and young adults' L2 reading on their L2 reading comprehension and cloze test results. Due to the very small number of studies included, I will not go into detail on these results and move to the following, more recent five meta-analyses in this field.

Kim (2012) summarized 21 papers. The results were analyzed separately for cognitive and affective outcome variables and age groups. Overall, the results show positive effects on both components over all age groups. Adults show the highest cognitive effects (d = 1.8), followed by elementary school children (d = 1.33). The results for the affective components were a little lower, but also medium to high with d = 1.13 for college students and d = 0.61 for elementary school children.

Effects for junior high and high school students were not as high. The effect sizes were also calculated for different components over all groups. Listening had the highest effect size with d = 1.94, followed by reading comprehension (d = 1.06) and writing (d = 1.04), reading speed (d = 0.80) and vocabulary (d = 0.73). Out of the affective components analyzed (motivation, attendance, attitude and anxiety), motivation was the only component with significant medium effects (d = 0.80).

Nakanishi (2015) included 34 studies in her meta-analysis and found medium to large effects (d = 0.46 for between groups and d = 0.71 for pre-post designs) of extensive reading on language competence. She additionally analyzed if these were affected by age, length of the reading classes and test type. Adults and university students had the highest effect sizes (d = 1.48 and d = 1.12), the other groups showed medium effects. The treatments seem to be more effective the longer the time period of their implementation. Additionally, the effects of extensive reading on reading speed were higher than on reading comprehension. Nakanishi (2015, p. 26) stated herself that some results have to be judged with care, because the confidence intervals contain zero. Yet, they deliver a first overview of extensive reading research with a representative number of studies.

Jeon and Day's (2016) meta-analysis of 71 samples from 49 studies published between 1980 and 2014 analyzed the overall effects of extensive reading on reading comprehension, reading rate and vocabulary and to what extent these effects are moderated by the context (ESL vs. EFL), the library size, the length of the treatment and the text types used. Jeon and Day (2016) criticized the previous meta-analyses summarized above for having too small samples. Additionally, they note that differing forms of cognitive and affective variables have been combined to one domain. Moreover, they criticize the inclusion of studies that do not meet the criteria of extensive reading. Jeon and Day found small to medium overall effects of extensive reading on language competence (d = 0.57 for between group contrasts and d = 0.79 for pre-post-contrasts) in their meta-analysis. These effects were significantly moderated by the year of publication (increasing effects the more recent the studies), the setting (results for EFL settings were higher than for ESL settings), age groups (adults showed the highest effect sizes d = 0.70, followed by children d = 0.52 and then adolescents d = 0.35) and text type (web texts were more effective than paper texts). The other factors did not prove to be statistically significant.

Liu and Zhang (2018) used 21 studies published between 2007 and 2016 to analyze the effect of extensive reading on students' vocabulary, also taking into account age group, length, control group, reading material treatment, test use and test reliability. Since the pre-post-test group only contained four studies, I will only report the results of the between group designs here. The results again showed a large effect of extensive reading on vocabulary learning (d = 1.32). As in the other meta-analyses summarized so far, these effects were highest for very young children (d = 5.18), followed by adults (d = 1.96) and lowest for junior high and high school students (d = 0.58 and d = 0.82). Additionally, the

authors analyzed that three months was a sufficient length; six months and one year even had lower effect sizes. These results are not only in contrast to Nakanishi (2015), but are also only explained by saying that “longer length of instruction may lead to quick forgetting” (Liu & Zhang, 2018, p. 11). We have to consider that these results could be partially biased, because the results are based on 11 studies with a period of one semester, eight with two semester and only two with one year and therefore have to be interpreted with great care. Studies using graded readers had the highest effect (d = 1.48). Also, the effects of different additional treatments were analyzed: comprehension questions d = 4.77, vocabulary instruction d = 1.46, vocabulary exercise d = 1.35, dictionary usage d = 1.3, book report d = 0.57, and others d = 0.32. Additional tasks are not in line with the core idea of extensive reading; therefore, it is surprising that these studies were used in this meta-analysis. In this case, the amount of extensive reading compared to these additional tasks could have significantly influenced results. Unfortunately, Liu and Zhang (2018) provided no information on the amount of pure extensive reading in these studies and if there were selection criteria or reasons for integrating these studies. Moreover, since the confidence intervals are partially very large and sometimes contain zero, these results have to be looked at carefully, something that is also not addressed by the authors. Additionally, many factors were analyzed on the basis of only 21 studies yielding the question to what extent some of these factors are based on the same studies or very small sets of studies, leading to multicollinearity issues.

The most recent study by Hamada (2020) is also only based on 22 studies published between 1997 and 2017. His results also show an overall, but small, effect of extensive reading on reading comprehension (d = .55). This effect was moderated by proficiency (high d = .71; low d = .30). In line with Nakanishi (2015), he also found stronger effects for longer interventions.

Combining the results of these meta-analyses, extensive reading seems to have a significantly larger positive effect on L2 language competence than other teaching methods (Hamada, 2020; Jeon & Day, 2016; Kim, 2012; Krashen, 2007; Liu & Zhang, 2018; Nakanishi, 2015). Additionally, all studies found age group effects with high effects for adults and partially also primary school students, but not for junior high and high school students (Jeon & Day, 2016; Kim, 2012; Liu & Zhang, 2018; Nakanishi, 2015). These results need to be interpreted carefully, because the meta-analyses partially use the same studies (e.g. Al-Homoud and Schmitt (2009) and Yamashita (2008)). Setting this problem aside, these results are interesting in the light of L2 reading development: The authors hypothesized that their findings could be due to more advanced reading skills and strategies of adults which are transferred to L2 reading. This is interesting in the context of the L1 reading experience to L2 reading transfer discussions, namely Linguistic Threshold vs. Interdependence Hypothesis (cf. section 2.2.4), strengthening Koda's (2005) suggestion that L2 transfer could depend on reading experience and ,thus, age. Additionally, almost all studies are of Asian origin and can only cautiously be applied to other L1

populations due to many differences (e.g. culture and writing systems). Jeon and Day (2016) themselves addressed this Asian focus as a limitation to their study and Nakanishi (2015) justifies in a footnote: “One of the main reasons I do not have studies conducted outside of Japan is that I simply could not find any.” (p. 12)

All in all, it is very difficult to draw conclusions from these studies and meta-analyses for this study, because research is sparce and most of these studies were conducted in very different environments. Also, turning to studies of a different L2 language than English is not possible. Studies with English L1 university students, learning a European L2 language either do not involve control groups (e.g. Hardy, 2016) or interpret extensive reading differently, like replacing text book assignments with reading one German romantic novel (Maxim, 1999).

I will therefore turn to results from Asian L1 students for some central issues. Comparing “pure” extensive reading to extensive reading with additional tasks (similar to the intensive reading intervention in this study) has shown extensive reading to have a larger positive short-term impact (one semester) on cloze test results but no significant long term (two semesters) differences in effect could be found between the two (Smith, 2006). Mason (2004) also did not find significant differences between their different experimental groups. This is seen as a positive result by the authors, reasoning that the groups do not seem to differ and extensive reading, therefore, is as efficient for language learning as other classes. From a pure number game perspective this reasoning is logical, but it is not statistically acceptable. Not being able to reject the null hypothesis (i.e. no significant differences between groups) does not mean that the null hypothesis can automatically be accepted, i.e. there is no statistical proof that there is no significant difference between these groups. One possibility of meeting this problem, at least to some extent, is to refer to the confidence intervals, unfortunately these were neither discussed nor reported. These results are, nevertheless, interesting, because they show that not all studies find significant superior effects of extensive reading.

Yamashita (2008) found positive effects of extensive reading on reading ability measures (a special extensive reading test) but not on linguistic ability measures (cloze test), speculating that the development of these linguistic skills might need more time. Again, these studies fit well to the design of this study but use Asian university and college students as participants.

Taking this state of the art into account, this study will contribute to extensive reading research in the following aspects:

- It delivers research results from the European area with L1 and L2 both being alphabetic writing systems, supplementing results from the Asian and Arabian area.

- Most participants of extensive reading studies so far are university students, i.e. adults. Therefore, this study helps to expand the studied population to middle school aged children at the very beginning of L2 learning.
- The authors also call for larger sample sizes, something that this study addresses by conducting the interventions multiple times with different students, boosting sample size.
- Extensive reading interventions in this study were strictly set to suit the core principles of extensive reading in contrast to reading with additional tasks (here called intensive reading), as outlined in the following sections.
- This study is designed with two different control groups – groups that have other English classes and groups that do something completely different. This design is new in the field.
- This study is also not only based on one intervention class. There are three extensive reading classes and six intensive reading classes.

2.3.4 Essentials of a Second Language Extensive Reading Program

That extensive reading is defined and perceived differently has already been addressed in the previous sections. Instead of constructing definitions of extensive reading, many authors resort to naming characteristics. Table 2.1 gives an overview of some essential characteristics for teaching extensive reading classes.

Table 2.1 Principles of Extensive Reading Programs

Williams (1986): Top Ten Principles of EFL Teaching
1. In the absence of interesting texts, very little is possible.
2. The primary activity of a reading lesson should be reading texts.
3. Growth in language ability is an essential part of the development of reading ability.
4. Classroom procedure should reflect the purposeful, task-based, interactive nature of real reading.
5. Teachers must learn to be quiet: all too often, teachers interfere with and so impede their learners' reading development by being too dominant and by talking too much.
6. Exercise-types should, as far as possible, approximate to cognitive reality.
7. A learner will not become a proficient reader simply by attending a reading course or working through a reading textbook.
8. A reader contributes meaning to a text.
9. Progress in reading requires learners to use their ears, as well as their eyes.
10. Using a text does not necessarily equal teaching reading.
Nation (2015): Practical Guidelines Concerning Extensive Reading
1. Include an extensive reading program as part of your language course.
2. Make sure that learners spend enough time each week on extensive reading, either around 3/16 of the total course time or better still enough time to meet the words often that they need to learn.
3. Make sure that there are two strands to the extensive reading program – (a) the strand where they read texts at the right level for them (around 2% of the running words are unfamiliar) and (b) the fluency strand where they read easily familiar texts quickly.
4. Support the fluency development strand by getting learners to do a timed reading course.
5. Support vocabulary learning from extensive reading by getting the learners to do dictionary look-up, preferably while reading electronic texts.

6. Support vocabulary learning from extensive reading by getting the learners to note unfamiliar words on word cards for later independent study. 7. Link some of the extensive reading to extensive listening, and to speaking and writing about what has been read. 8. If necessary, provide training in the guessing from context and word card strategies.
Day & Bamford (1998, 2002): Top Ten Principles for Teaching Extensive Reading
1. The reading material is easy / well within the linguistic competence of the reader. 2. A variety of reading material on a wide range of topics must be available. 3. Learners choose what they want to read. 4. Learners read as much as possible. 5. The purpose of reading is usually related to pleasure, information and general understanding. 6. Reading is its own reward. 7. Reading is usually faster than slower. 8. Reading is individual and silent. 9. Teachers orient and guide their students. 10. The teacher is a role model of a reader.
Nation and Waring (2020): Steps when Planning an Extensive Reading Program
1. Make the reader aware of the value of extensive reading. 2. Find your learners' present vocabulary level. 3. Provide plenty of interesting and appropriate reading texts. 4. Get the learners to find their reading level. 5. Set, encourage, and monitor large quantities of extensive reading. 6. Support and supplement extensive reading with fluency development and language-focused reading. 7. Help learners move systematically through graded reader levels.

Ray Williams' (1986) top ten principles of teaching reading are not aimed at extensive reading yet. They are even seen as principles of intensive reading by Day and Bamford (2002) that they then developed their extensive reading principle pendant to. Day and Bamford (1998) is one of the most prominent books on extensive reading and the authors' ten principles (Day & Bamford, 2002) have influenced many publications and the way extensive reading has been taught in the past 20 years. Even though Nation and Waring (2020) value the promotion of extensive reading, which the work of Day and Bamford has led to, they criticize that this list led to the belief that there is only one correct way of running an extensive reading program. Nation and Waring, therefore, present and discuss principles for an extensive reading program themselves, naming steps to take when planning such a program and reasons why extensive reading should be integrated in a language course (Nation, 2015; Nation & Waring, 2020). Between Nation and Waring's and Day and Bamford's principles lie 20 years of research on second language extensive reading, especially using quantitative methods, which reflects in the reasoning, concreteness and priorities of the principles. Yet, these principles are also limited to the research available to date (see section 2.3.3). The next sections of this chapter will concentrate on different aspects of extensive reading that are addressed in these principles and elaborate on them.

2.3.4.1 Student Activity

The role of the reader in an extensive reading program is, of course, to read as much as they can (see table 2.1 Day & Bamford (2002), principle 3). The recommendation is about one book per week

(Day & Bamford, 2002); this, however, clearly depends on the length of the books. Nation and Waring (2020), therefore, additionally quantify a minimum amount of reading with 5 500 words per week, equating to 55 minutes of reading when calculating with an average of 100 words per minute. The rather simple idea behind large amounts of reading is the assumption of a link between time spent reading and the development of reading and language skills (Day & Bamford, 2002). Thus, it is implicated that the more input the reader has, the more development will take place.

To stay motivated, learners should choose what they want to read (see table 2.1 Day & Bamford (2002), principle 4). In doing so, pupils do not only experience that they need to take responsibility for their learning process but also experience autonomy by being able to make choices according to their interests (Day & Bamford, 2002). Additionally, readers need to learn to find their reading level and chose books appropriately (see table 2.1 Nation & Waring (2020), principle 4). Nation and Waring (2020) suggest that students determine their vocabulary size and read a few pages of books at that level. He introduces the sentence "**R**ead something quickly and **E**njoyable with **A**dequate comprehension so you **D**o not need a dictionary." (p.13) to teach the students as a key reference for finding their level of reading. A consequence of these two demands is that extensive reading is individual and silent (see table 2.1 Day & Bamford (2002), principle 8), otherwise readers would not read at their own pace and not make personal reading experiences.

In this study, students in both intensive and extensive interventions are encouraged to read as much as they can. In the first lesson they are introduced to the variety of books available. Moreover, they are shown that the different levels of the graded readers are not equivalent and encouraged to read one or two pages of the book first, to find out if they are able to understand the text. For more information on the implementation of the reading intervention courses see chapter 3.

Nation and Waring's (2020) READ acronym contains an additional aspect also named by Day and Bamford (2002) in principle 8: Reading should be fast and students should be encouraged and trained to do so (see table 2.1 Nation (2015), principle 4 and Nation & Waring (2020), principle 6). Yet, Nation (2015) also suggests that students should work with unknown words by looking them up in a dictionary and learn these using word cards (see table 2.1 principle 4 and 5). The question whether students should rather use a dictionary to give them the chance of learning frequently occurring new words in the text or if fluency should strictly be placed before accuracy is debated. Koda (2005) also reported that studies have delivered conflicting results, but especially with adult L2 learners, using a dictionary seems to enhance vocabulary learning. The discussion of whether using a dictionary during reading is useful or not is summarized very pointedly by Hudson (2007): "As with the research in other areas of reading, whether dictionaries are effective depends upon the context of the learner, the learner's

ability vis-à-vis the language threshold, the type of word that is looked up and the importance of the word for comprehension" (p. 250).

In this study the dictionary use is one of the differences between the extensive and intensive reading course. Students in the extensive reading course are not provided with dictionaries and encouraged to guess from the context and ignore unknown words. When understanding becomes too difficult, due to too many unknown words, they are encouraged to choose an easier book. In the intensive reading course, on the contrary, looking up unknown words and keeping a vocabulary list is obligatory (see chapter 3 for more detailed information).

2.3.4.2 Role of the Teacher

The role of the teacher in an extensive reading program begins by explaining what extensive reading is about, what advantages it has and how it is done (see table 2.1 Nation & Waring (2020), principle 1). During the course, the teacher supports the students individually in their reading by helping them find appropriate books to read and speaking to them about their reading experiences (see table 2.1 Nation (2015), principle 9; Nation & Waring (2020), principle 2, 5 and 7). Additionally, the teacher should monitor the reading process of the students (see table 2.1 Nation & Waring (2020), principle 5). Apart from that, Nation and Waring (2020) encourage teachers to sit quietly and read as well to function as a role model for the students (see table 2.1 Williams (1986), principle 5; Day & Bamford (2002), principle 10).

Therefore, the role of the teacher in the reading course is rather passive. The main work is planning an appropriate extensive reading course and integrating it in the language course. The central question of this planning phase is which reading the students should engage in and which reading material is appropriate for the specific readers participating in the program. These questions will be addressed in the following sections.

2.3.4.3 Additional Tasks in an Extensive Reading Course

Authors have very opposing opinions about additional tasks in extensive reading courses. Both, Ray Williams and Day and Bamford believe that reading positively influences language ability and that students should spend time reading. But Ray Williams (1986) stresses the importance of different tasks supporting the development of reading skills (see table 2.1 Ray Williams (1986) principle 4, 6), the interaction of the reader and the text (see principle 7 and 8) and the importance of a combination between listening and reading (see principle 9). These aspects are not mentioned in the extensive reading principles of Day and Bamford (2002). Extensive reading is about reading as much as possible

without further tasks slowing down the reading process and distracting from reading and reading being its own reward. Full comprehension is not the aim of reading; readers should just understand enough to enjoy the story or extract the necessary information. Yet, some of the aspects Ray Williams (1986) names are mentioned again by Nation (2015). Nation lists principles on training students' fluency and vocabulary learning (see principle 4-6), by suggesting that extensive reading should be linked to extensive listening and speaking and writing tasks (see principle 7) and the training of reading strategies (see principle 8). A complete understanding of the text is seen as more important again. Therefore, depending on the particular view on extensive reading, the intensive reading course in this study could also be seen as an extensive reading course with additional skills and language training, as presented by Nation (2015). Nevertheless, I will stick to the term intensive reading because students engage intensively with the books they read, spending most of their course time working on the tasks provided for each book.

2.3.4.4 The Role of Extensive Reading in a Language Course

The different perspectives on extensive reading also influence the role extensive reading has in the language classroom: On the one hand, it can be seen as a style of teaching reading, next to skimming, scanning and intensive reading. On the other hand, in a bigger context, it can be seen as an approach to teaching reading (Day & Bamford, 1998).

Nation (2007) sees extensive reading as one of "four strands" in a language course. The underlying ideas of this structure are, first, the more time you spend doing something, the better you become at it (time-on-task). Second, language learning involves four skills that need to be represented and skilled: listening, reading, writing and speaking. Extensive reading, again, is one of four styles of reading. Third, reading for accuracy and reading for fluency are two different skills that need to be trained differently. He names four strands that need to be taught:

1) Meaning-focused input = extensive reading and listening
   The main focus of this strand is to understand what is read or heard, to gain knowledge by doing so and to enjoy the text.
2) Meaning-focused output = speaking and writing
   Learners notice deficits more often when producing output than when solely taking in input. Therefore, this strand has an important feedback function for the learner.
3) Language-focused learning = explicit language learning
4) Fluency development = becoming faster and better in the four main skills

Nation (2007) argues that in a language course each strand should be represented to an equal amount of time, the same holds for the skills in each strand. Extensive reading makes up ½ of the first

strand and ¼ of the fourth strand, thus, extensive reading should make up about 3/16 of a language course.

The reading intervention time in this study makes up 2/7 of the total English lessons, exceeding Nation's suggestion. Yet, the course only takes place for six weeks (1/3 of the semester), therefore, extensive reading only makes up about 2/15 of the time spent teaching English in the school year, which is below the suggested relation. In the intensive reading intervention, students additionally have to produce meaning-focused output (strand 2). Questions from students during the interventions and the answers in the portfolios show very clearly that students become aware of deficits in understanding through these additional tasks. Additionally, some tasks of the intensive reading interventions focus on explicit learning of grammatical features or vocabulary (strand 3).

Nation and Waring (2020) as well as Day and Bamford (1998) state that there are multiple ways of looking at extensive reading, two of them being the cognitive (fluency, comprehension, …) and the pedagogical (how much and what is read, follow-up assignments, …) view. This dissertation takes both views into account by analyzing the effects of extensive reading on language competencies (cognitive view) as well as students' reading preferences (pedagogical view). Nation and Waring criticize that Day and Bamford do not clearly distinguish between these two views. I see this point, but in my opinion the previous explanations of the principles of extensive reading have shown that these views cannot be viewed independently. Since extensive reading is a way of teaching reading, it has a clear pedagogical component and direction that should be based on cognitive theory and empirical evidence. The latter is something where I see a strong development in this field, but still great potential as well. Nation and Waring (2020) name a long list of potential areas of research at the end of their book that are still unanswered. The research of this study will contribute to the following questions that Nation an Waring raised:

- What are the favorite books and genres students read?
- Do students prefer 'native texts' or graded readers?
- How do learners from different language groups / ability levels benefit from extensive reading?
- Can all learners benefit from extensive reading?
- How do gains in extensive reading compare with those from other 'rich' forms of input?
- How does extensive reading impact other language skills? Which skills? Why? Are there more effective methods than extensive reading affecting these skills?

2.3.4.5 Reading Material

The different principles of extensive reading speak of appropriate or motivating 'material' or 'texts' the readers should be able to choose from (see table 2.1 Williams (1986) principle 1; Day &

Bamford (1998, 2002) principle 1 and 2; Nation and Waring (2020) principle 3). Appropriate material is defined by the authors as texts that contain almost no unknown words – a maximum of 2% (Nation & Waring, 2020) or two words per page (Day & Bamford, 2002). Thereby, the texts are comfortable and motivating to read. Additionally, there should be a wide variety of different genres and topics available for the students to choose from (Day & Bamford, 2002; Nation & Waring, 2020).

There are basically two options at hand, when selecting texts for an extensive reading program – original L1 literature or graded readers written for L2 readers, which are written to be at the readers' level. Especially literature written for L1 children is named as a popular option, because of their simple language (high frequency words, vocabulary and pattern repetition), shortness and therefore appropriate length, illustrations supporting understanding and the parallels to first language acquisition (Webb & Macalister, 2013). Yet, these texts also contain difficulties: There is usually not a set word family that the vocabulary in children's books is from, because the books are written independent from one another. Moreover, the L1 vocabulary of children might exceed the vocabulary of the L2 reader and too many unknown words will slow down reading speed. Consequently, the texts might not contain enough known words to reach the necessary 98% of known words suggested by Nation and Waring (2020) for sufficient comprehension. Moreover, the texts are age specific and might not be interesting any more for L2 learners (Nation & Waring, 2020; Webb & Macalister, 2013). Nation and Waring (2020), therefore, strongly argue to use graded readers written specifically for L2 learners as reading material.

To evaluate the use of children's texts for L2 reading, Webb and Macalister (2013) analyzed vocabulary size and the possibilities of incidental learning when reading L1 literature. Using corpus-driven analyses, the authors compared vocabulary between books for L1 children, graded readers for L2 language learners and L1 readers for older children. Their results show that L1 children's literature and L1 readers for older children did not differ significantly in vocabulary and required a vocabulary of + 10 000 words. L2 graded readers only require approximately 3 000 words. The authors come to the conclusion that L1 literature cannot fulfill the aim of using texts that are easy enough to understand 98% of the words. However, they also state that there are other factors influencing comprehension like background knowledge or illustrations.

The criteria of the number of headwords, by which graded readers are commonly classified by is not enough. Gillis-Furutaka (2015) witnessed this problem when speaking to her students about their reading experiences. Even though her students were reading at their competence level, they still had problems understanding books and became frustrated. She ran a qualitative study, using interviews and think-aloud protocols to find out what made reading these graded readers difficult. As factors that limited understanding and that need to be taken into account when writing and classifying graded

readers, she identified the age of the reader, the necessary cultural background knowledge, idiomatic and figurative language that was used, literary devices that make it difficult to understand the story, missing or misleading illustrations and confusing plot structures. Having used graded L2 readers as a teacher, I can relate to these points and the outcomes of the interviews. Additionally, I sometimes have the impression that the simplification of vocabulary leads to more confusion concerning the plot and narrative at times, because these become more implicit and descriptions are not as precise anymore.

Allan (2008) compared graded readers (B1 and B2 level) and authentic corpora (BNC – British National Corpus) concerning the occurrence of collocations. All in all, she found a similar number of occurrences of the selected word clusters in B1 and B2 graded readers and the BNC. Differences between the texts only showed for longer clusters. The sample studies also showed that the set of collocations in graded readers is a subset of the BNC and that this subset is smaller for B1 than for B2. Allan's conclusions are that B2 graded readers – B1 graded readers only to some extent – may provide authentic input for L2 students. Yet, as she claims herself, these findings are only based on few sample analyses and need to be validated in more detail.

From a cognitive standpoint, especially when only focusing on vocabulary, the preference of graded readers is justified; from a pedagogical and practical point of view, there is more than meets the eye. First, the students in this study are beginners in their L2. Many readers available for them are still too difficult. Drawing back on very easy children's books with only a few words per page and many illustrations was, therefore, necessary in this study to ensure that books on the level of early beginners were available. Second, students - at least in this study - have very different reading experiences and abilities in their L1. The small font and long texts in graded readers for L2 reading is not motivating for many of these students. Books with large print and short stories are necessary – something available in L1 children's literature. Third, a claim for an extensive reading program is that there is a large variety of different texts and topics available. Especially non-fiction texts and comic books are difficult to find as graded readers, but are available for beginning L1 readers and as children's books. Fourth, graded readers might seem closer to what students read in their normal classes and might not be as motivating. In Germany, graded readers that fit exactly the vocabulary of the school book and even use the same characters as in the student's school books are available. These texts might be too close to the usual (intensive) reading done in class so that motivation and a feeling of pleasure do not arise. Fifth, there is a theoretical claim and there are a few theoretical analyses leading to the strong promotion of L2 graded readers, but these assumptions have not been validated by research yet. Therefore, one could argue that children's books and books written for early L1 readers are more appealing and authentic for L2 readers, because they differ from their usual classroom texts. Reading these books could be more motivating, which is an important asset of successful language learning.

There are no studies comparing extensive reading groups using 'authentic' L1 literature to those using graded readers. On the contrary, prominent positive examples of students successfully engaging in extensive reading (Day & Robb, 2015), or prominent studies of extensive reading programs (e.g.Elley & Mangubhai, 1981; see also section 2.4.3.1), named as positive examples (e.g. Nation & Waring, 2020), used L1 children's books as extensive reading material. The material was or is used due to a lack of graded readers in the target language, but these studies show that it worked. Therefore, these L1 books cannot be as inappropriate as stated.

For the reasons discussed above, not only a wide variety of genres and topics are used as reading material, but also a wide variety of book types: baby board books, children's books, early L1 readers, L2 readers at various levels and young adult L2 literature (see appendix for a full bibliography and chapter 3 for the management of the library). Thus, the students can decide, what they prefer to read. Additionally, Renandya, Krashen, and Jacobs (2018) argue for the use of more serial books that keep students interested in reading, because they identify with the characters and stories. If books seemed interesting to include in the library and these belonged to a series, I tried to make sure to include various books of these series, if available. Therefore, numerous fictional and non-fictional books of different series are included in the library of this study. This study will have an explorative look at the reading preferences of students in the intensive and extensive reading interventions that are presented in section 4.7 and discussed in chapter 5.

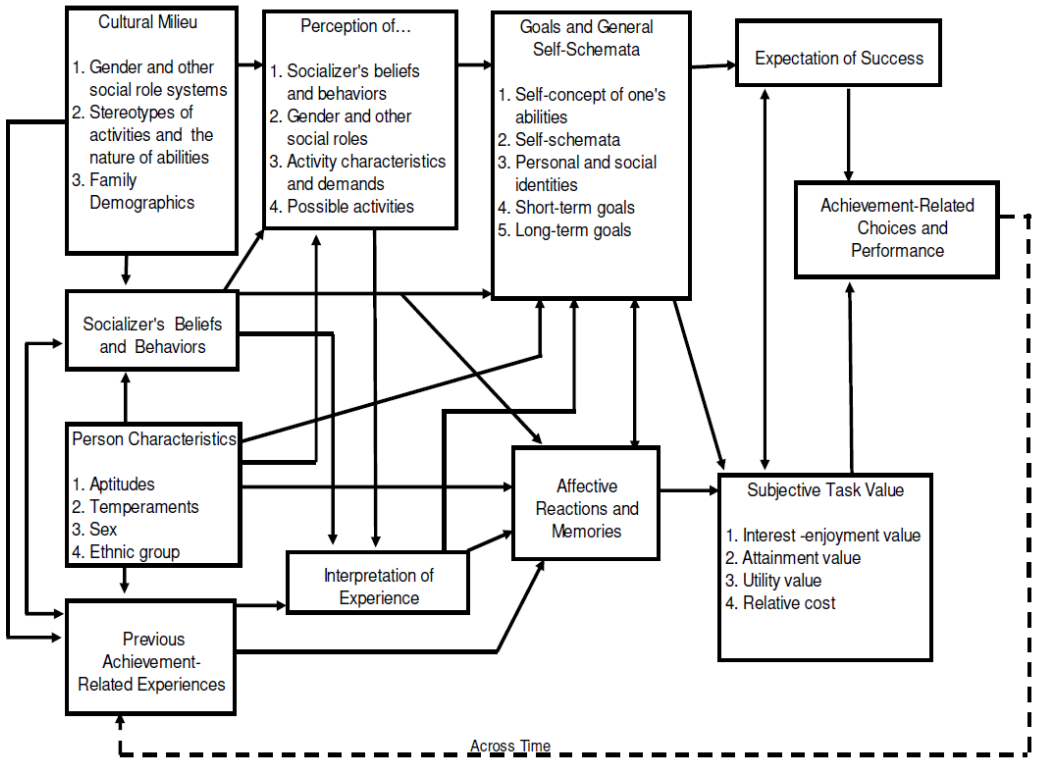
2.3.4.6 Motivation

Section 2.2 on reading processes and the reading model presented there, had a strong or exclusively cognitive-linguistic view on reading, like most reading models. Yet, motivation is also an important, if not core, asset of extensive reading programs. On the one hand, motivation has proven to be a crucial factor for improving students reading amount and comprehension and, consequently, their reading skill development (Grabe, 2009; Wigfield & Eccles, 2000). Therefore, motivation could influence the effects of intensive and extensive reading interventions on reading comprehension and language competence gains. On the other hand, classroom settings can influence reading motivation (Grabe, 2009; Wigfield & Eccles, 2000); On that account, intensive and extensive reading interventions could also have an impact on students' motivational factors. The following motivational models show possible relationships between these factors.

From a psychological point of view, one way of defining achievement motivation is as the "the individuals' expectancies for success and the value they have for succeeding" (Wigfield, 1994, p. 50), commonly named expectancy x value theory. A model that was developed on the basis of this idea is Eccles and Wigfield's (2020) Expectancy Value Model of Achievement Choices (see figure 2.8). This

model is of a general nature and not specifically aimed at reading, but can be used to explain which motivational factors could be involved in what way when looking at attitude towards reading and reading behavior.

The Expectancy-Value Model of Achievement Situations (Wigfield & Eccles, 2000) and the further developed Situated Expectancy-Value Model by Eccles and Wigfield (2020) is, as the name already states, an expectancy-value model that focuses on achievement situations as we find them in school. The core of this model is the idea that the *subjective task value* and *expectations of success* influence *achievement-related choices and performance*. The influence of different behavioral and normative beliefs that influence these expectations and values is represented in the various factors influencing these two central components of the model, see figure 2.8.



Reprinted from From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation, Contemporary Educational Psychology, 61, p.2 Eccles, J. S., & Wigfield,A. ©(2020), with permission from Elsevier.

Figure 2.8 Situated Expectancy value model of achievement choices (Eccles & Wigfield, 2020, p. 2)

For this study, I would like to concentrate on the right side of this model and directly apply it to the reading interventions. The *subjective value* that the reading intervention has for a student and the *expectation* he or she has of being successful in this intervention will influence the reading process and reading performance: If a student believes that L2 reading will help him improve his L2 language competence and therefore his grades (expectancy) and this competence or these grades are very important to him (value), then this will positively influence his attitude towards reading, lead to a

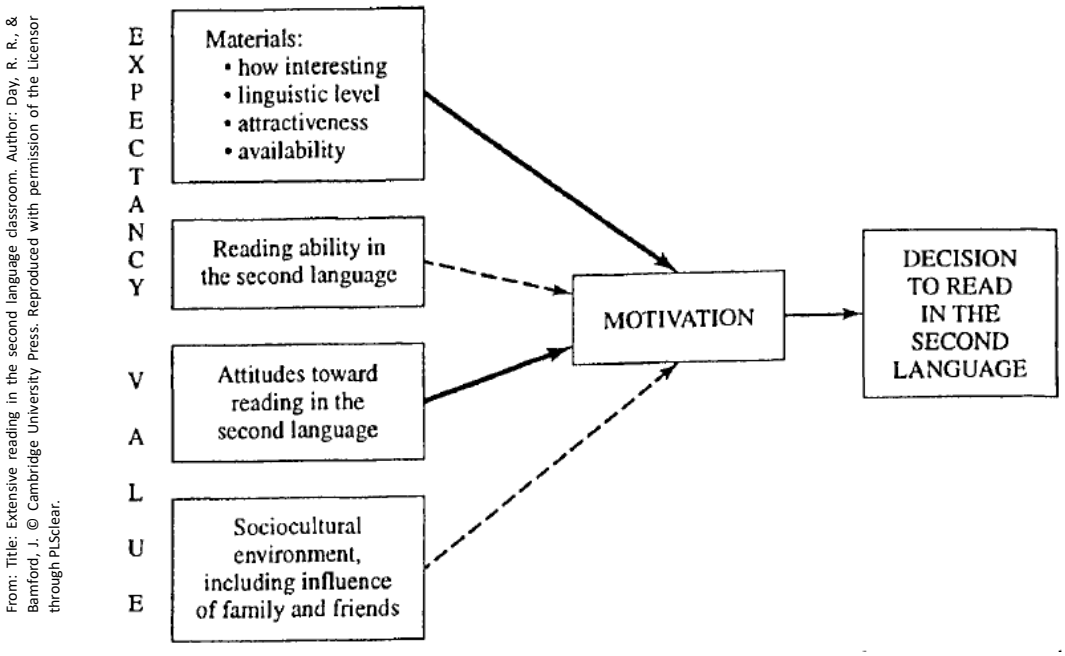
higher intention to read and then also to reading. If he believes that his peer group does not like students that read and have good grades (expectancy) and belonging to the peer group is very important to this student (value) then this is going to negatively influence his subjective view of reading and probably also going to influence his intention to read and reading itself. Which of these two effects is stronger, will depend on the relative importance of these two. In the extensive and intensive reading interventions in this study, where students work on their own and others do not necessarily notice how much somebody has read or how well their portfolio and tasks are answered, the subjective norm will probably not be as strong a factor as in an open in-class discussion.

These two factors are, among others, influenced by *goals and general self-schemata*. The effects of intensive and extensive reading interventions could be different, depending on the students' goals and self-schemata, as the following example shows: A student might value learning to read English, or becoming good in English, but have a very low English self-concept and the goal to not let others know that she is not a good reader (performance-avoid goal orientation). This could lead her to the expectation that she will not manage to read and understand the books presented in the library. In consequence, this student might rather value a task where it does not show that she is not good (extensive reading) than having to hand in tasks on her reading (intensive reading). Moreover, she might turn to very easy books, below her level and, consequently, might not learn as much as students with a higher self-concept. This is just one fictional example and possibility of how self-concept and goals can influence reading and reading related achievements. What becomes clear is that goals and self-concept are central factors influencing reading and reading behavior. These are therefore the two motivational factors I will focus on in this study and they will be presented in detail in section 2.3.5.

Day and Bamford (1998) also refer to the expectancy x value theory when talking about reading motivation induced by extensive reading programs, but their terminology and interpretation is a little different: Day and Bamford describe the expectancy of extensive reading courses as the material (how interesting, linguistic level, attractiveness and availability) and the L2 reading ability. The value is made up of the attitude towards reading in the L2 and the sociocultural environment.

In my opinion, this interpretation is not completely compatible with common general expectancy-value models (Ajzen & Fishbein, 1980; Eccles & Wigfield, 2020; Wigfield & Eccles, 2000). First, I do not agree that material is an *expectancy*, I believe it rather to be a *value*. One of the components of *value* is intrinsic value (see figure 2.8) – the enjoyment of reading a book - and another is utility value – if the book fits to one's individual plans. Therefore, if reading material is interesting, attractive, available and at my level, it will have a high *value* to me. Second, I do not see why attitude towards reading is only a value (see figure 2.9). Ajzen and Fishbein (1980) clearly state that attitudes are a result of expectancy-value considerations. The expectation (how good my L2 reading is) can

influence my attitude (e.g. “I don’t like reading in English, because I am not good at English and therefore can’t understand the stories.”), just as well as my values (how important reading English is to me) can influence my attitude (e.g. “I like reading English books, because it will help me to become better in English”). Third, as can be seen very clearly in the expectancy-value model, the sociocultural environment has an influence on both – expectancy and value.

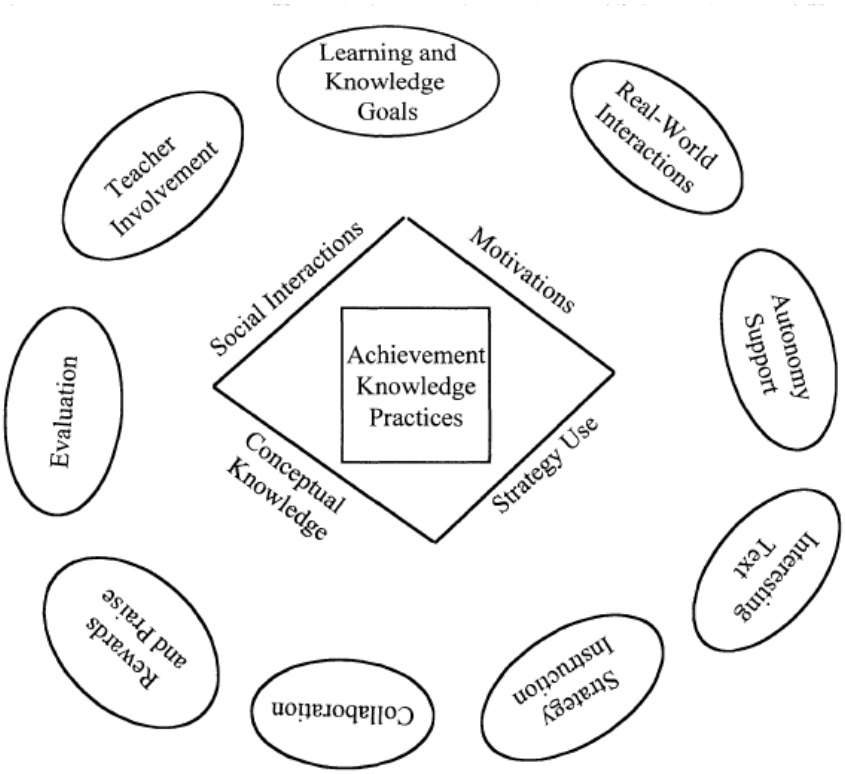


From: Title: Extensive reading in the second language classroom. Author: Day, R. R., & Bamford, J. © Cambridge University Press. Reproduced with permission of the Licensor through PLSclear.

Figure 2.9 Model of the major variables motivating the decision to read in a second language (Day & Bamford, 1998, p. 28)

A didactic model describing reading engagement is presented by Guthrie and Wigfield (2000), see figure 2.10. The outer circle represents factors of reading instruction that influence the inner square. Factors of the outer triangle can moderate these effects. This model is especially useful for showing differences between the intensive and extensive reading courses and their potentially different effects in more detail: Both reading intervention types use the same books and are therefore not likely to differ, concerning *real-world interactions* and *interesting texts*. Moreover, students work individually and silently in both intervention types, which means that there are no differences in *collaboration*. Intensive and extensive reading interventions are likely to differ in the salience of *learning and knowledge goals* (see section 2.3.5.1), the *autonomy and support* (dictionary use, control through weekly portfolio correction and feedback), logically the *strategy instruction*, *reward and praise* (pizza party for intensive reading classes and weekly feedback, also in form of praise), *evaluation* and *teacher involvement* in a way that the portfolios provided feedback and weak students received individual support for their correction. Therefore, not only from a cognitive point of view, but also from

a motivational and psychological point of view, intensive and extensive reading interventions are likely to strongly differ concerning their effects on reading processes and achievements.



From: Title: Engagement and motivation in reading. Author: J. T. Guthrie & A. Wigfield. In Title: Handbook of reading research, Vol. 3. Author: M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr. © Taylor and Francis Group LLC (Books) US. Reproduced with permission of the Licensor through PLSclear.

Figure 2.10 Engagement model of reading development (Guthrie & Wigfield, 2000, p. 410)

2.3.5 Effects of Goal Orientation and L2 Self-Concept on L2 Reading and Comprehension

As the previous sections have shown, goal orientation and self-concept are important factors in the reading and language learning process. The following sections will provide a summary of the psychological concepts goal orientation and self-concept. Moreover, previous research of the different effects presented in the previous models and analyzed in this study is summarized.

2.3.5.1 Goal Orientation

Background Information on Goal Orientation

Students can engage in activities, like reading, for various reasons (motives). One way of describing these different motives is by referring to different goal orientations. Different strands and stages of research use similar terms for goal orientations. I will, therefore, give a short overview of the terminology first. The first goal orientation models presented were two factor models, but whereas Dweck (1986) talks about *learning goal orientation* and *performance goal orientation*, Nicholls (1984) talks about *task orientation* and *ego orientation*. The two-factor model was then refined into a three-factor model by differentiating performance goal orientation into *performance-approach goal*

orientation and *performance-avoidance goal* orientation in addition to *mastery-approach orientation* (Elliot, 1999; Elliot & Church, 1997; Elliot & Harackiewicz, 1996) or using Vandewalle's (1997) terminology: *learning goal orientation*, *performance-prove goal orientation* and *performance-avoid goal orienatation.* There have been further developments with more factors (e.g. Elliot and McGregor (2001) with a 2x2 model; Elliot, Murayama, and Pekrun (2011) with a 3x2 model and Daumiller, Dickhäuser, and Dresel (2019) or Vansteenkiste, Lens, Elliot, Soenens, and Mouratidis (2014) with further devisions of learning and performance-approach goals), but the three factor model has remained the primary version used in research. For a detailed summary of the historical development and differences between these different models see Vandewalle, Nerstad, and Dysvik (2019) and for a meta-analysis on terminology and operationalisation see Hulleman, Schrager, Bodmann, and Harackiewicz (2010). I will refer to the termininolgy and descriptions used by Spinath and Schöne (2003), because these correspond to the SELLMO scales used for measuring goal orientation in this study. These are *learning goal orientation*, *performance-approach goal orientation*, *performance-avoid goal orientation* (3x2 model) and an additional factor *work avoidance*.

If students read to broaden their horizon and to discover new topics or stories, they pursue a *learning goal orientation*. There are also reading programs where students collect points for every book they read, then are invited to pizza and popcorn parties, they are rewarded with medals and trophies and receive a place on the wall of fame. In this case, the motive of some students to read many books might primarily be to demonstrate their abilities and perform well in comparison to others or even win the school reading competition, namely a *performance-approach goal orientation*. Some students might also read many books to prepare for an upcoming class competition because they fear that they might not succeed and people might notice their deficits. To avoid this negative situation, the students practice reading beforehand to avoid attracting negative attention. This motive to avoid situations that might show incompetence or inabilities is called *performance-avoid goal orientation*. Of course, this means that the situation itself might be avoided as well – a student believes that he will perform badly in the reading competition and, therefore, does not attend class on that day. Some students might also show a low reading engagement, but not with the motive to hide inability, but because they pursue other goals that are more important to them, for example, extensively practicing an instrument or engaging in sports. This would be a *work avoidance goal*. As the example already shows, work avoidance does not necessarily mean the avoidance of effort in general, though this could be the case, it could also be the avoidance of work in one learning situations to pursue other goals.

Even though there seems to be a strong connection between goal orientation and intrinsic and extrinsic motivation (Heyman & Dweck, 1992), associating learning goal orientation with intrinsic motivation and performance goal orientation with extrinsic motivation is not suitable for all contexts (Spinath et al., 2012), for example when competition is a crucial element of the task itself, like playing

a game of chess. If I want to be a good chess player, I am not interested in loosing. Winning chess matches is a crucial element of being good and learning to play well. It is, therefore, clear that I have a performance goal orientation, but could have an intrinsic motivation – I am not playing chess for the sake of winning, but for the fun, which includes winning. Yet, this equation can work in other contexts, for example, when looking at the example of the reading competition again: If reading is practiced to perform well in the competition, then it is probably driven by extrinsic motivation, whereas reading for fun is an intrinsic motive. Thus, it is probably sound to assume that in this study, interventions that lead to an adaption of learning / performance goals can be associated with a higher intrinsic / extrinsic motivation.

Even though goal orientations are commonly analyzed as singular components, people can pursue different goal orientations at the same time and different tasks could enhance pursuing both learning and performance goals (Spinath et al., 2012). For example, students could be proud to attend the exclusive pizza party (performance-approach goal orientation), but also have an interest in the books they read and enjoy discussing the content with their friends at that party (learning goal orientation).

Effects of Goal Orientation on Achievement

Extensive reading intervention programs emphasize the importance of reading being its own reward, students' reading on their own at their own speed and choosing texts of their interest. In contrast, intensive reading interventions demand close reading and often involve tasks where students are supposed to prove their competence that is then graded or rewarded. According to the expectancy-value hypothesis (Eccles & Wigfield, 2020; Wigfield, 1994; Wigfield & Eccles, 2000), students with a high learning goal orientation (expectancy: goal to improve reading or language skills) and maybe also students with a high performance-avoid goal orientation (expectancy: goal to avoid showing poor reading or language skills) might therefore achieve better results in the extensive reading intervention (higher subjective task value) than when participating in the intensive intervention (lower subjective task value). Students with a high performance-approach goal orientation (expectancy: goal to show good language or reading performance skills) might achieve better results in the intensive reading intervention (higher subjective task value) than when participating in the extensive intervention (lower subjective task value).

This leads me to the following hypothesis:

5) Goal orientation moderates the effect of the reading interventions: Students who have a stronger learning goal orientation profit more from the extensive reading intervention (higher competence gain) than students with a higher performance-approach goal

orientation, who profit more from the intensive reading intervention (higher competence gain). These effects are due to a better fit of the task to the student's motives.

Meta-analyses have shown that learning goal orientations and performance-approach goal orientations are positively associated with achievement (r = .14 and r = .10), whereas a performance-avoid goal orientation is negatively associated with achievement, with r = -.13 (van Yperen, Blaga, & Postmes, 2014). These correlations are only small and do not show an advantage of pursuing a learning goal orientation compared to a performance-approach goal orientation for achievement. Payne, Youngcourt, and Beaubien (2007) arrived at slightly different results in their meta-analysis: They did not find effects of performance-approach goal orientation on learning or academic performance, but high correlations between state learning goals and performance-approach goals.

Previous research has also found various correlations between goal orientation and effort and persistence during a task, feedback seeking and feedback-processing behavior, metacognition and training transfer. Learning goal orientation was always associated positively with these factors, whereas performance-approach goal orientation was uncorrelated and performance-avoid goal orientation correlation results were negative (Hulleman et al., 2010; Vandewalle et al., 2019). Since the language tests, that the students fill in before and after the reading comprehension, are constructed to be rather complex so that differences in competence can be measured, students with a learning goal orientation could be at an advantage in these tests. This study uses pre-posttest differences for analyzing effects, therefore, this should not be a problem, because only competence gains are used for the analyses. Thus, students who might perform well in the tests due to a high learning goal orientation will not achieve higher competence scores.

Additionally, students pursuing a learning goal orientation set goals that are aimed at developing their skills and engage in more deep level processing strategies, like extending ideas and applying knowledge. Students pursuing a performance-goal orientation tend to set goals that involve comparing well to others and use cognitive strategies that process information rather on the surface level, like memorizing and rehearsing facts (Vandewalle et al., 2019). Moreover, if a task requires a high working memory, students with a learning goal orientation are at an advantage when solving the task and are also more likely to cooperate and share their knowledge (Vandewalle et al., 2019). Learning goal orientation, therefore, could facilitate the use and development of adequate (higher-level) processing strategies while reading and boosting the effects of reading interventions, leading to higher competence gains. I thus hypothesize that

6) learning goal orientation has an overall positive effect on students' competence gains in the different language tests. This effect is independent of the intervention group type.

2.3.5.2 Self-Concept

If a student feels that her English abilities are high, that she is able to improve or that she has a higher competence than others, then this student has a high English *self-concept*. If a student receives an English task and believes he will perform well in this task, then he has a high *self-efficacy*. These positive self-beliefs, have shown to influence the students' academic life and achievements positively in various ways (Marsh & Craven, 2006; Valentine, DuBois, & Cooper, 2004). On the one hand, positive experiences through extensive reading activities that enhance the students' self-esteem concerning their English abilities could lead to a more positive self-concept in English (skill-development view). On the other hand, the student's already existing positive or negative self-concept could influence their achievements in English tests (self-enhancement view). Therefore, the English self-concept of the students is considered an important factor in this study and measured before and after the English intervention.

The following sections define self-concept, describe general aspects of this psychological concept, point out similarities and differences to self-efficacy and summarize studies investigating the impact of intensive and extensive reading interventions on students' self-beliefs.

Definition

Self-concept, if dated back to James (1890), has already been in the focus of psychological research for over a century (Möller & Trautwein, 2015). Consequently, numerous definitions with large variances have been brought forward since (Bong & Clark, 1999). Bong and Clark (1999) deducted a number of underlying aspects that these definitions have in common and came to the conclusion that self-concept is "one's perception of the self that is continually reinforced by evaluative inferences and that it reflects both cognitive and affective responses" (Bong & Clark, 1999, p. 140). Bong and Skaalvik (2003, p. 2) simply define self-concept "as a composite view of oneself".

Schöne et al. (2012), whose SESSKO scales are used in this study, define self-concept as the entirety of cognitive representations of one's own abilities in academic performance situations. This definition is quite similar to the ones above, with the difference that it only names cognitive representations and not affective responses as an entity of self-concept. This differentiation between a cognitive representation of the self (with its attributes) and affective components (which result from evaluating the cognitive self-concept) is well established (Möller & Trautwein, 2015). Furthermore, an additional differentiation of the cognitive component of self-concept in descriptions and evaluations

is suggested (Bong & Clark, 1999). The affective component strongly depends on the result of evaluative processes: Two students that are not very good in English (negative cognitive self-concept), might show very different affective reactions. The first student, for example, might not have a low affective self-concept, because she is very good in math and science and these subjects are very important to her; additionally, the English performance might not be very important to her parents. Maybe she might even perceive her low cognitive English self-concept as proof for her mathematical and scientific abilities. The second student might have parents for whom it is very important that the student participates well in their English class and learns English well for his future. He might struggle in many different classes and knows that English is very important for a good job in an international company. He sees that, already at school, he cannot compete with his classmates, probably leading to a lower affective self-concept than the first student's. For this reason, Schöne et al. (2012) only included items measuring the cognitive component of self-concept in their scales.

Although these three definitions are very broad, they differ in their conceptualization of self-concept and do not name specific aspects that help to understand what self-concept is concretely. As displayed in the two examples above, there are several factors that affect a person's self-concept. Shavelson, Hubner, and Stanton (1976, p. 411) believe that "[s]even features can be identified as critical to the construct definition. Self-concept may be described as: organized, multifaceted, hierarchical, stable, developmental, evaluative, differentiable". Additionally, the self-concept of a person is influenced by social comparisons (frames of reference), causal attributions, the way someone believes to be viewed by others (reflected appraisal from significant others), previous experiences and the resulting expectations of mastering future challenges or the perceived ability to reach certain goals (mastery experience) and whether somebody is able to perform well in areas that are important to him or her (psychological centrality) (Bong & Skaalvik, 2003). I will go into detail about some of these aspects in the following sections.

Concept

As already mentioned above, self-concept can be divided into a cognitive component and an affective component (Bong & Clark, 1999; Möller & Trautwein, 2015). Additionally, the previous definition of Shavelson et al. (1976) characterizes the concept as multifaceted and hierarchical. This means that the different experiences and perceptions of self, are grouped into categories that are organized in a hierarchical structure. At the lowest domain level, there is the perception of the self in specific situations. Together, these experiences make up the self-concept in a subdomain, for example certain subjects or physical abilities and appearance. Combined, they lead to a self-concept in different areas, for example an academic, an emotional, a social and a physical self-concept. Together, these

various self-concepts form the general self-concept (Möller & Trautwein, 2015). Empirically, this view could only partially be validated (Marsh & Craven, 2006). Some studies have even found negative correlations between self-concepts, suggesting a differentiation of the academic self-concept into a math and a verbal self-concept (Bong & Clark, 1999; Möller & Trautwein, 2015).

Additionally, age is an important factor when assessing self-concept. Marsh and Craven (2006) report that if self-concepts highly correlate for young children, this does not change with age. Initially uncorrelated self-concepts, on the contrary, develop into stronger negative correlations with increasing age (differentiation-distinctiveness hypothesis). Therefore, the older a child gets, the more differentiated and stable the self-concepts seem to be (Marsh & Craven, 2006; Möller & Trautwein, 2015).

Comparison Frames: Internal/External Frame-of-Reference-Model and Big-Fish-Little-Pond-Effect

To develop a self-concept in a specific domain, students need information to judge and evaluate their performance in this domain. This information can derive from criterial comparisons, but these are usually not available. A much stronger frame of reference for the self-concept are therefore social, but also individual or domain comparisons.

Evidence for the strong social orientation and its impact on self-concepts is the Big-Fish-Little-Pond-Effect. A big fish feels really strong and superior, if he swims in a little pond, but when put into the large ocean, he might realize that he is only a very small inferior creature. These effects have also been found concerning the self-concept in school classes: Students with the same achievements and capabilities have a higher self-concept when they are placed in a low-achievement group, than if they are among high-achievers (Bong & Clark, 1999; Bong & Skaalvik, 2003; Möller & Trautwein, 2015).

The fact that self-concepts develop on the basis of social comparisons is a widely accepted view. Therefore, positive relationships between (domain-specific) self-concept and academic achievement are often presented as evidence for existing social comparisons in studies. However, as already mentioned above, previous studies have found zero or negative correlations between math and verbal self-concepts (Bong & Clark, 1999; Marsh & Craven, 2006) or even similar subjects like L1 and L2 (Möller, Zitzmann, Helm, Machts, & Wolff, 2020). This is surprising because these effects also show, when the actual achievement in both subject areas is high (Möller et al., 2020; Möller & Trautwein, 2015). To explain this effect, it is hypothesized that, in addition to the social and external frame of reference, students use an additional, internal frame of reference between different domains. If a student performs better in math than in English, he/she will probably devaluate his/her achievements in English, even if these are high as well. This effect is called the Internal/External Frame-of-Reference Model, short I/E-Model. Since the students in this study also participate in other mathematical or

scientific classes, positive and self-concept enhancing experiences in those classes might have a negative impact on their English self-concept.

Thus, I hypothesize that

7) English (reading) interventions will have a positive impact on the students' English self-concept (self-concept gain), whereas non-English interventions will have a negative impact on the students' English self-concept.

Self-Enhancement vs. Skill-Development

The two opposing views of self-enhancement and skill-development relate to the question whether a positive self-concept leads to positive academic experiences (self-enhancement) or whether positive academic experiences lead to a positive self-concept (skill-development). There is empirical evidence for both views.

A strong argument for the skill-development view is the big-fish-little-pond effect that clearly shows that different experiences and social comparisons lead to the development of different self-concepts. This could mean that the reading interventions in this study, where students make first (positive) experiences with extensive or intensive reading in English and are in a different class environment than in their regular classes, could have a positive effect on their English self-concept.

However, there is also evidence for the self-enhancement view. Valentine et al. (2004), for example, showed in their meta-analyses that even when controlling for previous achievement, self-concept still had a small but significant influence on future achievement. This effect was especially high for achievements that were in the same domain as the measured self-concept. Therefore, in this study, you would expect students with a high English self-concept to have higher competence gains on English proficiency tests than students with a low English self-concept, independent of the intervention. Thus, the following hypothesis will be tested:

8) Students' self-concept at the beginning of the interventions (T1) will correlate significantly positively with their test score gains and significantly negatively with their average grade difference (meaning a positive relationship of self-concept and grade improvement due to lower grades being better than higher grades). This effect is independent of the experimental group the students are in.

The Reciprocal-Effects-Modell (Marsh & Craven, 2006) suggests that these two opposing views both seem to have their justification. There is also evidence that the extent to which either the effect of the skill-development or the self-enhancement effect is stronger is age specific. The skill-

development model seems to apply better to younger children, whereas the self-enhancement model was found to be more dominant in studies with older children (Bong & Clark, 1999). This is not surprising, since younger children still need to develop their academic self-concept, whereas older children might already have a quite differentiated self-concept that has an impact on their academic actions and results. Since I am assessing grade 5 and grade 6 students, this difference is interesting for my study. Even though students have already made first experiences learning English in primary school, they start learning English extensively when entering secondary school in grade 5. Thus, their self-concept in English may not be as developed in grade 5 as in grade 6 and the skill development effect might show to be stronger.

Self-Concept and Self-Efficacy

There is very little literature available on the associations between extensive reading interventions and self-beliefs. Therefore, it is necessary to also draw on literature measuring self-efficacy. This makes it necessary to first shortly explain similarities but also decisive differences between these concepts.

The two self-belief concepts, self-concept and self-efficacy, are sometimes used synonymously while they are differentiated in other research (Bong & Clark, 1999). In a meta-analysis, Valentine et al. (2004) could not find any differences between effects of self-beliefs (self-concept, self-esteem and self-efficacy) on academic achievement. Yet, the type of self-beliefs differed in level of measurement (global, academic or specific). Marsh et al. (2019) call this problem jingle-jangle fallacies: "two scales with similar names might measure different constructs (jingle fallacy) while two scales with apparently dissimilar labels might measure similar constructs (jangle fallacy)" (p. 332). Comparing self-efficacy to self-concept is also difficult on a practical level, because the variety of self-concept definitions lead to a variety in self-concept scales. However, this variety is not as big for self-efficacy scales, leading to contrasting results (Bong & Clark, 1999). Self-efficacy seems to be closely related, if not equal to the cognitive dimension of self-concept. There is also empirical evidence for this assumption, for example, self-efficacy was able to predict self-concept scores (Bong & Clark, 1999). Moreover, the internal structure of these two concepts seems to be different. As already mentioned, one aspect of self-concept is that domain-specific academic concepts divide into uncorrelated verbal and math self-concepts but verbal and math academic self-efficacy are highly correlated (Bong & Clark, 1999; Bong & Skaalvik, 2003). Therefore, there seem to be a different underlying hierarchical order for these two domains. Consequently, studies on the I/E-Model usually found support for the model when looking at students' self-concept, but not for self-efficacy (Bong & Skaalvik, 2003).

Additionally, self-efficacy, as defined by Bandura (Bandura, 1977), is strongly focused on future expectations of mastering a specific task. Self-concept, in contrast, is more general and less task specific (Bong & Clark, 1999; Bong & Skaalvik, 2003; Marsh et al., 2019). This also reflects in methodological aspects of research of these two concepts: Both concepts are usually measured using self-reports (Bong & Skaalvik, 2003), but self-efficacy items usually involve the judgement of certain and specific tasks that a student should solve, whereas self-concept items involve a judgement of their general ability in a certain domain (Bong & Clark, 1999; Bong & Skaalvik, 2003). The question is, whether these concepts would actually resemble each other, if they were measured on the same domain level. Additionally, self-concept is perceived as a more stable concept than self-efficacy (Bong & Skaalvik, 2003), which could be a consequence of the differences in measurement. These different foci also show in the choice of dependent variables: Studies on the impact of self-concept usually use standardized test scores or grade point averages, whereas studies on self-efficacy effects use specific school tasks (Bong & Clark, 1999; Bong & Skaalvik, 2003). This makes it difficult to compare these results concerning the impact of the two constructs. Additionally, studies on the impact of self-concept are often correlational designs, whereas self-efficacy studies are often experimental (Bong & Clark, 1999). These aspects illustrate that different research conventions have developed in the two fields.

Another difference lies in the comparison frame of reference. As already described above, a person's self-concept usually develops from social or individual comparisons of the self to others or to a previous self. Self-efficacy judgments, on the contrary, seem to develop by making comparisons to absolute standards (Bong & Clark, 1999; Bong & Skaalvik, 2003; Marsh et al., 2019). This also shows in stronger Big-Fish-Little-Pond-Effects for self-concept than self-efficacy (Marsh et al., 2019). Since scales on self-concept and self-efficacy are often worded to trigger these comparisons, Bong and Clark (1999) remark that "[t]here is much room for debate whether social comparison differences between self-concept and self-efficacy are inherent in construct definitions or are largely an artifact of different measurement traditions" (p. 143). Therefore, a big strength of the scales by Schöne et al. (2012) used in this study is that they differentiate between a criterial, individual, social and absolute self-concept. Congruously, in their reported correlations between the self-concept scales and other concepts, self-efficacy had the highest correlations with criterial and absolute self-concept (r = .33), followed by the individual self-concept (r = .31) and the smallest correlation with the social self-concept (r = .26).

2.3.5.3 Previous Research on Self-Beliefs and Reading Interventions

Self-beliefs about one's own competence are an important asset of reading motivation (see section 2.3.4.6 on motivational and reading models). Still, as far as I know, there are no studies measuring the effect of extensive reading interventions on students' self-concept. In general, research

on the effects of extensive reading on self-beliefs is very sparce (Burrows, 2012; Nation & Waring, 2020; Park, 2020). Therefore, this section reports findings of the few studies available on the relationship of extensive reading and self-efficacy.

Self-Efficacy and Reading

Lake (2014) measured reading self-efficacy and reading interest (described as a more general and dispositional reading self) at the beginning and end of L2 extensive reading classes and L2 reading classes that additionally had to write blogs. These groups were not extensive reading classes, because the reading amount was too small and functioned as control groups. In total 244 female Japanese students, age 18-19, participated in the study. T-test results showed that reading self-efficacy and reading interest improved in both class types, but only significantly in the extensive reading classes ($t(123) = -7.06$, $p < .001$, $r = .54$ for self-efficacy and $t(119) = -4.73$, $p<001$, $r = .40$ for reading interest). Additionally, there was a significant positive relationship between reading interest and reading self-efficacy ($r = .52$, $p < .001$). In conclusion, extensive reading seems to have a positive effect on students' self-beliefs (self-enhancement) and this effect is stronger for more specific measures like self-efficacy than for more general measures like reading interests.

Burrows (2012) also looked at the impact of different L2 reading interventions on student's self-efficacy in his dissertational thesis. With 322 Japanese students, his study is the largest presented here. The participants were divided into extensive, intensive, reading strategy and extensive/strategy groups. Three times over the course, the students answered questionnaires on reading comprehension, reading self-efficacy, perceived utility of extensive reading, perceived utility of reading strategies and sources of reading self-efficacy. His study produced three major results concerning the relationship of self-efficacy and reading interventions: First, participants in the reading strategy and combined reading strategy / extensive reading groups had significantly higher self-efficacy gains than in the other two groups. Second, there was a positive relationship between gains in reading comprehension and gains in self-efficacy. Third, the reading amount was not related to gains in self-efficacy. There are some aspects that make it difficult to deduct assumptions from these results for my study. First, it is difficult to say in how far these results are generalizable and transferable to European students, since "Asian cultures have a weaker academic self-efficacy than their western counterparts" (Burrows, 2012, p. 84). Second, the questionnaire used in this study was developed by the author himself and also piloted, but there is no information on the piloting procedure and results. Therefore, it is unclear if internal consistency and divergent and convergent validity of the self-efficacy questionnaire is sufficient and in what way the self-efficacy measured relates to other scales and concepts. Third, intensive reading was only used as a control group, where students learned

vocabulary, read the same text aloud as well as silently and then completed exercises. This is very different from the intensive reading intervention in my study. Despite these points, the results of Burrows (2012) deliver some first hints at which possible effects and relationships can be expected in this study; Burrows did not find significant gains of self-efficacy for the extensive and intensive reading groups. Yet, results show that all groups improved their mean self-efficacy. The mean self-efficacy of extensive reading in the pre-test was non-significantly higher than in the other groups, but also increased from T1 to T2 to about the same level as the strategies groups. Therefore, the question whether extensive reading has a positive influence on self-efficacy remains unanswered after these results, but positive effects are possible.

McLean and Poulshock (2018) measured the effect of three different reading interventions on students' reading self-efficacy. Their participants were 59 Japanese first-year EFL learners at university (age 18-19) studying at the Department of Economics. Except for 5 students, the participants were all males. The students were divided into three groups: (1) the word target group, who were instructed to read at least 2 500 words per week outside of class, (2) the sustained silent reading (SSR) group, who needed to read 15 minutes a week in class and one book per week outside of school and (3) the comparison group, who were instructed to read one book per week outside of school. English reading self-efficacy (RSE) was measured using the instrument developed by Burrows (2012) in his dissertation. In paired sample t-tests for each group, only the word-target group showed a significant increase in RSE ($t(18) = 3.78$, $p = .001$). Additionally, the number of words read during the interventions did not correlate significantly with the RSE results ($p = .08$, $r = .19$). McLean and Poulshock (2018) do not discuss these results in detail, but mention that the word target group experienced more extrinsic motivation that was then internalized, according to the self-determination theory, and lead to higher self-efficacy. Unfortunately, they do not provide evidence or reasoning, why they can make these assumptions. I assume that one reason why McLean and Poulshock found different results than Burrows (2012) might lie in the nature of the reading the students engaged in. The RSE scales used in this study ask students how well they believe they can read and understand children's books, song lyrics, subtitles in American movies, a party invitation, an article in a newspaper, and so on. (See Burrows, 2012, p. 470). The word target group was the only group that was not instructed to read a book per week and was therefore more free in their choice of text and maybe engaged in a wider range of reading material. As already mentioned when defining self-efficacy, this self-belief construct is commonly used in a task specific way, as in this case. Therefore, I believe it would be interesting to reflect on what specific self-efficacy the RSE scales measure – in my opinion they assess the perceived ability of coping with various text types. You would only expect to find pre-post effects, if students have the chance to experience and practice reading different texts in the reading programs, which to a large extent depends on the corpus used.

Reading Attitudes

Park (2020) used the concept of reading attitude as in the model by van Schooten and Glopper (2002) to investigate the influence of intensive and extensive reading interventions on reading attitudes. Reading attitudes is again a different belief concept, which means moving even farther away from self-concept than already done with self-efficacy. Yet, these few studies represent the next related studies available in the field of self-beliefs and reading interventions. Thus, I would at least like to briefly present them.

Van Schooten and Glopper (2002) base their reading attitude model on the Model of Planned Behavior (MPB), distinguishing between five components: cognitive attitude like "Reading adolescent literature is a good way to relax yourself." (p. 186), affective attitude like "I'm not interested in adolescent literature." (p. 186), subjective norms like "A good student should be devoted to adolescent literature." (p. 187), perceived behavioral control like "I would like to read adolescent literature that I really like, but I do not know what books there are and which I would enjoy." (p. 188) and behavioral intentions like "I think it's a good idea to bring a work of adolescent literature if I have to wait in a waiting room or if I travel by train." (p. 188). The difference to self-concept and self-efficacy is on the one hand obvious for the cognitive attitude, which involves general opinions. The other components, on the other hand, show a stronger relation to self-beliefs: The affective attitude items could in some cases reflect the affective component of the self-concept. The perceived behavioral control and behavioral intentions could reflect personal implications and reactions of the cognitive self-concept, e.g. the thought "compared to other students I am not good at reading adolescent literature (social self-concept), therefore, to improve I should use time in waiting rooms or on the train for reading (behavioral consequence)". Van Schooten and Glopper (2002) found that cognition, affect and intentions influenced the time Dutch secondary students (grade 7-9) spend reading adolescent literature, whereas the subjective norm and the perceived behavioral control hardly influence the reading behavior.

The bridge between these L1 attitudes and L2 reading attitudes was built by Yamashita (2004), who investigated the relationship between L1 and L2 reading attitudes of Japanese learners of English. For this purpose, she developed a questionnaire differentiating between two components: an affective "I feel anxious if I don't know all the words." (p. 7) and a cognitive "I think reading many books is advantageous to getting a job." (p. 8) one. Whereas the cognitive attitudes were phrased as general beliefs in Van Schooten and Glopper's scales, here they are phrased as personal beliefs "I think…". These questionnaires were presented to 59 Japanese university students that rated the different statements for their Japanese L1 and their English L2. Factor analyses of the two scales showed that the items loaded at two subfactors for each scale, both for L1 and L2: comfort and anxiety for the anxiety statements and value and self-perception for the cognitive statements. Correlations of these

scales revealed that attitudes correlated positively, but descriptive statistics showed that these attitudes were more positive for L1 reading than for L2 reading. Additionally, proficiency does not seem to have an influence on reading attitudes. When looking at the impact of these attitudes on extensive reading, operationalized as the number of pages read during an extensive reading program, only comfort and self-perception significantly correlated with the number of pages read. Yamashita (2004) concludes on the basis of these results that attitude seems to affect students' extensive reading performance, but since this is not a pre-/posttest design it is difficult to say whether the amount of reading is a consequence of the reading attitudes or whether the reading attitudes were already a consequence of previous, pre-experimental, experiences and part of a reciprocal process.

Park (2020) used an adapted version of Yamashita's (2004) scales in his study on the impact of reading interventions on reading attitudes. He investigated the influence of reading interventions on reading attitudes with 73 Korean school students (age 15-16). He found that extensive reading, compared to intensive reading interventions, increased positive reading attitudes. He was also able to confirm the results that proficiency levels did not have a significant influence on the development of reading attitudes.

Conclusions

Studies from the psychological field on the effects of self-concept and self-efficacy on academic performance are well established, but they do not investigate concrete pedagogic interventions, let alone L2 reading interventions. Therefore, it is only possible to assume from general effects and theories, which effects might be found for L2 reading interventions in this study. There are the studies presented above on relationships between reading interventions and self-efficacy or reading attitudes, but it is difficult to relate these to the established psychological field or even to each other. Park (2020), for example, wrote "[t]he model was selected because it effectively illustrates L2 learners' reading attitude utilizing the three-dimensional reading attitude framework." (p. 340), but van Schooten and Glopper (2002), whom he is referring to, not only use a five-dimensional model in their study, but also conduct their study with Dutch L1 students and their L1 reading. Park does not mention why this can be transferred to L2 reading. He furthermore states that "the current study adapted Yamashita's (2004) instrument, which is firmly grounded in van Schooten and de Glopper's (2002) three-dimensional framework of reading attitude" (p. 244) without further explanation. As already mentioned above, I found major differences in the operationalization of the two scales used in these studies. Furthermore, Yamashita (2004) does not seem to perceive her scales as 'grounded' in van Schooten and Glopper (2002), as this model is not even mentioned in her paper. These two examples unfortunately are not the only examples of inconsistencies in previous research (see also Nation and

Waring (2020) sections on 'Insufficient Reporting' and 'Citing the Work of Others' in chapter 11). Moreover, studies on extensive reading are often conducted in Asia, therefore, psychological concepts and their effects might not be comparable to those in Europe. This study will contribute to the field by analyzing the reciprocal relationship of self-concept and language/reading skill development for different reading interventions in the European contexts.

2.4 Summary, Research Question and Hypotheses

Reading is a complex process involving various skills and processes (see section 2.2.2 and 2.2.3). L1 reading experience and L2 language competence add complexity to the L2 reading process (see section 2.2.4). In this study the students participate in extensive reading interventions (see section 2.3.2), where they simply read as much as possible, and intensive reading interventions (see section 2.3.1), where their reading is accompanied by a portfolio with language and comprehension tasks and exercises. It stands to reason that these interventions involve different forms of reading, consequently influencing the reading process and comprehension in different ways. The following table summarizes differences between intensive and extensive reading interventions on the basis of cognitive and motivational theories.

Table 2.2 Intensive vs. Extensive Reading

Intensive Reading Interventions	Extensive Reading Interventions	Sections
• same reading material, all books of the library available to all students		2.3.1
• students chose their books individually and worked/read at their own pace		2.3.2
• dictionary available	• no dictionary available	
• worksheet for each book	• no additional tasks	2.3.4
• evaluation of the books (evaluation sheet)	• no evaluation of the books	2.3.4
• pizza party as a reward	• no rewards	2.3.4
• individual assistance of weak students who need help with their portfolios	• only explanations or help when students come with questions	2.3.4
Implications for the following cognitive and motivational differences:		
• implicit[9] (reading) and explicit[10] (e.g. vocabulary tasks) L2 learning	• only implicit L2 learning	
• reason for reading/goal: understanding the story, finding the information necessary to solve the task	• reason for reading/goal: pleasure and understanding	2.2.1 2.2.3
• students need to apply and learn reading skills and strategies by filling in the tasks	• students might apply skills and strategies to generate an understanding of the text, but this is neither guided nor enforced by additional tasks	2.2.1
• less reading = less reading practice because of time needed for exercises	• maximum reading = maximum reading practice (= important for developing automatic and rapid lower level processing skills)	2.2.2

[9] Implicit L2 learning "involves learning processing skills and language knowledge without being aware of attending to specific information that is learned, and it relies on extensive amounts of input" (Grabe, 2009, p. 60).

[10] "Explicit learning, unlike implicit learning, involves conscious attention to, and awareness of, the specific skills or language knowledge that a reader is focusing on." (Grabe, 2009, p. 61)

• students receive feedback and are challenged if their text model understanding of the text is correct • reading, but also writing is skilled • additional, feedbacked tasks give opportunity to show an reflect on what has been learned → stronger fit to performance-approach goal orientation	• students can develop a wrong text model of the text without realizing this; no correction • only reading is skilled • reading being its own reward is emphasized and students do not have to fill in tasks → stronger fit to learning goal orientation or performance-avoid goal orientation	2.2.4.5 2.3.4.4 2.3.5.1
• learning goal orientation positively impacts improvement through reading interventions to a stronger focus on skill-development and a deeper engagement in the learning process		2.3.5.1
• self-concept is more stable with age → can still be influenced and developed to a positive self-concept at this early age → both interventions provide the possibility to make individual and criterial comparisons, instead of strong social comparisons, to develop a (positive) English self-concept		2.3.5.2

Moreover, section 2.3.4.6 presented general and reading models that focus on motivational aspects influencing the reading process and section 2.3.5 focused on how goal orientation and self-concept can influence reading (achievement) and vice versa. The following diagram (figure 2.11) depicts the variables assessed in this study and their possible relationship on the basis of the models and theses discussed.

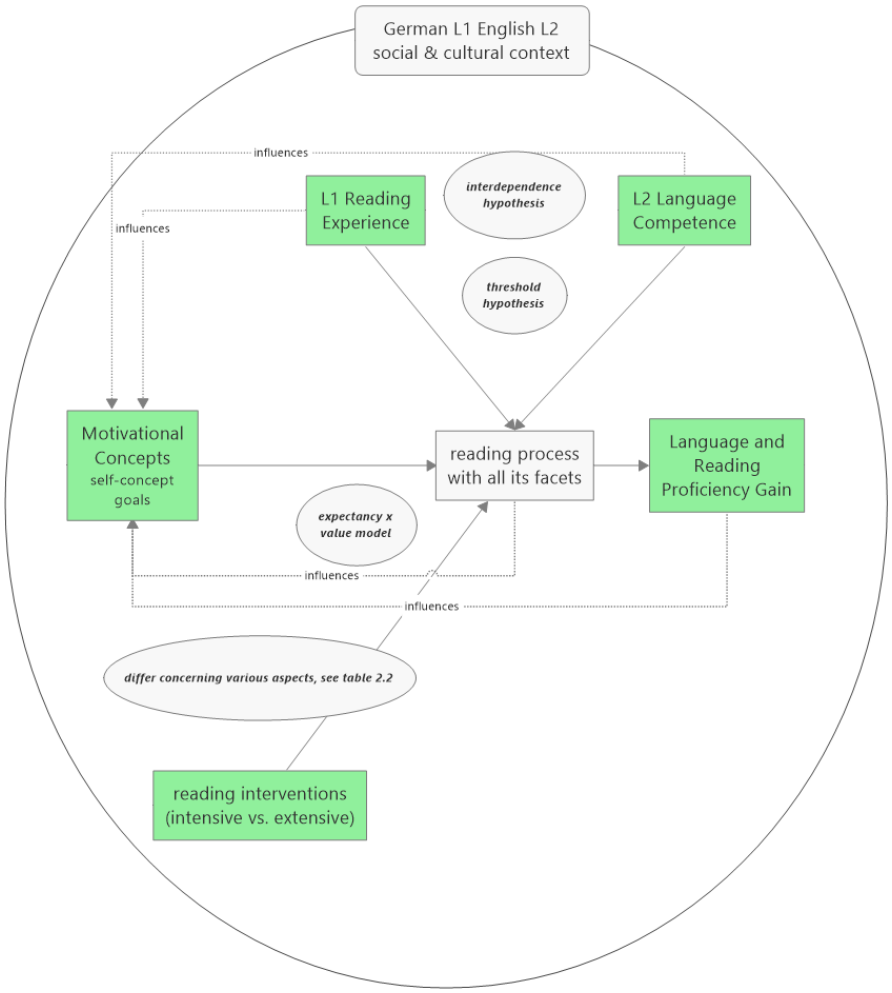


Figure 2.11 Impact of Factors in this Study

The L2 reading process with all its facets is in the center of this impact model. Goals and self-concept, L1 reading and L2 language experience and the reading interventions are expected to influence the reading process – on the one hand directly and on the other hand as an interaction (expectancy x value theory). If and what impact L1 reading experience and/or L2 language competence have on L2 reading processes is addressed by the language Threshold and Interdependence Hypothesis. In total, a successful L2 reading practice then leads to a higher reading and language competence. Moreover, self-concept (and maybe also goals) does not only influence language competence, but could also be influenced by growing language competence due to successful reading interventions (reciprocal effects model). The entire model is set in a social and cultural context that influences the different variables, in this case in Germany with English as the L2 language to be learned

In this study I am going to analyze the direct effect of different reading interventions (intensive vs. extensive), different goal orientations and self-concepts, L2 language competence and L1 reading experience on L2 language and reading competence (green boxes in the model; ANCOVA and MANCOVA analyses). Additionally, I am also going to analyze the relationship of the psychological factors and competence gain (correlations) and direct effects of the reading interventions on self-concept and goal orientation score gain (MANCOVAs). Moreover, the model suggests possible interactions between these psychological and cognitive factors and the reading intervention type (moderation analyses). Finally, I will analyze to what extent significant factors impact language and reading proficiency gain (hierarchical regression analysis with significant factors of previous analyses).

On the basis of the differences between intensive and extensive reading and various motivational and cognitive factors influencing the reading process discussed up to now, I repeat my hypotheses that have been developed in the previous sections here:

1) Both, L1 reading experience (measured as a self-report of L1 books read per week) and L2 language experience (measured as scores in language tests at T1) will be significantly associated with the competence gain for students in the English language tests and in grade improvement.
    a. For the students in the non-English intervention groups, I expect L1 reading experience and L2 language proficiency both to have a positive impact on the competence gain.
    b. For students that receive additional training, especially for the reading interventions, I do not expect these effects to be as high or significant, or even negative.
2) Independently of the intensive or extensive reading intervention, the students' competence gains in the pre- and posttests is positively related to the number of years of English they have had at primary school.

3) Students participating in the extensive reading intervention will perform better in the fictional and non-fictional reading comprehension test than students in the intensive reading intervention, who will perform better in the cloze test and the preposition test.
4) In the extensive reading intervention students will prefer reading books where they can draw on background knowledge to understand them, whereas this will not show as strongly for intensive reading.
5) Goal orientation moderates the effect of the reading interventions: Students who have a stronger learning goal orientation profit more from the extensive reading intervention (higher competence gain) than students with a higher performance-approach goal orientation, who profit more from the intensive reading intervention (higher competence gain). These effects are due to a better fit of the task to the student's motives.
6) Learning goal orientation has an overall positive effect on students' competence gains in the different language tests. This effect is independent of the intervention group type.
7) English (reading) interventions will have a positive impact on the students' English self-concept (self-concept gain), whereas non-English interventions will have a negative impact on the students' English self-concept.
8) Students' self-concept at the beginning of the interventions (T1) will correlate significantly positively with their test score gains and significantly negatively with their average grade difference (meaning a positive relationship of self-concept and grade improvement due to lower grades being better than higher grades). This effect is independent of the experimental group the students are in.

This chapter gave an overview of factors involved in intensive and extensive reading against the backcloth of different research perspectives and theoretical approaches. As presented, previous research on L2 extensive reading is sparce and has concentrated on Asian and, therefore, mostly logographic or syllabic L1 and English L2 constellations and/or developing countries. Moreover, most research is focused on university students (see section 2.3.3). The study of this dissertation will contribute to the research field of extensive and intensive reading in less explored populations. In the following I will present the design and results of the study.

3 Project Description

Literacy is the ticket of entry into our society,
it is the currency by which social and economic positions are waged,
and it is the central purpose of early schooling.
(Bialystok, 2001, p. 152)

3.1 Participants and Missing Data

In total 364 5th and 6th grade students from a German Grammar School, participated in this study. The students were between 9 and 12 years old (*M=10.49; SD=0.55*), 43.1% of the students are male and 52.2% female.[11] 62.2% of the students reported that they started learning English in first grade, 24.4% in second grade, 8.7% in third grade, 3.2% in fourth grade and 1.5% reported that they did not have any English lessons at their primary school. Since the content and depth of English lessons at primary schools is very different, the students were additionally asked, if they had already read words or sentences in those lessons, which was affirmed by 80.3% of the students. 87.1% of the students reported to engage in L1 reading during their free time. This number is in line with previous research of the KIM-Studies (Medienpädagogischer Forschungsverbund Südwest, 2018) were 17% of the students reported to never read. Therefore, in this aspect, the students in this population seem to be quite representative. Table 3.1 gives an overview of the distribution of the participants across the different intervention groups.[12]

Not all cases of the subsamples in Table 3.1 add up to N=364 due to missing data. As a matter of fact, only 244 (67%) of the data sets were complete, but since not every scale is needed for every analysis, the complete cases for the analyses in this study are higher. For the items on language proficiency, 314 (86%) data sets were complete. For the psychologic scales, 284 (78%) data sets and for grades 360 (99%) data sets were complete. Therefore, the large amount of missing data is not due to a high drop-out rate or missing tests, but rather due to single missing items in some scales. Wherever possible, the missing data of the psychological scales was aggregated according to the instructions in the manuals, increasing the ratio of complete data sets.

The composition of the intervention groups was randomized, in consequence the distribution of gender, migration background and pre-knowledge is about equal across all intervention groups, as table 3.1 shows.

[11] Some students did not report their gender or drew a box labeled "other" and checked that, therefore the percentages do not add up to 100%.

[12] In this paper the term "intervention groups" is used as a hyperonym for all groups of this study. These divide into two experimental groups "extensive reading intervention group" and "intensive reading intervention group" and two control groups – "non-English intervention group" and "other English intervention group". For more details see section 3.2.1.

Table 3.1: Participants divided by intervention groups

			n (%)	*n (%)*
non-English intervention (control group)	English in primary school*	no	23 (21.9)	111 (30.5)
		yes	82 (78.1)	
	gender	boy	50 (47.2)	
		girl	56 (52.8)	
	migration background**	no	81 (77.1)	
		yes	24 (22.9)	
other English intervention (control group)	English in primary school*	no	21 (16.8)	131 (36.0)
		yes	104 (83.2)	
	gender	boy	54 (43.5)	
		girl	70 (56.5)	
	migration background**	no	89 (71.2)	
		yes	36 (28.8)	
extensive reading intervention (experimental group)	English in primary school*	no	6 (15.8)	40 (11.0)
		yes	32 (84.2)	
	gender	boy	18 (46.2)	
		girl	21 (53.8)	
	migration background**	no	32 (84.2)	
		yes	6 (15.8)	
intensive reading intervention grade 5 (experimental group)	English in primary school*	no	9 (22.0)	42 (11.5)
		yes	32 (78.0)	
	gender	boy	19 (46.3)	
		girl	22 (53.7)	
	migration background**	no	27 (67.5)	
		yes	13 (32.5)	
intensive reading intervention grade 6 (experimental group)	English in primary school*	no	9 (24.3)	40 (11.0)
		yes	28 (75.7)	
	gender	boy	16 (43.2)	
		girl	21 (56.8)	
	migration background**	no	26 (70.3)	
		yes	11 (29.7)	

*Hast du in der Grundschule schon Wörter oder Sätze auf Englisch gelesen?
**Sprecht ihr in der Familie noch eine andere Sprache außer Deutsch?

3.2 Design

3.2.1 Intervention Groups and Control Groups

The interventions were integrated into an individual assistance program that generally takes place in the second half of grade 5 and the first half of grade 6 at the grammar school where this study was conducted. The assistance program is integrated into the school curriculum and time table. The program consists of a variety of courses that help students to either catch up or give them additional input in their main courses, or to engage in a topic not represented in the core curriculum. The students participate in three courses in grade 5 and three in grade 6, each course is taught 90 Minutes per week for six weeks, adding up to a total of 540 minutes for each course. The intensive and extensive reading intervention were two of these courses. Other courses offered are: Advanced English, English Foundation, Advanced Math, Math Foundation, Essay Writing (German), Spelling (German), Music and Dance, Robotics, Juggling and Badminton. The group sizes vary between 10 to 20 students and are all taught by teachers with the corresponding subjects; I taught the reading interventions. The English reading courses were designed and integrated into the program for the purpose of this study.

The teacher who organizes the individual assistance programs assigns the students to the different classes at the beginning of each school semester. As a guidance for her decision, she asks the students and their teachers to name preferences for certain courses. To ensure an unbiased and mixed student sample for each reading intervention group, the students were not able to choose the reading intervention, they were randomly assigned to this course. It was only decisive, if the student and parents had approved to the participation in the study, which was the case for almost all students. Furthermore, the teacher assigning the students was advised to assign students of different proficiency and from different classes to randomize possible effects of confounding variables, like interest, English proficiency, motivation, effects of the main English class or the English teacher.

The various intervention groups of the individual assistance program can be divided into three resp. four main groups; the type of group is the main independent variable of this study:

1. Experimental Group: students that received the reading intervention in one of the trimesters and had non-English classes in the other trimesters (see figure 3.1, student 1)
2. Control Group 1 (other English intervention): students that took part in either the Advanced English or the English Foundation Course in one of the trimesters and non-English courses in the other trimesters (see figure 3.1, student 2)
3. Control Group 2 (non-English intervention): students that neither took part in the reading intervention, nor another English class throughout the three trimesters (see figure 3.1, student 3)
4. Mixed Group: students that were part of the intervention group and another English class (see figure 3.1, student 4).

	Trimester 1	Trimester 2	Trimester 3
Student 1	advanced maths	reading project	spelling
Student 2	English foundation	maths foundation	essay writing
Student 3	juggling	essay writing	music and dance
Student 4	maths foundation	advanced English	reading project

Figure 3.1 Example of possible combinations for students.

3.2.2 Interventions

3.2.2.1 Extensive Reading

Following the different principles and suggestions for running an extensive reading program (see chapter 2.2), students of the extensive reading intervention were introduced to the idea of extensive reading in the first lesson: They were advised to read as many English books as they can. They were shown how to check if books were at the right level and assured that they do not have to understand everything. They should freely choose what they wanted to read and were told that they would neither have to pass a test or answer comprehension tasks, nor be rewarded for their reading. The only rule was that they had to read quietly on their own. During the students' reading, the teacher helped students choose their books, talked to students about what they had just read, or sat in the class and read quietly one of the library books herself.

The students borrowed and returned every book that they choose with a library software, even if they only read it in class, making it possible to keep track of their amount of reading and their reading habits, not only for the teacher, but also for the students themselves. Dictionaries were not provided to help with vocabulary, but the students could ask questions, if there was something they did not understand. No exercises were provided and there was no guided reading. With a length of 90 minutes and assuming a reading speed of about 100 words per minute, the students had plenty of time to reach the recommended minimum amount of 5 500 words per week during this course.

3.2.2.2 Intensive Reading

To keep as many variables as constant as possible between the two interventions, students of the intensive reading intervention were also able to choose their books freely from the same range of books, but had different tasks to work on in a portfolio for each of their chosen books. The cover page of the portfolio was a bookshelf, motivating students to color in books they have read (see figure 3.2). It should create a feeling of pride for every book they were able to collect in there. The next section of

the portfolio is for new vocabulary (see figure 3.3). The students could either fill in the vocabulary pages on their own as they encountered new words, or they had tasks on their worksheets where they had to write down new vocabulary for this section. The students had to look up the German translation on their own and also make new sentences using the new word.

The next section in the portfolio are tasks that the students needed to complete after finishing a book. My aim when designing the tasks was to create worksheets where the students need to closely engage and work with the reading material, but are still free in choice concerning their reading material, and to some extent also the tasks they want to solve and the speed in which they have to finish. We[13] created matching task sheets for every library book that consisted of three comprehension and three language tasks (see figures 3.4 and 3.5). The tasks range from simple comprehension questions to creative tasks like writing an own story. The students were able to choose which task(s) they wanted to do, as long as they completed at least one language and one comprehension task. Therefore, they had a freedom of choice which book they wanted to read and which tasks they wanted to engage in. After completing a book, students needed to fill in an evaluation sheet on the book and the tasks, giving feedback on why they chose this book and how they perceived the tasks (see figure 3.6).

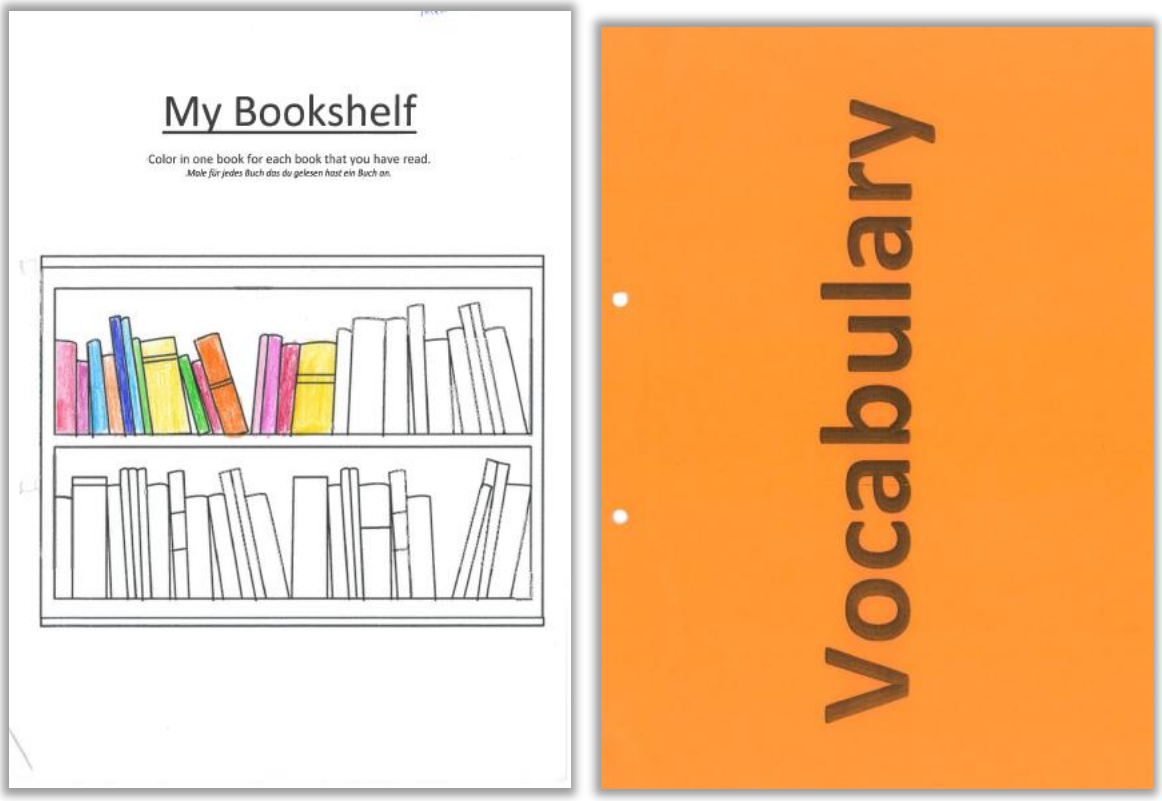


Figure 3.2 Bookshelf from Arda, Grade 6

[13] I would like to thank my student assistant Phylicia Weitlauff at this point, who was an extraordinary help in developing all the different work sheets for every single book of my 386 library books and coding thousands of tests.

English Word:	German Word:
(to pushing)	drücken

Original Sentence: Children race across the lawn, pushing hard-boiled eggs* with spoons!

Own Sentence: I pushing the button.

English Word:	German Word:
often	oft

Original Sentence: And President Kennedy's son John Jr. often played

Own Sentence: I learn often for school.

* under his dad's desk

English Word:	German Word:
daughter	Tochter

Original Sentence: President Bonnerly's daughter Penelope had a puppy

Own Sentence: The daughter from my grandma is my mother

* named Rascuito.

English Word:	German Word:
visit	besuchen

Original Sentence: Pope John Paul II was the first Pope to visit the

Own Sentence: On weekends I visit my grandma.

* famous house.

English Word:	German Word:
builder	Bauunternehmer

Original Sentence: In July 1792, a builder named James Hoban won the

Own Sentence: The builder builds houses.

English Word:	German Word:
honor	auszeichnen

Original Sentence: It was named in his honor.

Own Sentence: The President honors my mum.

English Word:	German Word:
called	anrufen/rufen

Original Sentence: For years people simply called it "the white house".

Own Sentence: My mum called my name often

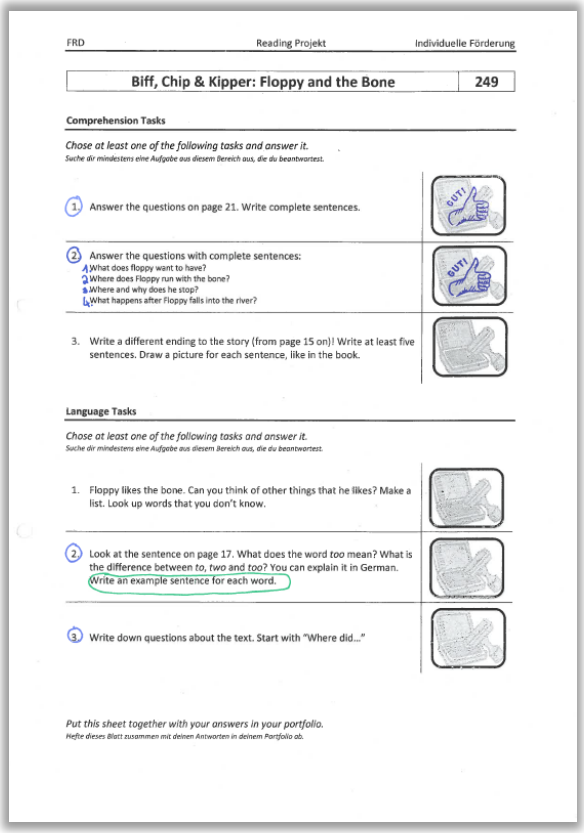
English Word:	German Word:
turkey	Truthahn

Original Sentence: Each November, the President saves one turkey

Own Sentence: My family likes turkey for dinner

* context. S from becoming Thanksgiving dinner!

Figure 3.3 Vocabulary Section from Inka, Grade 5



FRD Reading Projekt Individuelle Förderung

Biff, Chip & Kipper: Floppy and the Bone 249

Comprehension Tasks

Chose at least one of the following tasks and answer it.
Suche dir mindestens eine Aufgabe aus diesem Bereich aus, die du beantwortest.

1. Answer the questions on page 21. Write complete sentences.

2. Answer the questions with complete sentences:
   1. What does floppy want to have?
   2. Where does Floppy run with the bone?
   3. Where and why does he stop?
   4. What happens after Floppy falls into the river?

3. Write a different ending to the story (from page 15 on)! Write at least five sentences. Draw a picture for each sentence, like in the book.

Language Tasks

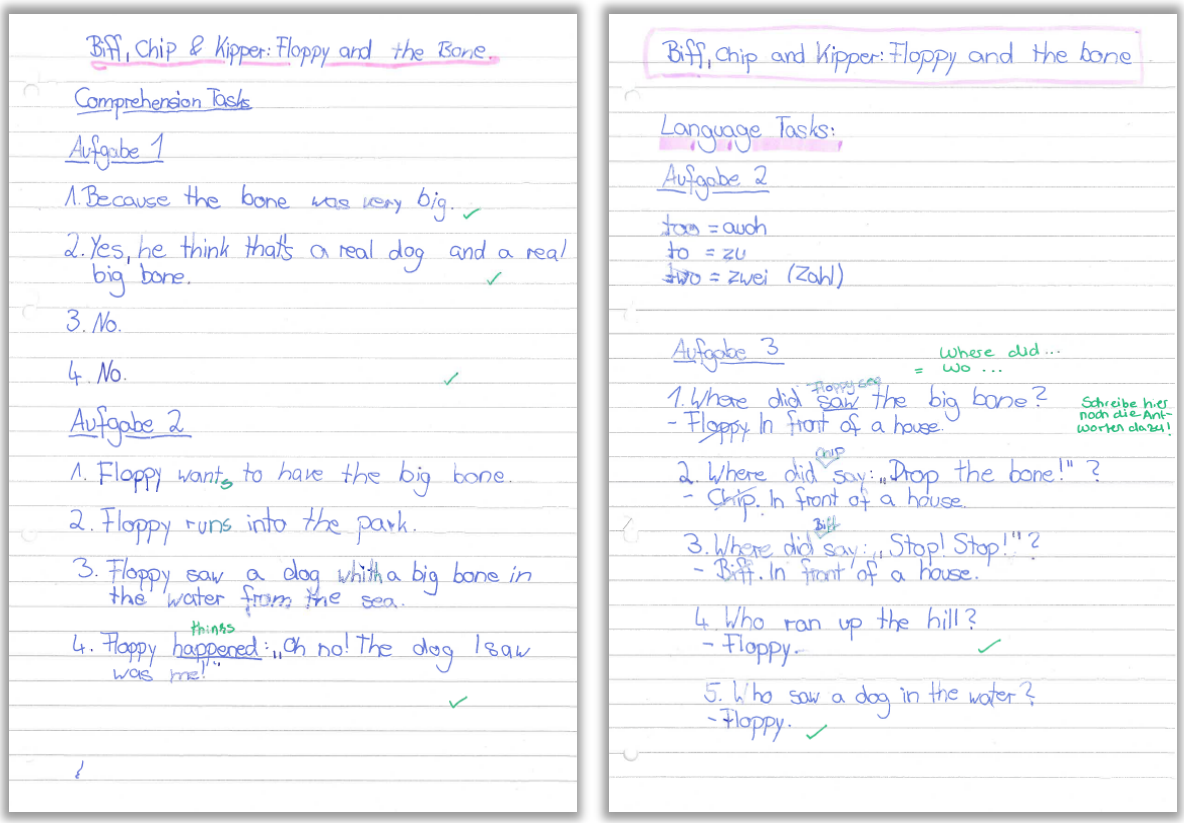
Chose at least one of the following tasks and answer it.
Suche dir mindestens eine Aufgabe aus diesem Bereich aus, die du beantwortest.

1. Floppy likes the bone. Can you think of other things that he likes? Make a list. Look up words that you don't know.

2. Look at the sentence on page 17. What does the word *too* mean? What is the difference between *to, two* and *too*? You can explain it in German. Write an example sentence for each word.

3. Write down questions about the text. Start with "Where did..."

Put this sheet together with your answers in your portfolio.
Hefte dieses Blatt zusammen mit deinen Antworten in deinem Portfolio ab.



Biff, Chip & Kipper: Floppy and the Bone.

Comprehension Tasks

Aufgabe 1

1. Because the bone was very big. ✓
2. Yes, he think that's a real dog and a real big bone. ✓
3. No.
4. No. ✓

Aufgabe 2

1. Floppy wants to have the big bone.
2. Floppy runs into the park.
3. Floppy saw a dog with a big bone in the water from the sea.
4. Floppy thinks happened: „Oh no! The dog I saw was me!" ✓

Biff, Chip and Kipper: Floppy and the bone

Language Tasks:

Aufgabe 2

too = auch
to = zu
two = zwei (Zahl)

Aufgabe 3

Where did ... = wo ...

1. Where did Floppy saw the big bone?
- Floppy. In front of a house.
2. Where did Chip say: „Drop the bone!" ?
- Chip. In front of a house.
3. Where did Biff say: „Stop! Stop!" ?
- Biff. In front of a house.
4. Who ran up the hill?
- Floppy. ✓
5. Who saw a dog in the water?
- Floppy. ✓

Schreibe hier noch die Antworten dazu!

Figure 3.4 Tasks from Emma, Grade 6

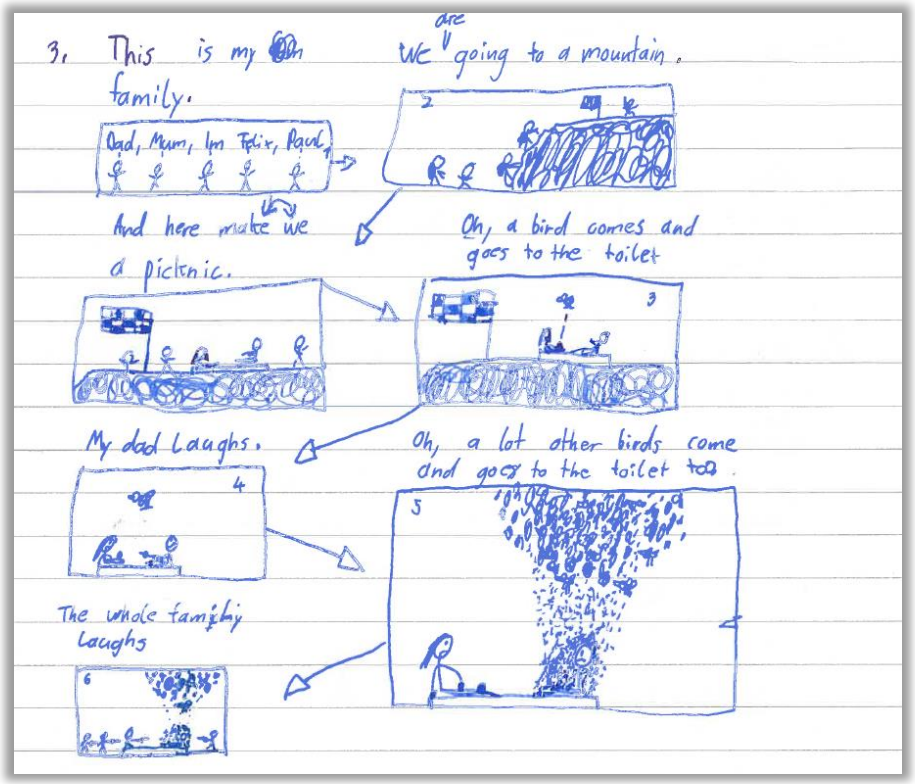


Figure 3.5 Tasks from Florian, Grade 6

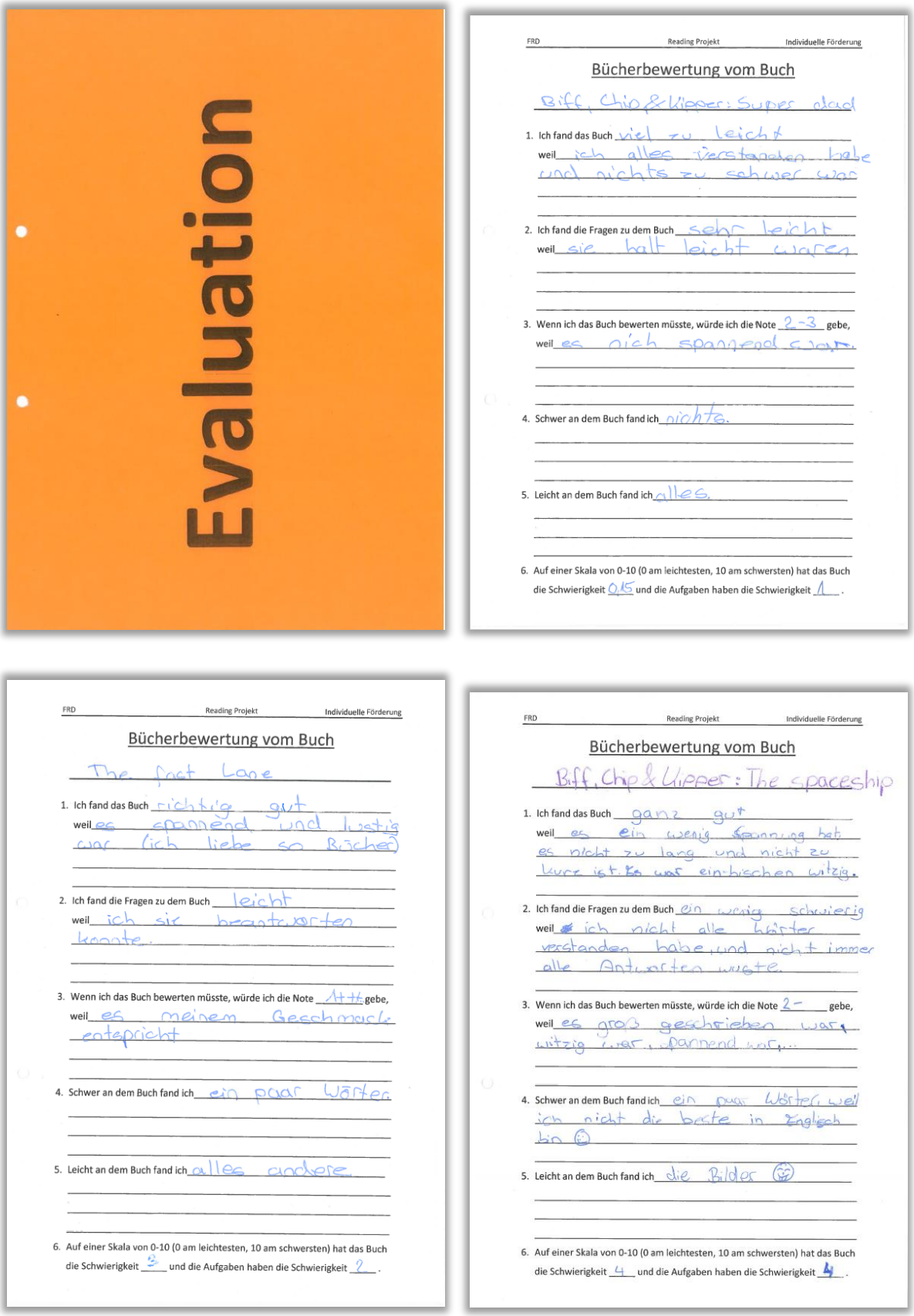


Figure 3.6 Evaluation from Lili, Grade 6

A teacher trainee additionally assisted the students during the tasks[14]. After every lesson, I collected, corrected and feedbacked the portfolios of the students (see figure 3.7). At the beginning of each lesson, the students spent at least 15 minutes going through the feedback and making corrections, before they were allowed to go on reading the next books or completing the next task.

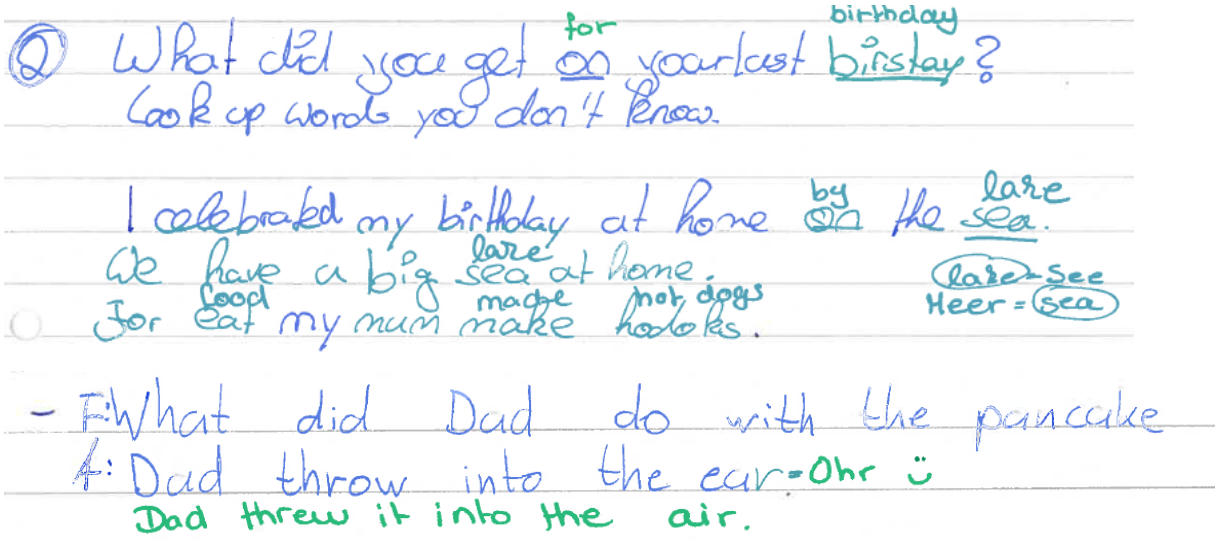


Figure 3.7 Corrections for Fenna, Grade 5

To achieve more acceptance for the corrections and the feedback, the corrections were made in green instead of red and fully correct or corrected answers were rewarded with a stamp. If the students managed to collect a certain number of stamps, the group was granted a pizza party in the last lesson. Just as in the extensive reading intervention, their reading amount and preference were tracked with the library program (see section 3.2.3).

3.2.2.3 Other Interventions

The other English courses were not always taught by the same teachers; therefore, the content of the lessons vary. The content of the English interventions for the advanced groups was creative writing, grammar revision at a higher-level and understanding their favorite English songs. The English foundation group focused primarily on grammar topics using additional material provided by the school book publishers and explanations on the board.

In the German courses, pupils also worked with material provided by school book publishers, mostly binders with diagnostic tests and then additional exercise material. For more spelling practice

[14] I thank Luis Flaig, my former student who was doing a voluntary social year at our school during the time of the study. Luis was a great help during the intervention sessions and also jumped in running the library in the breaks when I was not there.

the students also wrote small dictations, for essay practice the students wrote interesting stories on topics like friendship or ghosts. The teacher talked to the students about what the structure of a story should be like and what stylistic devices and language they should use. The robotic course worked with Calliope mini, a small computer with only one circuit board used in education. With the help of a digital learning course, the students engaged in the first steps of programming. Unfortunately, I did not receive feedback on the content of the other courses.

3.2.3 English Books and Library

The basic material for the reading interventions was a collection of 386 English children's books and easy readers (see section 9 for a full bibliography). I chose the books with the aim to reflect a wide range of text types and interests, therefore the book collection does not only include easy readers from German school book publishers for L2 learners, but also a wide range of books originally aimed at L1 beginning readers. Whereas the easy readers are longer stories with more text but very controlled vocabulary and grammar, the British and American books for L1 children have less text and many supporting pictures, but are usually more complex concerning vocabulary and grammar. Adding numerous original British and American books to the library also made it possible to widen the range of topics from the typical easy reader stories about family and friendship or small adventures to non-fictional books about animals, cultures, scientific topics or biographies. I additionally wanted some books that were much easier than the easy readers for L2 beginners, so that also weak students with limited vocabulary could engage in the reading process and gain confidence in reading by experiencing that they were able to read an entire book, even if it is very short and easy. In my search for suitable books, I turned to baby board books and very short easy readers for L1 first readers. I chose the British and American children's books by viewing the books at the Children's English Library Stuttgart and recommendations on Amazon.

The books were kept in a cabinet in the school foyer (see figure 3.8). All students of this study, independently of the intervention group they were in, were able to borrow books during the break, were allowed to take them home and were able to return them either during the breaks or into a letter box at the cabinet. I kept track of the lending behavior with the library software PS-Biblio from Paidsoft: Each student was equipped with a library card with a barcode and each book with a barcode as well. A laptop with the program and a USB barcode scanner were kept in the cabinet. The books were displayed in thematic boxes, because students pulling out and properly laying back books kept on a shelf did not work out in the first week and it was easier to move the boxes up to the classroom each week for the reading lessons.

Figure 3.8 Student's deciding to borrow a book and English library cabinet

3.3 Pre- and Posttests

3.3.1 Procedure

Every six weeks, directly before and after each new intervention period, a pre-/posttest was conducted with all students participating in the study (see table 3.2). All students wrote the test at the same time, during a class where the students were in their normal classes and supervised by a teacher who had lesson at that time. With a standardized information sheet, the teachers were informed to keep the students focused on the test, prevent copying and to not answer any questions or help. The teachers did not know who participated in which intervention group, nor were they informed in detail what the survey or the study was about. Within the frame of the 90-minute lesson, there was no time limit on the test and the students were given enough time to finish. The few students who did not have parental permission to participate in the study were kept busy with math tasks. As a reward, the students received a small sweet treat after finishing the test.

Table 3.2 gives an overview of the intervention groups for each six-week period. Tests were administered before and after these classes for all students. There were two tests after 2017/2 (3)[15] and before 2017/1 (1) due to the summer holidays. The same holds for 2017/1 (3) and 2018/2 (1) due to the Christmas holidays. For class 5, the additional classes started in the second half of the school year, after the Christmas holidays. The reading interventions were taught as extensive reading. After the summer holidays these students entered 6th grade and had additional classes in the first half of the school year, this time the reading intervention was taught as an intensive reading intervention. It was made sure that students who had already participated in an extensive reading class did not participate again in the intensive reading class. After Christmas the new 5th graders started again with additional classes, this time taught as intensive reading classes.

[15] meaning year 2017, semester two, intervention period three

Table 3.2 Overview of the different intervention groups and tests conducted

Year/Semester (Trimester)	Class	Number of Groups (Number of Students)			
		Extensive	Intensive	Other Engl.	Other Interv.
Test 1 2017/2 (1)	5	1 (18)	--	2 (21)	8 (84)
Test 2 2017/2 (2)	5	1 (15)	--	2 (15)	8 (93)
Test 3 2017/2 (3)	5	1 (13)	--	2 (17)	8 (93)
Test 4 and 5 2017/1 (1)	6	--	1 (13)	1 (12)	8 (99)
Test 6 2017/1 (2)	6	--	1 (13)	1 (12)	8 (99)
Test 7 2017/1 (3)	6	--	1 (15)	1 (10)	8 (99)
Test 8 and 9 2018/2 (1)	5	--	1 (14)	2 (25)	7 (77)
Test 10 2018/2 (2)	5	--	1 (14)	2 (20)	7 (82)
Test 11 2018/2 (3)	5	--	1 (14)	1 (11)	8 (91)
Test 12					

To ensure data anonymity after the survey, but allow data synchronization, the tests were equipped with a participant number and the student's name on a piece of tape. Before handing back the test, the students ripped off their name, leaving only their anonymous participant number on the test. Thus, the digitalized data and the stored questionnaires do not contain the student's names.

3.3.2 Test Structure

Each test consisted of a cloze test, a fictional and a non-fictional reading comprehension and a preposition exercise. I designed four parallel versions of these language proficiency tests, because it would have been quite likely to have large retest effects if the students had filled in the same test every six weeks. Moreover, some students talk about tests and exchange their answers afterwards and some do not, which might have led to strong biases of the results of the following tests if these had been identical. The psychological scales SELLMO and SESSKO were only part of the first and last test of each intervention semester, because filling in the scales each time would lead to redundancy and probably inaccurate results due to a loss of motivation for answering the questions conscientiously. The rest of the information on demographic data, media consumption and reading preferences were integrated into tests that did not have a lot of items already. The following table (3.3) gives an overview of the content of each test. The content and construction of the different test parts will be explained in the following sections.

Table 3.3 Combination of test parts

Tests	1	2	3	4	5	6	7	8	9	10	11	12
	Cohort 1 Grade 5				Cohort 1 Grade 6				Cohort 2 Grade 5			
Cloze Test	X	X	X	X	X	X	X	X	X	X	X	X
Non-Fictional Reading Comprehension	X	X	X	X	X	X	X	X	X	X	X	X
Fictional Reading Comprehension	X	X	X	X	X	X	X	X	X	X	X	X
Preposition Exercise	X	X	X	X	X	X	X	X	X	X	X	X
SELLMO	X			X	X			X	X			X
SESSKO	X			X	X			X	X			X
Demographic Data			X							X		
Reading Behavior			X							X		
Media Consumption							X				X	
Reading Preferences				X							X	

3.3.3 Test Design

When designing the tests for the students, my main aim was to compose tests that

1) test the students' language and reading proficiency to be able to assess their competence gain from test to test
2) assess the students' beliefs and motivation concerning the subject English
3) collect information on further variables that could play a role when assessing the students' competence gain in English

The discussion of what reading is and different reading processes in the previous chapter (see section 2.3) already showed how complex reading is. In how far it is possible to even measure single skills or strategies is questionable (Khalifa & Weir, 2009). Khalifa and Weir (2009) therefore suggest that one might

> be better served by identifying which types of reading are most appropriate to different levels of proficiency and attempting to ensure by the text selected, the way items are constructed, and other salient performance conditions set […] that the cognitive processing demand to complete such tasks are commensurate with the skilled reading process (p. 40)

This is especially relevant when looking at the validity of cloze tests, but of course also for the reading tests designed for this study. I will elaborate on this issue in the following sections.

I decided to use a cloze test to measure a general language proficiency (see section 3.3.3.1). To measure reading comprehension, I chose to include a fictional and a non-fictional reading comprehension test (see section 3.3.3.2 and 3.3.3.3), because students are confronted with both text types in the reading interventions. Additionally, a preposition test was added to the test set (see section 3.3.3.4). Time wise, it would have been impossible to integrate several tests to different grammatical skills, therefore I decided to test prepositions, because the appropriate use of prepositions requires syntactic as well as semantic knowledge. Prepositions are a challenging category. Sometimes they are classified as a lexical sometimes as a functional category. Some are semantically transparent (e.g. *book on/under the table)* others are lexically determined (e.g. *decide on something*). What is more, even if there are lexical equivalents (*on/auf, in/in*) it is not necessarily the equivalent which is used in the other language (e.g. *auf dem Tisch, on the table, im Internet, on the internet*). Therefore, extensive input in English is necessary to develop a notion of what preposition should be used when. I assume this mixture of tests to deliver an accurate reflection of the student's abilities in areas relevant for analyzing the benefit of the implemented reading programs.

Language testing and language proficiency cannot be evaluated without taking characteristics of the test-takers into account (Khalifa & Weir, 2009). Therefore, apart from the standard demographic data, I additionally collected information about the student's self-concept and goal orientation in English (see section 3.3.3.5 und 3.3.3.6), see also section 2.3.5 for further background details.

3.3.3.1 Cloze Test

The cloze test procedure, originally developed for judging text difficulty (Taylor, 1953), has developed into a tool for judging foreign language proficiency in the 1970s (Watanabe & Koyama, 2008). The big benefit of the cloze test procedure is seen in its easy construction (Aitken, 1977), but its validity has been questioned in the following years and investigating what a cloze test really measures has been subject to extensive research. Alderson (1979) suggests that the cloze test procedure measures rather low-level processing skills (see section 2.2), because correlation results with grammar and vocabulary tests (r = .65 - .70) proved to be higher than with reading comprehension tests (r = .58 to .63). Gellert and Elbro (2012) challenge this view by stating "that there are gaps and there are gaps" (p. 17), meaning that some gaps might only require parsing of certain semantic and syntactic structures at a sentence level, whereas other gaps might require inferences beyond sentence level and therefore a more global understanding. This procedure would call for variable spaced gaps that delete words demanding a certain amount of comprehension.

All in all, the reliability and validity of cloze tests has proven to be very heterogeneous in previous research (Watanabe & Koyama, 2008). Oller (1973) reports relatively high correlations between .83 and .89 with the second language placement exams UCLA, ESLPE and the standardized English proficiency test TOEFL. What is interesting about these results is that the measures for listening comprehension showed the highest correlations with the cloze test results. Balik (1980) explains this phenomenon with the similarity between the cognitive challenges of having to fill in natural gaps in communicative settings to the gaps in cloze tests. He argues that when listening to a test or taking part in a conversation, words are lost due to auditive misunderstandings or a listener might encounter unknown words. These gaps have to be filled in order to understand what is being said and, even though one form is oral and the other written, both tasks require a general language proficiency. I do not totally agree with this point of view, because gaps can be left empty in a conversation and it is still possible to make sense of what is being said, whereas in cloze tests participants are forced to fill these gaps to complete the test. Taking the research about cloze tests into account, I would argue that cloze tests measure a mixture of language competences similar to what is necessary in conversations. That is why I have decided to include a cloze test for measuring an overall language proficiency, but to supplement it with further tests measuring prepositional knowledge and reading comprehension.

It could be argued that even simple texts could still be too difficult for beginning learners in 5th and 6th grade, but cloze test results seem to be rather independent of the text difficulty (Oller, 1973). Alderson (1979) even showed in his research that increasing difficulty of cloze tests also leads to an increasing correlation between cloze test results and other proficiency tests.

Further validity issues like the impact of the deletion pattern and different scoring methods on the test's validity were analyzed in previous research. A meta-analysis of Watanabe and Koyama (2008) showed that rational deletion patterns (variably spaced – only words are omitted that require a certain amount of understanding of the text) had a higher reliability than random deletion patterns (every n-th word), which is in line with Gellert and Elbro (2012). When using a random deletion pattern, tests using the deletion of every seventh word were more reliable to those omitting every tenth or twelfth word. A rational deletion pattern would have been difficult for designing parallel tests, therefore, deleting every seventh word is the deletion pattern that I opted for when designing the cloze tests for this study. This also made it possible to adapt already established cloze tests that used this deletion pattern.

When comparing scoring methods, clozentropy[16] produced results with the highest reliability, but is difficult to assess. Therefore, most studies included in the meta-analysis used either the AC

[16] Watanabe and Koyama (2008, p. 110): "For clozentropy scoring (CE), a corpus of answers is developed and the frequency of each answer on the corpus is calculated. Then, this information is cross-checked by administering the same cloze test to

scoring method (any acceptable answer is counted) or the EX method (only the original deleted word is counted). Even though results do not seem to differ significantly (Oller, 1973), the AC scoring method is perceived to be fairer (Watanabe & Koyama, 2008). The AC method is the scoring method used in this study.

To design parallel cloze tests, in a first step, I chose ten different texts of presumably similar difficulty. They were mainly adapted from school books for 5th and 6th grade that the students did not know, because they were taken from older versions of their book or from another publisher. School book texts were chosen because it is likely that the vocabulary and grammar knowledge, needed to comprehend the texts, is similar and therefore I assumed that the cloze tests have about the same difficulty. In a second step, every 7th word in the texts was omitted, contracted words were counted as one word. Additionally, if the gap was a name (see second gap), which was not possible to guess from context and needed for further comprehension, the next word was omitted. Some texts were shortened, so that each text had exactly 25 gaps, see figure (3.9) for an example.

Da hat wohl jemand ein paar Wörter vergessen…

Unten findest du einen Text, bei dem jedes 7. Wort fehlt. Ergänze das Wort, das deiner Meinung nach am besten in die Lücke passt. In jede Lücke passt nur ein Wort.

Beispiel:
Bill and Ben are cats. Every __day__ after breakfast I open the door __for__ them. Then Bill and Ben go __out__ …

The Shock Team

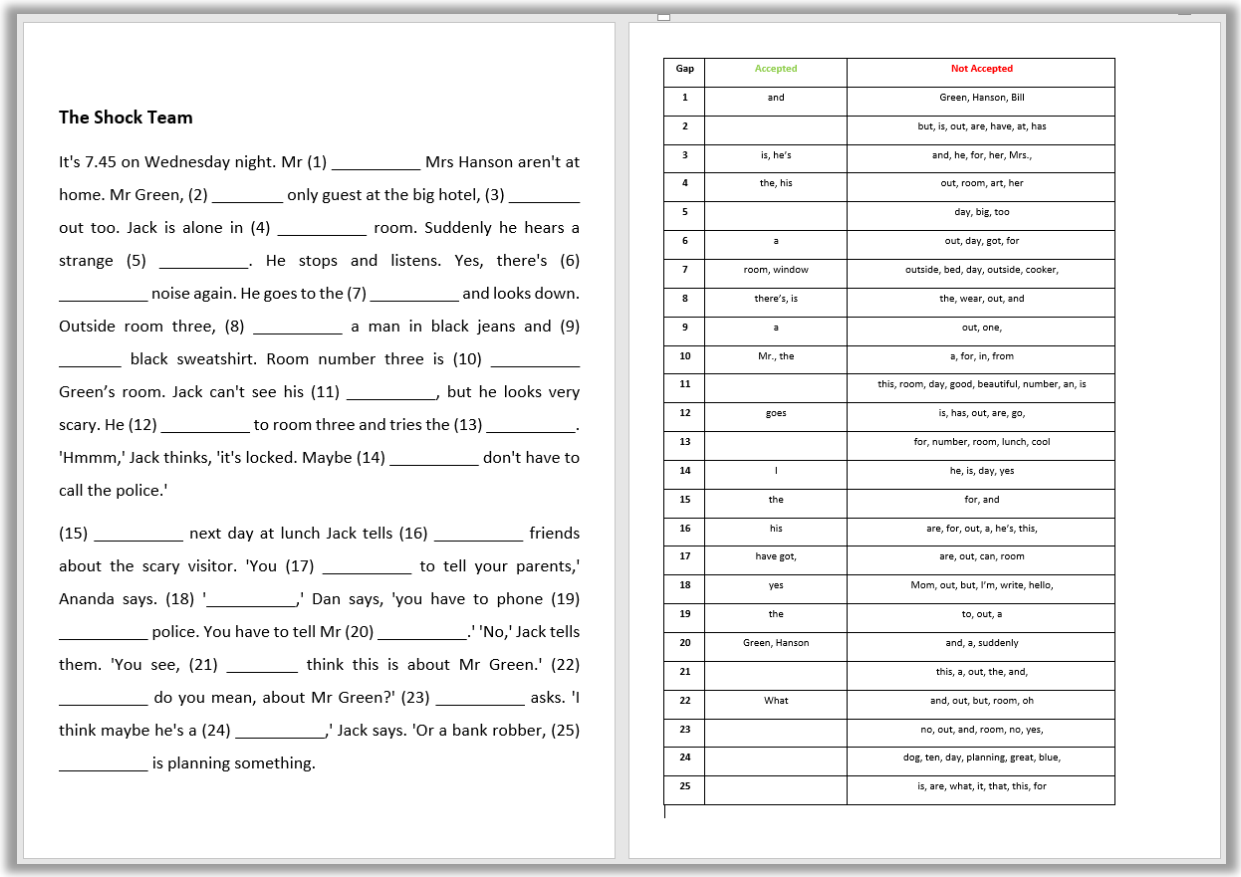
It's 7.45 on Wednesday night. Mr _____ Mrs Hanson aren't at home. Mr Green, _____ only guest at the big hotel, _____ out too. Jack is alone in _____ room. Suddenly he hears a strange _____. He stops and listens. Yes, there's _____ noise again. He goes to the _____ and

Figure 3.9 Cloze Test

In a third step, the tests were piloted with a group of sixth grade students. A gap was considered answered correctly (= 1 point) if the word made grammatical and lexical sense in the context of the

native speakers and compiling those responses. The answers are, finally, weighted according to a logarithm of the frequency of each response." This procedure was developed by D. K. Darnell (1970).

text, even if it was not a match to the original version. The following figure (3.10) gives an overview of all correct and incorrect solutions provided by students for one of the cloze tests. After checking the parallel test reliability in a final step, the four tests with the highest reliability were chosen for the study (see chapter 3.3.4.2 for further detail).



The Shock Team

It's 7.45 on Wednesday night. Mr (1) _____ Mrs Hanson aren't at home. Mr Green, (2) _____ only guest at the big hotel, (3) _____ out too. Jack is alone in (4) _____ room. Suddenly he hears a strange (5) _____. He stops and listens. Yes, there's (6) _____ noise again. He goes to the (7) _____ and looks down. Outside room three, (8) _____ a man in black jeans and (9) _____ black sweatshirt. Room number three is (10) _____ Green's room. Jack can't see his (11) _____, but he looks very scary. He (12) _____ to room three and tries the (13) _____. 'Hmmm,' Jack thinks, 'it's locked. Maybe (14) _____ don't have to call the police.'

(15) _____ next day at lunch Jack tells (16) _____ friends about the scary visitor. 'You (17) _____ to tell your parents,' Ananda says. (18) '_____,' Dan says, 'you have to phone (19) _____ police. You have to tell Mr (20) _____.' 'No,' Jack tells them. 'You see, (21) _____ think this is about Mr Green.' (22) _____ do you mean, about Mr Green?' (23) _____ asks. 'I think maybe he's a (24) _____,' Jack says. 'Or a bank robber, (25) _____ is planning something.

Gap	Accepted	Not Accepted
1	and	Green, Hanson, Bill
2		but, is, out, are, have, at, has
3	is, he's	and, he, for, her, Mrs.,
4	the, his	out, room, art, her
5		day, big, too
6	a	out, day, got, for
7	room, window	outside, bed, day, outside, cooker,
8	there's, is	the, wear, out, and
9	a	out, one,
10	Mr., the	a, for, in, from
11		this, room, day, good, beautiful, number, an, is
12	goes	is, has, out, are, go,
13		for, number, room, lunch, cool
14	I	he, is, day, yes
15	the	for, and
16	his	are, for, out, a, he's, this,
17	have got,	are, out, can, room
18	yes	Mom, out, but, I'm, write, hello,
19	the	to, out, a
20	Green, Hanson	and, a, suddenly
21		this, a, out, the, and,
22	What	and, out, but, room, oh
23		no, out, and, room, no, yes,
24		dog, ten, day, planning, great, blue,
25		is, are, what, it, that, this, for

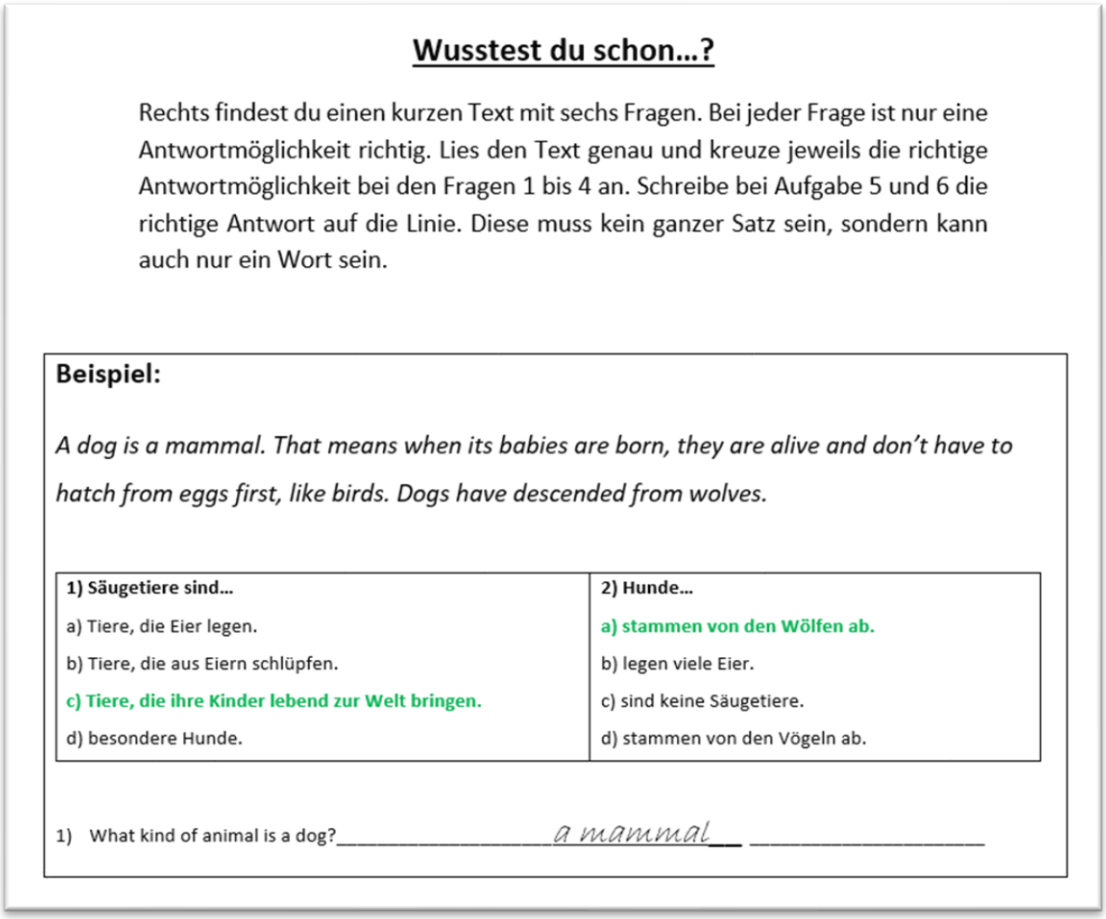
text adapted from Harger and Schwarz (2013) p. 65

Figure 3.10 Cloze Test Solutions

3.3.3.2 Non-Fictional Reading Comprehension

The non-fictional reading comprehensions were adapted from TeeVee (2016). The texts were designed to be about 100 words long with four multiple choice and two open comprehension questions. The comprehension questions were translated from the English original into German, to ensure that the students did not fail to answer a question due to poor comprehension of the question, but rather because they did not understand the text. Additionally, the students really had to understand the content to answer the German questions and could not just look for an English answer in the text. Open questions were added to the test after a first piloting of the translated original test versions failed, these were phrased in English. Furthermore, some words in the text were replaced with less difficult synonyms. The following figure (3.11) shows the introductory example the students

were given in their tests. To check if the different comprehension tests designed are really parallel, they were piloted with students who did not participate in the study (see chapter 3.3.4.2 for further detail)



Wusstest du schon…?

Rechts findest du einen kurzen Text mit sechs Fragen. Bei jeder Frage ist nur eine Antwortmöglichkeit richtig. Lies den Text genau und kreuze jeweils die richtige Antwortmöglichkeit bei den Fragen 1 bis 4 an. Schreibe bei Aufgabe 5 und 6 die richtige Antwort auf die Linie. Diese muss kein ganzer Satz sein, sondern kann auch nur ein Wort sein.

Beispiel:

A dog is a mammal. That means when its babies are born, they are alive and don't have to hatch from eggs first, like birds. Dogs have descended from wolves.

1) Säugetiere sind…	**2) Hunde…**
a) Tiere, die Eier legen.	**a) stammen von den Wölfen ab.**
b) Tiere, die aus Eiern schlüpfen.	b) legen viele Eier.
c) Tiere, die ihre Kinder lebend zur Welt bringen.	c) sind keine Säugetiere.
d) besondere Hunde.	d) stammen von den Vögeln ab.

1) What kind of animal is a dog? a mammal

Figure 3.11 Non-fictional reading comprehension

3.3.3.3 Fictional Reading Comprehension

The fictional reading comprehension tasks were designed by adapting texts of similar difficulty from different reading comprehension worksheets for teachers. The texts were shortened or lengthened to be 150 to 200 words long. Five true/false statements (three true and two false) were invented for each test. The students were additionally asked to correct the false statements. Each test could earn 7 points - five for the correct selection of true or false and one point for each correction of the false answers. The tests were piloted beforehand to ensure that they are parallel (see chapter 3.3.4.2). The following figure (3.12) shows the example the students were given in their tests as an introduction.

Worum geht es in dem Text?

Auf den folgenden Seiten findest du zwei Texte mit jeweils 5 Aussagen. Entscheide, ob die Aussagen richtig (true) oder falsch (false) sind. Falls die Aussage falsch ist korrigiere den Satz und schreibe ihn richtig auf die Linie.

Beispiel:

Hi Beth!
Would you like to go shopping on Saturday? I want to buy some new shoes. Let's meet outside the bank. Then we can go to the shopping mall.
Love
Helen xx

1) Helen wants to go shopping on the weekend..

x true ☐ false: _____ .

2) Helen wants to meet outside the mall.

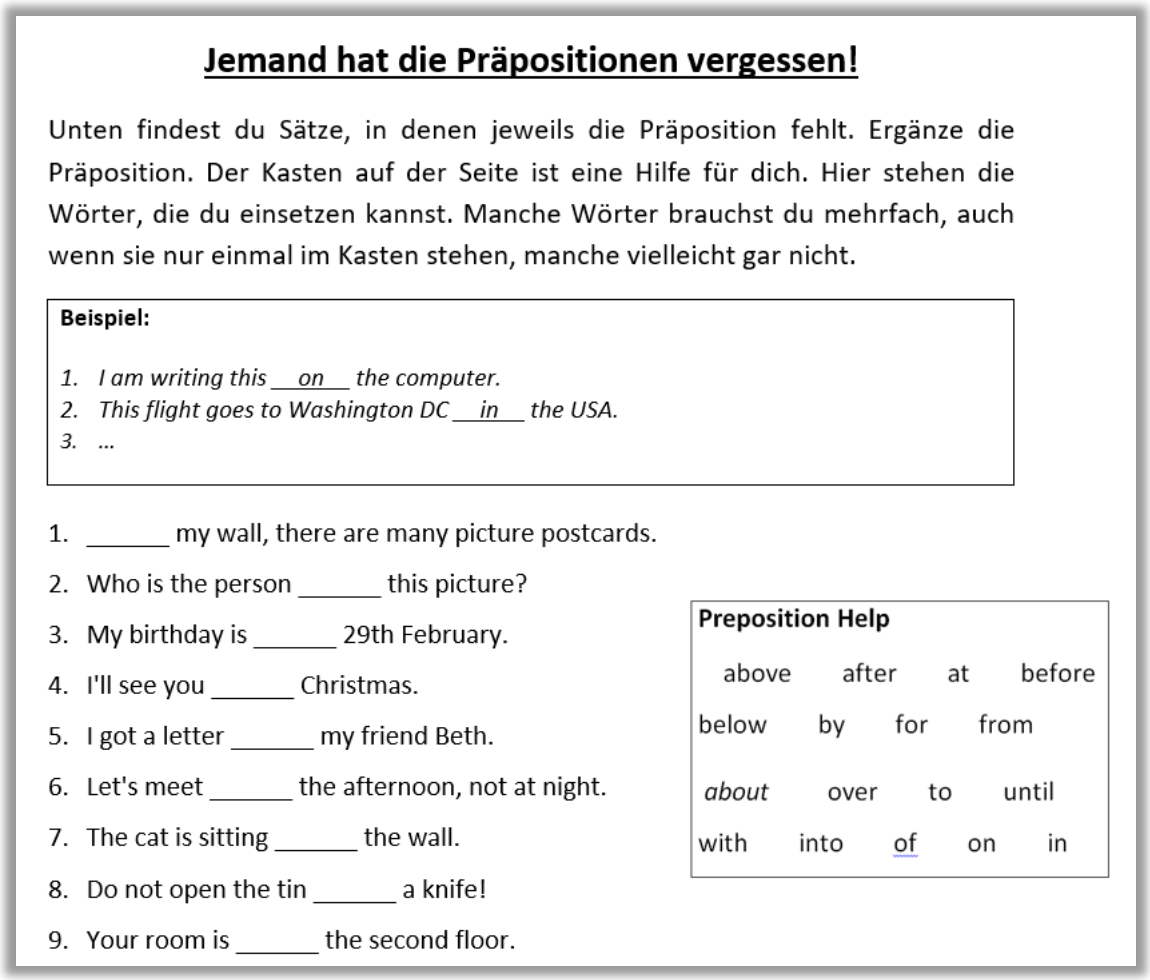
☐ true x false: _____Helen wants to meet outside the bank_____ .

Figure 3.12 Fictional reading comprehension

3.3.3.4 Prepositions

To design four parallel preposition tests, a single long test with 107 items – gap filling exercises (see figure 3.13) - was designed and given to a group of students that did not participate in the study to fill in. To ensure that the students were not tested on their ability to differentiate prepositions from other words, but to use them correctly, a box containing the prepositions needed to fill in the gaps was provided.

In a first step, the item difficulty of each item was calculated by dividing the number of participants that answered the item correctly by the number of total participants. Items that were extremely easy (>.8) or difficult (<.1) were omitted, the rest of the items were then divided into three difficulty levels. In a second step, the items of different difficulty levels were distributed across four tests to an equal number, leaving each test with 20 items: three with low difficulty, nine with medium difficulty and eight with high difficulty. In a third step the four resulting tests were checked for parallel test reliability (see chapter 3.3.4.2).



Jemand hat die Präpositionen vergessen!

Unten findest du Sätze, in denen jeweils die Präposition fehlt. Ergänze die Präposition. Der Kasten auf der Seite ist eine Hilfe für dich. Hier stehen die Wörter, die du einsetzen kannst. Manche Wörter brauchst du mehrfach, auch wenn sie nur einmal im Kasten stehen, manche vielleicht gar nicht.

Beispiel:

1. *I am writing this __on__ the computer.*
2. *This flight goes to Washington DC __in__ the USA.*
3. *...*

1. _____ my wall, there are many picture postcards.
2. Who is the person _____ this picture?
3. My birthday is _____ 29th February.
4. I'll see you _____ Christmas.
5. I got a letter _____ my friend Beth.
6. Let's meet _____ the afternoon, not at night.
7. The cat is sitting _____ the wall.
8. Do not open the tin _____ a knife!
9. Your room is _____ the second floor.

Preposition Help

above after at before
below by for from
about over to until
with into of on in

Figure 3.13 Preposition Test

3.3.3.5 Psychological Scales: SELLMO and SESSKO

The SELLMO (Skalen zur Erfassung der Lern- und Leistungsmotiviation) scales were developed by Spinath et al. (2012) to measure students' goal orientation. These consist of four scales: learning goal orientation, performance-approach goal orientation, performance-avoid goal orientation and work avoidance. The target group of the scales are students from third grade up to university students. The average time needed for the test is supposed to be 7-15 minutes. To fill in the scales, the student need to rate 31 items (e.g. *In der Schule geht es mir darum neue Ideen zu bekommen.*[17]) on a five-point Likert-scale ranging from 1 (strongly disagree) to 5 (strongly agree).

The SESSKO (Skalen zur Erfassung des schulischen Selbstkonzepts) were developed by Schöne et al. (2012) to measure students' self-concept. These scales also consist of four individual scales measuring the students' self-concept on the basis of the three reference standards (social, criterial and individual) and one scale without reference (absolute). The scales are also aimed at students from grade

[17] translation: For me school is about getting new ideas.

3 onward up to university students and, according to the authors, need around 7-15 Minutes to be filled in. In 22 items, the students needed to rate their abilities. The frame of reference is framed by the phrasing of the item, e.g *Ich kann im Englischunterricht weniger … mehr als meine Mitschüler(innen)*[18] or *Ich komme mit den Aufgaben im Englischunterricht schlechter zurecht als früher … leichter zurecht als früher*.[19] The original SESSKO items are actually only phrased for school in general and were rephrased for English lessons in particular for this study by substituting *at school* with *in the English lessons* for each item.

The divergent and convergent validity of both scales was shown by the authors and split-half and retest reliability is given, see the manuals (Schöne et al., 2012; Spinath et al., 2012) for detailed descriptions reporting results of internal consistency and factor analyses.

3.3.3.6 Further Questions

Demographic Data

The students were asked nine questions concerning their personal background. The first question asked about their gender and the second question their age. Question three, four and five are about their English language experiences at primary school: for how many years they learned English there, when they started and if they had already read words or sentences in English. Question six asks if they speak another language apart from German at home and if so which one. The last questions are about their migration background, asking if their parents (mom and dad) or they themselves were born in another country. The following list shows the original German questions the students were asked:

1) Ich bin ein □ Junge □ Mädchen

2) Wie alt bist du? _____ Jahre

3) Wie viele Jahre hast du in der Grundschule Englisch gelernt? _____ Jahre

4) In welcher Klasse hast du begonnen Englisch zu lernen? _____ Klasse

5) Hast du in der Grundschule schon Wörter oder Sätze auf Englisch gelesen?

□ Ja □ Nein

6) Sprecht ihr in der Familie noch eine andere Sprache außer Deutsch?

□ Nein □ Ja : wenn ja, welche? _____

7) Ist deine Mutter in Deutschland geboren? □ Ja □ Nein

[18] translation: In Englisch class I can do less … more than my classmates.
[19] translation: Compared to before, I cope better … worse with the tasks in my English class.

8) Ist dein Vater in Deutschland geboren? □ Ja □ Nein

9) Bist du in Deutschland geboren? □ Ja □ Nein

Reading Behaviour

The students were also asked nine questions about their reading behavior. The first question is how many books they themselves own and the second question is how many books their family owns. The third question asks how many English books they own and the fourth question how many English books – except for those from the English library at school – they have read so far. Questions five and six ask if their parents read in their free time and if yes how often. Questions seven and eight are the same questions about their own reading and question nine asks what they like to read, when they read.

These are the nine questions in their original German wording:

1) Wie viele Bücher besitzt du? Falls du es nicht genau weißt, so ungefähr. _____ Bücher
2) Wie viele Bücher gibt es in eurer Familie (so ungefähr)? _____ Bücher
3) Wie viele englische Bücher besitzt du? _____ Bücher
4) Wie viele englische Bücher hast du - außer denen aus der englischen Bibliothek in der Schule - bis jetzt gelesen? _____ Bücher
5) Lesen deine Eltern in ihrer Freizeit Bücher? □ Ja □ Nein
6) Wenn ja, an wie vielen Tagen in der Woche lesen sie in ihren Büchern?

   □ 1-2 Tag □ 3-4 Tage □ 5-6 Tage □ jeden Tag

7) Liest du in deiner Freizeit? □ Ja □ Nein
8) Falls ja, wie viele Stunden in der Woche liest du? _____ Stunde
9) Was liest du? (Du kannst auch mehrere Sachen ankreuzen)

   □ Bücher □ Zeitschriften □ Comics □ Zeitungen □ im Internet □ anderes, nämlich _____

Media Consumption

The students were also asked about their additional media consumption, to see if there were additional sources of English input, apart from books. They were asked about watching videos in English (question one), playing English video games (question two), chatting with English speaking friends (question three) or other activities not mentioned in the first questions. They were also asked to name the hours a week they spent with these activities.

These are the four questions in their original German wording:

1) Schaust du in deiner Freizeit englische Videos im Internet (z.B. auf Youtube)?

□ Ja □ Nein Falls ja, wie viele Stunden in der Woche? _____ Stunden

2) Spielst du englische Videospiele in deiner Freizeit?

□ Ja □ Nein Falls ja, wie viele Stunden in der Woche spielst du? _____ Stunden

3) Chattest du in deiner Freizeit mit anderen Leuten auf der Welt auf Englisch?

□ Ja □ Nein Falls ja, wie viele Stunden in der Woche chattest du? _____ Stunden

4) Gibt es sonst noch etwas, was du regelmäßig auf Englisch machst?

□ Ja, nämlich _____ □ Nein

Falls ja, wie viele Stunden in der Woche? _____ Stunden.

Reading Preferences

To find out what books the students preferred reading, they were asked to help improve the English library and assist when deciding which books should be bought next by rating the following aspects.

In einem Buch ist mir wichtig das …	*In a book it is important to me that …*
… ich alle Wörter verstehe.	… I understand all the words.
… es viele Bilder hat.	… it has a lot of pictures.
… die Sätze kurz sind.	… the sentences are short.
… die Sätze etwas länger sind.	… the sentences are long.
… die Schrift groß ist.	… the font is big.
… dass die Geschichte spannend ist.	… the story is exciting.
… das Buch kurz ist.	… the book is short.
… das Buch lang ist.	… the book is long.
… es mehrere Bücher aus der Serie gibt.	*… there are several books from one series.*
… in dem Buch Wörter übersetzt sind.	*… there are glossaries for unkown words.*
… in dem Buch auch Aufgaben zu dem Text sind.	*… there are additional tasks to the text.*
… in dem Buch nur die Geschichte und keine zusätzlichen Aufgaben sind.	*… not additional tasks.*

3.3.4 Quality Criteria of Test and Measurement Theory

The previous sections have provided a first overview of the tests and questionnaires the students had to fill in regularly. The following sections elaborate on the measures I took to meet the three main quality criteria objectivity, reliability and validity, and seven further quality criteria for the four tests that were designed for this study.

3.3.4.1 Objectivity

A test is objective, if it is independent of the test administrator and the evaluator (Moosbrugger & Kelava, 2012).

Objectivity of application was ensured by standardizing the instructions for students and teachers. All the instructions needed (anonymization procedure, test structure, procedure, time and importance) were presented to every student in written form at the beginning of the test and with example questions before each task. The teachers supervising the tests also received a written instruction, where they were advised to prohibit cheating, to read the first page of the test together with the students and to not answer any questions. Moreover, the students took the test in a subject where they were in their homeroom class composition and not during the individual assistance program. Therefore, students of all experimental and control groups took the test in the same setting and since the teachers did not know which student participated in which program this setting is a double-blind design.

Objectivity of analysis is also granted due to clear criteria sets for marking the tests. Since the majority of the tests consisted of multiple-choice questions and Likert-scales, only few parts were prone to error and needed detailed attention concerning grading criteria. While correcting the cloze tests, we made lists of solutions that we accepted and of those that we declined as unsuitable solutions, to ensure a coherent grading across tests. This also holds for the open reading comprehension questions and the preposition gap filling exercises. Since marking an answer as wrong or right was not always straight forward, the answer key was revised and discussed in a team after grading a first set of tests. If the Likert scales were not clearly ticked (more than one step ticked or marked in between two steps) the item was left out.

3.3.4.2 Reliability

A test is reliable if it is exact and free of systematic measurement errors (Moosbrugger & Kelava, 2012). The tests in this study consists of various subtests that have to be addressed separately concerning this criterion.

Since the language proficiency test parts were all developed for this study, their reliability had to be verified first. The aim was the development of different test versions of similar items (item twins) of the same difficulty. Students that did not participate in the study were asked to complete these test versions – one group for each test type - to check to what extent the variation of scores within one student is constant across the test versions (parallel test reliability). In practice, this means that the test versions are run with the same sample and the results of the different tests are then correlated: Rel(x) = Corr (x_A, x_B), with x_A and x_B being results of the parallel tests (Moosbrugger & Kelava, 2012).

To prevent students from copying and to eliminate possible sequence effects, the order of the test versions was varied from test to test, meaning that each student received a different order of the test versions. Table 3.4 – 3.9 summarize the correlations for the parallel test analyses of these different test versions.

Table 3.4 Parallel test reliability of cloze test versions

	(1)	(2)	(3)	(4)	**(5)**	**(6)**	**(7)**	**(8)**
(1) A Cool New Friend (Derkow Disselbeck & Schwarz, 1997)	1							
(2) A Pirate Story (Derkow Disselbeck & Schwarz, 1997)	.31	1						
(3) A Day in the Life of Jack Hanson (Harger & Schwarz, 2013)	.18	.68**	1					
(4) A Day in the Life of Bill and Ben (Harger & Schwarz, 2013)	.43*	.43*	.24	1				
(5) The Saturday Match (Derkow Disselbeck & Schwarz, 1997)	.58**	.57**	.44*	**.62****	1			
(6) I Hate Sport (Harger & Schwarz, 2013)	.62**	.50**	.16	**.59****	**.78****	1		
(7) White Rabbit (Dolch, Dolch, & Patterson, 1961)	.51**	.60**	.33*	**.60****	**.70****	**.75****	1	
(8) The Shock Team (Harger & Schwarz, 2013)	.48**	.59**	.39*	**.57****	**.65****	**.67****	**.48****	1

N = 26; *p < .05; **p < .01; test versions used for parallel tests printed in bold

As the correlation results of the cloze test versions in table 3.4 above show, four tests have almost acceptable correlations of *Rel > .60* (Evers, 2001): *The Saturday Match, I Hate Sport, White Rabbit* and *The Shock Team*. Some of the results are slightly below *Rel = .60*, but since *N=26* is very small and the results are all highly significant this is not very problematic. Additionally, the average

points achieved on each of the four test versions was calculated. The test version with the highest average was used in test one, the one with the lowest average in test four.

The result of the non-fictional reading comprehension was not sufficient for the first trial and therefore adjustments were made to the tests and new test versions were created and retested until the reliability was acceptable. Table 3.5 summarizes the results of the first trial in the first line and the second trial in the second line. Results in the first trial were far from acceptable – many tests had a correlation near zero, some even had negative correlation, only few correlations were significant at all. The first test versions, consisting of only four multiple-choice questions per version, were revised, complemented by two additional open answer questions, and retested with a different group. The results (second line in each row of table 3.5), are better for some tests, but not good enough to attain the four necessary parallel tests. Therefore, only the tests *Leprosy* and *Globe* were retained, new tests were created and taken into a third round with again another group of students (table 3.6).

Table 3.5 Parallel test reliability of non-fictional reading comprehension versions 1 and 2

	(1)	(2)	**(3)**	(4)	(5)	**(6)**	(7)	(8)	(9)	(10)
(1) Flowers	1									
(TeeVee, 2016)										
(2) Marco Polo	.40*	1								
(TeeVee, 2016)	.10									
(3) Leprosy	.23	-.07	1							
(TeeVee, 2016)	-.01	.05								
(4) Week	.45**	.28	.51**	1						
(TeeVee, 2016)	.41*	.04	-.36*							
(5) Dead Sea	-.16	-17	.34*	.12	1					
(TeeVee, 2016)	.05	.05	-.04	-.00						
(6) Globe	.42*	.43*	.25	.20	.31	1				
(TeeVee, 2016)	.04	.07	**.64****	.29	-.05					
(7) Senses	.15	.02	.44**	.19	.32*	-.10	1			
(TeeVee, 2016)	-.08	.12	.38*	.20	-.43*	.37*				
(8) Seasons	.13	.14	.17	.46**	.06	-.00	.14	1		
(TeeVee, 2016)	.26	.16	.04	.30	-.04	-.15	-.14			
(9) Gravitation	-.16	.23	-.17	.03	.19	-.08	.06	.-01	1	
(TeeVee, 2016)	.29	-.03	-.32	.25	-.01	-.10	-.11	.15		
(10) Molecules	.05	.03	.37*	.16	.24	.27	.50**	.51**	.05	1
(TeeVee, 2016)	.09	.19	**.65****	.12	.13	.27	.26	.11	-.13	

N = 28 and N = 23; *p < .05; **p < .01; acceptable results printed bold

Table 3.6 Parallel test reliability of non-fictional reading comprehension versions 3

	(1)	(2)	**(3)**	**(4)**	**(5)**	(6)	**(7)**	(8)	(9)
(1) Planets (TeeVee, 2016)	1								
(2) Meteor (TeeVee, 2016)	.45*	1							
(3) Heat (TeeVee, 2016)	.64**	.32	1						
(4) Wind (TeeVee, 2016)	.56**	.46*	**.70****	1					
(5) Leprosy (TeeVee, 2016)	.62**	.62**	**.60****	.71**	1				
(6) Warmth (TeeVee, 2016)	.61**	.50*	.55**	.56**	.39*	1			
(7) Globe (TeeVee, 2016)	.71**	.58**	**.62****	.61**	**.66****	.55**	1		
(8) Animals (TeeVee, 2016)	.44*	.41*	.21	.31	.42*	.56**	.58**	1	
(9) Birds (TeeVee, 2016)	.46*	.14	.33	.32	.35*	.33	.39*	.62**	1

N = 24; *p < .05; **p < .01; test versions used for parallel tests printed bold

Most of the results of the third versions of the non-fictional reading comprehension were significant (table 3.6) and with *Rel > .60* acceptable for the analysis of group difference (Evers, 2001). The matches with the highest pairwise correlations *Heat, Wind, Leprosy* and *Globe* were chosen as components for the four tests.

The fictional reading comprehensions also proved difficult at first. Not all tests correlated significantly and only a few tests had an acceptable high correlation (see table 3.7). Possible candidates would have been *Job Offer, Mail from Mike* and *Taking Train*, but a fourth candidate did not fit. *Strange Walk* for example correlated well with *Job Offer* and *Mail from Mike* but not with *Taking Train*. Adding further questions, as done for the non-fictional reading comprehension, was not an option because the texts are quite short and it was difficult to come up with further suitable questions. Additionally, difficulty analysis[20] of the different versions showed that some texts seemed to be more difficult than

[20] Analyzed by calculating the item difficulty (number of participants who answered the item correctly / number of total participants)

others were. To achieve a mixture of easy and difficult questions and a balanced difficulty across the versions, two texts each were matched and a second correlation was calculated with the combined results (see table 3.8).

Table 3.7 Parallel test reliability of fictional reading comprehension single versions

	(1)	(2)	(3)	**(4)**	**(5)**	**(6)**	(7)	(8)	(9)	(10)
(1) Peter Parker (Edwards, 2014)	1									
(2) A Bad Day (Edwards, 2014)	.43*	1								
(3) Mail from Greta (Edwards, 2014)	.28	.43*	1							
(4) Job Offer (Edwards, 2014)	.21	.56**	.56**	1						
(5) Mail from Mike (Edwards, 2014)	.25	.46*	.33	**.61****	1					
(6) Taking the Train (Quinault & Baumgärtner, 1991)	.29	.40*	**.68****	**.65****	**.58****	1				
(7) Kates Kite (Beile, 1996)	.11	.26	.48**	**.68****	.27	.66**	1			
(8) Bonzo the Dog (Beile, 1996)	.32	**.59****	.02	.29	.26	.18	.21	1		
(9) Strange Walk (own creation)	-.01	.50*	.53	**.62****	**.63****	.44*	.42*	.27	1	
(10) Letter about Mexico (own creation)	.05	.30	.01	.23	.08	.18	.36*	.21	.30	1

N = 24; *p < .05; **p < .01; acceptable results printed bold

Table 3.8 Parallel test reliability of fictional reading comprehension matched versions

	(1)	**(2)**	**(3)**	**(4)**	(5)
(1) Mail from Greta & Bonzo the Dog	1				
(2) Mail from Mike & Kates Kite	.53**	1			
(3) Taking Train & Strange Walk	**.63****	**.85****	1		
(4) A Bad Day & Job Offer	**.72****	**.73****	**.72****	1	
(5) Peter Parker & Letter about Mexico	.42*	.34	.31	.45*	1

N = 24; *p < .05; **p < .01; acceptable results printed bold

All pairwise correlations of the first four combinations are significant and high. Even though the first and second combination do not correlate above *Rel = .60*, the correlation is still highly significant and with *Rel = .53* only slightly below *Rel = .60*. The first four combinations of the test versions were therefore used for the four tests.

The validation of the correlations of the preposition test versions (see section 3.3.3.4 on how these were designed) also show acceptable results (see table 3.9).

Table 3.9 Parallel test reliability of preposition test versions

	(1)	(2)	(3)	(4)
(1) Version 1	1			
(2) Version 2	.65**	1		
(3) Version 3	.66**	.60*	1	
(4) Version 4	.68**	.68**	.65**	1

N = 24; *p < .05; **p < .01

All in all, the results of the pilot run of the language proficiency tests show a sufficient parallel test reliability for analyzing group differences of *Rel > .60* (Evers, 2001). Therefore, it can be assumed that the test versions are parallel and of the same difficulty. They should, thus, lead to the same result when taken by the same person.

To ensure similar test conditions across all pre- and posttests, all tests took place in the third and fourth lesson of the day. Moreover, all tests had the same instructions and example questions. The students were rewarded with sweets after every test.

3.3.4.3 Validity

A test is valid if the test really measures what it is designed to measure (Moosbrugger & Kelava, 2012). This is a very important quality criterion, because a test can be objective and reliable without being valid. Weir (2005) developed a framework for conceptualizing reading test validity. He names five forms of validity that need to be assessed when designing a test: context validity, theory-based validity, scoring validity, consequential validity and criterion-related validity (see figure 3.14 on the next page). Test taker characteristics influence the setting of the tasks and the demand for the task (content validity), but also internal processes of the test takers while taking the task (theory-based validity). Choices made at theses stages influence the scoring methods needed to assess the test (scoring validity), which then again influence the interpretation of the scores (consequential validity) and their value (criterion-related validity).

From: Title: Examining Reading: Research and Practice in Assessing Second Language Reading. Authors: Hanan Khalifa, Cyril J. Weir © Cambridge University Press. Reproduced with permission of the Licensor through PLSclear.

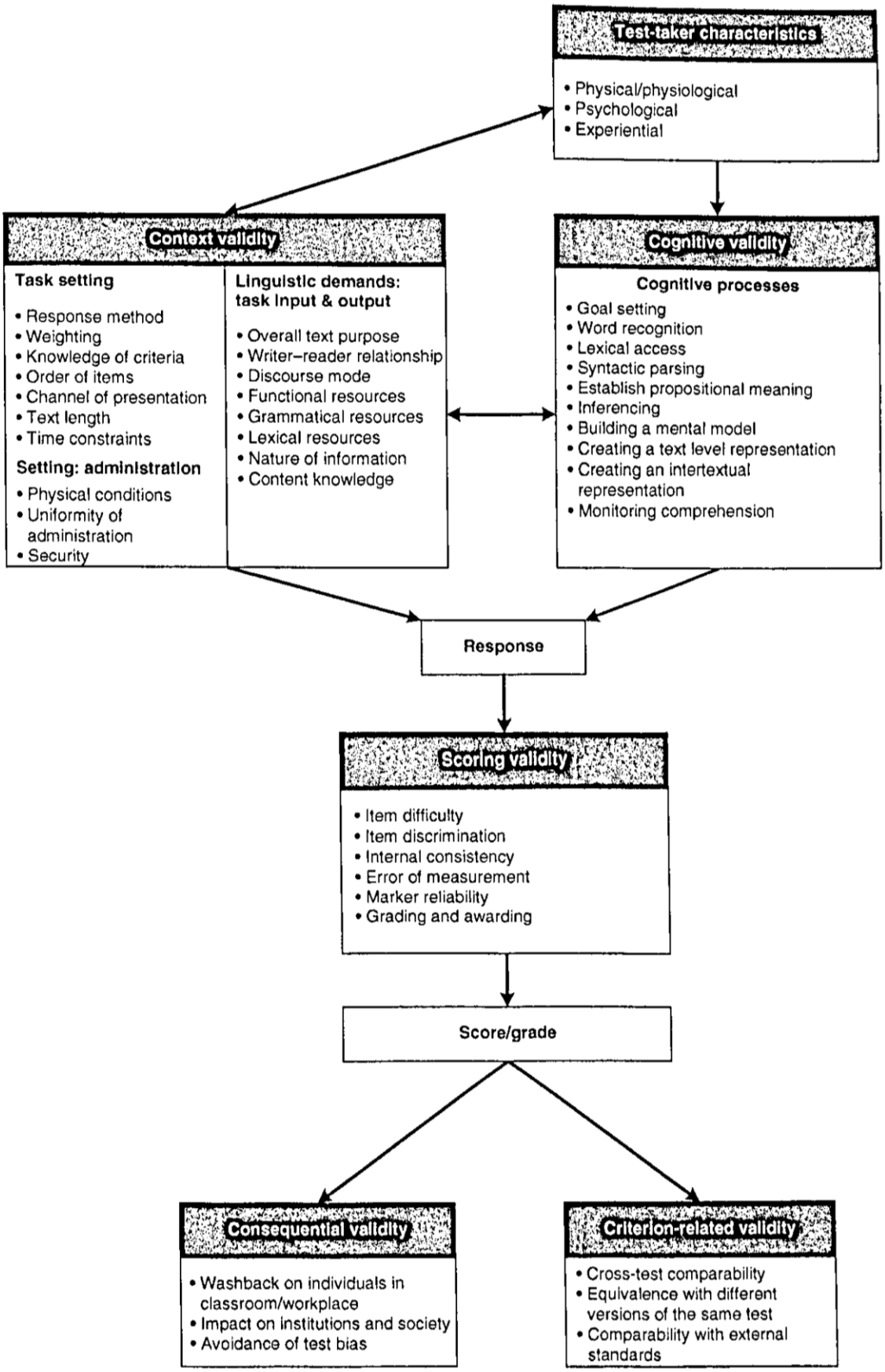


Figure 3.14 Framework Conceptualizing Reading Test Validity (Khalifa & Weir, 2009, p. 5)

Some of the five validities in this framework rather resemble other quality criteria in content and aim when using the structure of the three main and seven other quality criteria structure. Scoring validity, consequential validity and criterion-related validity corresponds to issues discussed in the previous reliability and following standardization and fairness sections. I used the terminology of the ten quality criteria, as for example described by Moosbrugger and Kelava (2012), for the superstructure of the section 3.3.4 and will therefore point to other sections when following the framework of Weir (2005) in this section on validity. When presenting how validity issues were faced in this study, I am going to focus on the two reading comprehensions and the preposition test. Since validity issues of cloze tests are rather specific and controversial, these have already been discussed in the cloze test section (see 3.3.3.1). The psychological scales are standardized scales and have been validated and discussed in the manual by the authors of the scales and will therefore not be discussed here (see section 3.3.3.5 for further information).

Test-Taker Characteristics

When designing a test, the physical/physiological, psychological and experimental characteristics of candidates need to be taken into account (Khalifa & Weir, 2009; Weir, 2005). The students participating in this study did not suffer from any severe physical or physiological impairments that had the potential to influence the results of the test. It is unclear, if students suffered from dyslexia, but since the group sizes are large and participants were randomized, I expect students with undiagnosed dyslexia to be distributed equally over the intervention groups. Additionally, it is not preventable that some students with short-term ailments like a cold or a headache participated in the tests, but these effects are probably equally distributed across the intervention groups as well and therefore neglectable.

The same holds for psychological influences, like a recent divorce of the parents, death of a pet or fight with a friend that might influence the results, but were not possible to control. As for psychological traits like motivation, the students were assessed with the SELLMO scales (Spinath et al., 2012) measuring goal orientation and the SESSKO scales (Schöne et al., 2012) measuring self-concept. The moderation effect of traits, that seem to be strongly related to academic achievements, were then taken into account as moderators when analyzing the results of the language tests.

Not only psychological, but also experimental factors, like education, cultural background, preparation and exam experience can influence the result of tests (Khalifa & Weir, 2009). Since all the students are in the same grade, at the same school, these factors are already controlled for by the cohort choice for this study. Additionally, the students of the experimental groups derived from different classes, randomizing teacher and experience effects. The students were advised that no

preparation was necessary for the tests and since the results had no consequences for the students, the pressure to achieve well on the tests was quite low.

Cognitive Validity or Theory-Based Validity

Cognitive or theory-based validity is an aspect of what is called construct validity in other frameworks. It means that the test's requirements should match cognitive processes that are necessary to fulfill the task (Khalifa & Weir, 2009). Here, the first question at stake is whether reading can be broken down into different subskills or components that can be tested individually or whether it is a unitary process – of course with different aspects to it, but not divisible. This discussion is mainly presented in chapter 2, but in total the current perspective is that when investigating the reading process as a cognitive performance it can be useful to break it down into different steps and components. Whether this is useful and legitimate when testing reading comprehension is rather controversial (Grabe & Stoller, 2020; Weir, 2005; Weir & Khalifa, 2008). There is, for example, no sufficient empirical evidence from factorial analyses that reading can be assessed with different independent subskills (Khalifa & Weir, 2009). According to Khalifa and Weir (2009), instead of focusing on different subskills, the focus when designing reading tests should therefore be on the analysis of the processes necessary to fulfill the given task and if these are in line with the test takers proficiency.

Weir (2005) differentiates between four different reading types or goals, each involving different skills and strategies (see left box in figure 2.6 of chapter 2). Since the texts used for my reading comprehension tests are very short, testing skimming and scanning strategies is not possible. The students expectably also do not have enough reading practice and language competence in their second language to even use these strategies. Additionally, when reading books in English or understanding texts in their school books, it is not the aim that they skim a text to get an overview or scan for information, they should read their text and understand what they are reading at the learning stage they are at. It could be possible though that the students use search reading strategies after having read the text for the first time, to find answers to the comprehension questions. Much more dominant in the everyday demands of the students is careful reading – on a global, as well as a local level. This matches the skills required in the reading comprehensions tests. Especially in the intensive reading interventions, students had to analyze certain structures, translate words they did not know or mimic the structure of the text in an own text, which involves a local understanding. In the extensive reading interventions, students had to develop an accurate comprehension of the text to understand what they were reading, even though the focus was not on understanding detail, but rather a global understanding. Focusing on these strategies goes hand in hand with the developmental stage: the students in 5th and 6th grade are still very occupied with low-level processing, additionally short texts

are more likely to be read using careful reading and bottom-up processing strategies (Khalifa & Weir, 2009).

Urguhart and Weir (2014) expanded their reading model with a goal setter (chooses the reading type suitable for the text and task) and a monitor (provides feedback about the reading process) –see also figure 2.6. Since the students know that they have to understand the text in the test, I assume that students rely on global careful reading strategies when reading the text (goal setter) and if this becomes difficult at times try to use local reading strategies to identify, for example, key words or structures, to restore global comprehension (monitor). Then, when reading the comprehension questions, they can either answer them remembering what they have read, or have to go back using search reading (goal setter) to find the correct answer. All in all, the language tests in this study match the reading types expected from the students in everyday situations at school, but are of different formats to ensure that a variety of skills and strategies are assessed. The cloze test and preposition test represent indirect task types, whereas the multiple-choice, true-false and short answer formats represent direct task types.

Context Validity

Context validity is also an aspect of construct validity and describes "the appropriateness of both the linguistic and content demands of the text to be processed, and the features of the task setting that impact on task completion" (Khalifa & Weir, 2009, p. 81) , which can be achieved by making the test as authentic and clear as possible (Weir, 2005).

Concerning the task setting, the rubric should be clear, unambiguous and easier than the tasks itself. Additionally, one should question if it could be reasonable to use the students L1 in the instructions, instead of the goal language of the language test (Weir, 2005). The rubrics of the language tests in this study were therefore kept short, using only 1-2 representative example(s). The understanding and clearness of the rubric was ensured by questioning students of the same grade, that did not participate in the study, beforehand on their understandability. The language of the tasks was phrased in German since the students' achievements on the tasks should not be influenced by difficulties in understanding the instructions and understanding what the purpose of the tasks is. The reading tasks mirror closely what the students were expected to learn in the reading interventions, namely to read and understand unknown texts.

Using multiple-choice formats when assessing reading comprehension is criticized, because these response formats are proscribed to only test recognition. They are also said to be quite restricted to assessing complex reading techniques and therefore real-life processing. Additionally, correct answers by guessing need to be taken into account (Weir, 2005). These arguments have their eligibility,

but I also have to take the English level of my students as beginning learners into account. Many students at this early stage are inhibited when they have to produce an oral or written output in English and expecting this competence when assessing reading comprehension - namely only what the students have understood – would, in my opinion, strongly bias the results in this case. Nevertheless, I wanted to address these undisputable disadvantages, by taking some measures. I therefore decided on phrasing the multiple-choice questions of the non-fictional reading comprehension in German to prevent the students from searching for the words of the questions and answers in the text and finding the correct answer without having understood the text. By answering the question in their first language, they need to establish a concept of their English reading first. Furthermore, I added two short answer questions that had to be answered in English to not solely rely on multiple-choice questions. The non-fictional reading comprehension questions were phrased in English and only required a true-false decision. To meet problems here, students were required to correct wrong statements.

The order of the tests was kept constant over all test periods. I decided to start with the cloze test and place the prepositional exercise between the reading comprehensions to give the students at least a little diversion between the tests. I also assumed the students to be cognitively exhausted after filling in the language tests and perceived the psychological scales to be less demanding, moving these to the end. The students did not have any time pressure when completing the tests, so that speed would not bias the results.

The demand on the material itself was that the texts should match the students' reading level and experience. Therefore, the texts were adapted from older school books, reading books or exercises that the students also encounter in this form in their school books or when reading books from the library. Of course, no texts were taken from books that the students could potentially already know. The English level of the students did not make it possible to use longer texts. Table 3.10 gives an overview of the scope and content of the texts used. Weir (2005) questions if a number of small texts with about 250 words is enough to see if students are able to cope with a larger amount of reading, as it is for example required in academic studies. I agree, but additionally assume for my tests that most students cannot cope with a larger amount of text as that used in the test (about 600 words in total for cloze test and reading comprehensions together), especially not without a large drop-out rate due to a too excessive demand and therefore a loss of motivation. Results would not be reliable anymore. Additionally, one longer text would not have made it possible to test different text types, strategies and topics (see also table 3.10).

Of course, the down-side of using these shorter texts is that skimming and scanning cannot be tested explicitly and the tests are limited to testing careful reading and maybe searching for detail, as

described in the previous section. Nevertheless, I do not see this as a reliability issue, because exactly this kind of reading is what the students need to learn and engage in at this early stage of learning English (see also chapter 2 on learning to read in the L2 and reading strategies). Therefore, the test characteristics match the students' reading level. Additionally, the topics of the cloze test and fictional reading comprehension are broad and should be of interest to both boys and girls.

The texts and questions of the non-fictional reading comprehension, in contrast, were designed to be as specific as possible so that students would not be able to understand the text and answer the questions with common knowledge. Of course, difficult scientific vocabulary should not stand in the way of comprehension at this stage, therefore words that proved to be too difficult but essential for understanding in the piloting studies were either substituted with an easier alternative, explained in brackets or translated. Moreover, especially because phrasing non-fictional reading comprehension questions would again involve using scientific vocabulary, the German phrasing ensured that it is not the students' comprehension of the questions but of the text that was tested and again increased validity.

Table 3.10 Words and Content of the Test Versions

Test Version	**cloze test**	**non-fictional reading comprehension**	**fictional reading comprehension**
1	188 words; TTR[21] ≈ .51 mystery story about a team of children watching a strange quest at a hotel	116 words; TTR ≈ .67 Leprosy	144 words and 179 words TTR for both parts ≈ .58 letter about a holiday in Spain and a dog reporting about his day
2	183 words; TTR ≈ .39 tale about a rabbit who wants to find a friend	110 words; TTR ≈ .52 experiment description about the earth as a globe	174 words and 183 words TTR for both parts ≈ .50 e-Mail about working on a farm for the summer and a story about a girl flying a kite
3	186 words; TTR ≈ .40 boy talking about how much he hates sports	89 words; TTR ≈ .65 movement of air due to heating and cooling of the earth	177 words and 219 words TTR for both parts ≈ .51 a dialogue on a train ride and a report about incidents when taking a walk
4	171 words; TTR ≈ .52 report about going to a football match with an integrated love story	99 words; TTR ≈ .63 swelling and shrinking of steel bridges through temperature changes	206 words and 194 words TTR for both parts = .50 a diary entry about a shopping trip gone wrong and a story about a climbing adventure and the contents of a phone call

[21] type-token-ratio (TTR) = $\frac{total\ number\ of\ different\ words}{total\ number\ of\ words}$

Scoring Validity

Scoring validity requests the scoring of a test to be reliable (Khalifa & Weir, 2009). Scoring validity in this framework is, as already mentioned, very similar to the concept of reliability and has already been discussed in the previous section. I will therefore only briefly summarize and point to some aspects here. The sample in this study is quite large with a low drop-out rate, allowing possibilities of various randomization that have been discussed in the previous sections. The tests were pretested and given to experts to judge the clearness of instructions, readability and layout. Especially the layout, acceptability and recognition value were boosted by printing the test as an A5 booklet. Instructions on the front page, for the different tests and the example items were, of course, kept constant for every test version. A first version of the solutions was examined for correctness by language experts. The first scorings, especially of the cloze test and open answer questions, were performed twice in cooperation with my student assistant to achieve a consent on what answers are considered as correct and every correct and wrong answer for each item was supplemented in the original answer key. Since anonymity was guaranteed to the students, the booklets neither contained information on the student's name, nor on the intervention group they belonged to. This information was merged in the data table at a later point, after all test results had been put in on the basis of the anonymity number.

Whereas parallel test reliability was designed to be acceptable, internal consistency proved to be an issue. Post reliability analysis of the full survey revealed only moderate reliabilities for the language tests: Cronbach's Alpha $.65 < \alpha < .79$ for the preposition tests; Cronbach's Alpha $.49 < \alpha < .65$ for the fictional reading comprehension and Cronbach's Alpha $.49 < \alpha < .63$ for the non-fictional reading comprehension, except for test 2 with $\alpha = .11$. Leaving the last value aside for now, it is questionable if Cronbach's Alpha is a good value for judging the reliability of these tests. First, Cronbach's Alpha is sensitive to the number of items on a scale (Field, 2018). Therefore, it is no surprise that the preposition test has the highest values with 20 items, followed by the fictional reading comprehension with 10 items and the non-fictional reading comprehension with 6 items. Second, as already discussed above, reading competence is a multifaceted construct and therefore not all items contribute to measuring one single construct. Nevertheless, the non-fictional reading comprehension scores are a bigger problem. The internal consistency result is far from acceptable and not in line with the results of the other tests. Unfortunately, test two is used for judging the competence gain of the first and of the second intervention period, fortunately it is not used for judging the competence gain of the entire semester. Its flaws therefore only become significant when looking at the direct effects, but not the delayed effects. I will take these results into account and refer back to them when discussing the results of the study in chapter 5.

Consequential Validity

Tests are taken for specific purposes and these purposes could have major consequences on the lives of the participants, especially if results do not meet the quality criteria. Since the tests were only designed for this study and therefore only used in this context, this validity issue is only of minor importance in this study. The students neither receive a result of their test, nor is it recorded or used for individual assessment. Still, when taking into account that deducted results of this study might influence further research, views and educational policies on this topic, this issue is worth discussing.

Aspects concerning the tests fairness, beneficing, appropriateness and forgery resistance are discussed in the following sections on further quality criteria and not in this section on validity.

Criterion-Related Validity

It can be assumed that tests with reading comprehension questions measure reading competence and fill in the blank preposition exercises measure prepositional knowledge. Thus, no concurrent validation was performed with the language tests.

The criterion-related validation of the psychological scales is much more relevant in this context and is reported and discussed in the corresponding manuals. The scales show convergent and divergent validity.

3.3.4.4 Further Quality Criteria

Scaling

For the data analyses I want to run, the scales of the dependent variables need to have at least an interval scale level. The language tests meet this criterion; due to the point system for each correct answer and a zero line they are even ratio scaled.

The scale level of the Likert-scales of the psychological tests then again proves to be more difficult. On the one hand, you can argue that the steps between the different answer-options are not necessarily equal intervals and therefore the scale only meets ordinal scale criteria. On the other hand, this classification would lead to tremendous additional effort and complexity when analyzing the scales. That is why Likert-scales are often tacitly, classified as metric scaled. Experts have not come to a conclusion on this discussion with reasonable arguments for and against each classification on both sides (Richard Williams, 2019). In this study, the Likert-scaled items will be treated as interval scaled variables.

Standardization

The SELLMO and SESSKO scales are normed by the authors for different age groups. According to the German DIN-norm 33430, normed tests have to be reevaluated every 8 years. The latest version and standard tables are from 2012 fulfilling this norm.

Test Efficiency

Due to a large number of test trials and participants and a tight financial budget, there was no possibility of using language tests that had to be paid for per use. Suitable free language tests with enough parallel versions were not available, which is why the language tests were developed specially for the purpose of this study. For the psychological scales, only scales that I was granted permission for to use (SELLMO and SESSKO) were used. Therefore, test costs were kept at a minimum.

The time that the students needed to fill in the tests was also closely calculated and pretested to ensure that the tests needed no longer than an average of 60 minutes to fill in, leaving slow students with 30 minutes of extra time in the 90-minute-long lessons. Questions that only needed to be surveyed once were distributed evenly across all pre- and posttests (see table 3.3).

For the economization of preparation time, evaluation effort and natural resources in terms of paper, it would have been much easier to have an online version of the tests. Unfortunately, this was not possible for several reasons: Due to copyright regulations, some material could only be used in a printed version. Additionally, over 100 students needed to be tested at the same time, which is far beyond my school's digital equipment and broadband internet capacity; and my experience is that even though my 5th grade students might have a high click competency on their smart phone, taking them to the computer lab to do something can be quite nerve-wracking. Therefore, over 2000 tests had to be printed, labeled and stapled to little booklets by hand and then corrected and captured manually, which was not very economic but without alternative.

Benefit and Appropriateness

A test is only beneficial, if it has a practical relevance and if it is more useful than harmful (Moosbrugger & Kelava, 2012). The practical relevance in the tests is to evaluate the effectiveness of different reading intervention types. In the context of English didactics, evaluating which methods have what effect on the students' English competence is important when it comes to teaching effectively and when deciding how resources like time and money are invested.

I do not believe that the students had any harm from taking the test, since filling in the test does not differ much from filling in a workbook page. The psychological skills also did not have any potential

of triggering distress. Therefore, in my opinion, the students were not psychologically or physically overstrained by participating in the test and 90 minutes, the time of one double-lesson, is an appropriate time span.

Forgery Resistance and Fairness

The link between the test content and the texts and tasks students encounter at school are, as already mentioned, very similar. Therefore, the tests have a high face validity, which is prone to be corrupted by the test takers. To ensure that the students still answered the questions as honestly as possible the students were reminded that the test was anonymous at the beginning of the test. Moreover, they were reminded that the results of the tests were important and that, even though I would not trace back results, they should be committed to the tasks and fill in the information truthfully. This unfortunately did not turn out as hoped for. I could judge from some of the comments on the test and answers that some questionnaires were not filled in truthfully. Two students were excluded from further interventions and from participating in the library, because we, as a school, did not want to leave the comments that they wrote into the test booklets without a reaction, but this was an exception. Judging from the way some questions were answered, I believe that some students did not engage in the tasks: When answering the cloze tests, they were fast at giving up and not answering anything at all. For other tasks, only the multiple-choice questions were answered. Since I am controlling my results for motivational factors and since I assume that these students are distributed evenly across the different intervention groups, I hope that these factors did not influence the results and took these issues into account when making the decision to exclude multiple outliers (see 3.5.2.1).

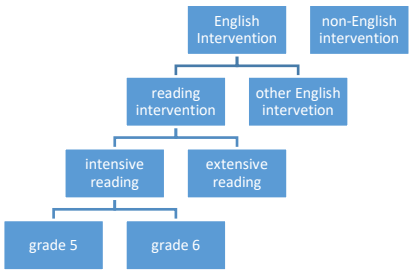
As already mentioned in the sections above, the texts and items of the language competence texts were chosen to appeal to girls as well as boys and therefore I believe the test to be fair concerning gender. Migration background and sociocultural status could play a role, due to a lack of cognitive/academic language proficiency (CALP) (Cummins, 2008) and might not only concern the German and English reading competence, but also background knowledge that might help to understand the texts and test taking strategies. This is a big issue for the entire academic achievement of educationally alienated children and could not be ruled out but just controlled for in this study.

3.4 Summary of Dependent and Independent Variables

Table 3.11 summarizes the variables examined in this study. The main independent variable in focus are the interventions. The table shows all possible contrasts, but in most cases only reading vs. English / non-English interventions and intensive vs. extensive reading are of interest and therefore

reported. Gender and if the students speak German and/or another L1, the students L1 leisure reading (in h) and their language competence in the different language tests at T1 are also independent variables – some in the function of covariates, some were tested as moderators.

Table 3.11 Summary of Variables

<table>
<tr><th>main independent variables
(fixed factor)</th><th>additional independent variables
(covariates)</th><th>dependent variables</th></tr>
<tr><td rowspan="2">intervention group with possible contrasts
 
</td><td>• gender
• German L1
additional independent variables
(moderators)
• leisure reading behavior
• L2 language competence (T1)
○ cloze test
○ preposition test
○ non-fictional reading comprehension
○ fictional reading comprehension</td><td>• overall results pre-/posttest (direct and delayed)
• subtest results pre-/posttest (direct and delayed)
○ cloze test
○ preposition-test
○ non-fictional reading comprehension
○ fictional reading comprehension
• school grade improvement</td></tr>
<tr><td colspan="2">• goal orientation (SELLMO)
○ learning goals (Lernzielorientierung)
○ performance approach orientation (Annährungs-Leistungszielorientierung)
○ performance avoid orientation (Vermeidungs-Leistungszielorientierung)
○ work avoidance (Arbeitsvermeidung)
• self-concept gain (SESSKO)
○ criterial (kriterial)
○ individual (individuell)
○ social (sozial)
○ absolut (absolut)</td></tr>
<tr><td colspan="3">Explorative Analyses of:
• distribution of goal orientations and self-concept within intervention groups before and after the intervention
• amount and type of borrowed books</td></tr>
</table>

Goal orientation and self-concept scales were analyzed as independent variables (main and interaction effects on language competence) and dependent variables (effect of intervention type on self-concept and goal orientation changes from T1 to T4). The gain in the four language tests and a total score (sum of all four tests) were dependent variables in this study. The direct pre-/posttest result is the difference between the test taken directly before the intervention and directly after the intervention (intervention period), whereas the delayed pre-/posttest is the difference between the first and the fourth test (semester period). See also table 3.2 for when which test was taken.

3.5 Data Analysis

3.5.1 Data Aggregation

To perform analyses with the data, some items first had to be aggregated to a single variable or to variables suitable for performing the desired statistical analyses. For the SELLMO and SESSKO scales, the manual (Schöne et al., 2012; Spinath et al., 2012) advises to add up items of subscales to obtain a total score for these. If no more than three items are missing, the manual of the SELLMO scales (Spinath et al., 2012) suggests calculating the mean of the remaining items and multiplying by the number of total items of the scale to obtain a total score. The procedure is the same for the SESSKO scales (Schöne et al., 2012), except that there is no suggestion to the minimum of items that have to be answered to perform this procedure. These corrections were performed for data sets with missing items. Reliability analysis showed that almost all results were acceptable to good (Cronbach's Alpha: .70 ≤ α ≤ .86 for SELLMO scales and .83 ≤ α ≤ .90 for SESSKO scales), except for two scales: SELLMO Lernziele T1 (α = .58) and SELLMO Vermeidungs-Leistungsziel-Orientierung T1 (α = .64). The reliability of these two scales cannot be improved by leaving out an item. I will use the result of these two scales in the analysis, but these results have to be interpreted carefully.

Raw scores of the different language proficiency tests also had to be aggregated to competence gain scores for each semester. This was done in three steps:

1. Due to the different number of total points for the different test parts, not the absolute but the relative value ($test\ score\ t_x = \frac{\text{achieved points } t_x}{\text{total points } t_x}$; $x \epsilon \{1; 2; 3; 4\}$) for each test t was calculated and used. This value describes the ratio of points achieved in a specific test in relationship to the total number of points that could have been achieved on this specific test, which has two advantages: First, it allows to compare students' results of different subtests and second, when averaging the results to a total test point score, each subtest delivers the same proportion to the total score, independent of the number of total points of the subtest.
2. To ensure that a slight difference in test difficulty or a competence gain of all students at a certain point in the semester will not interfere with the results when intervention groups of different semesters are combined to one intervention group, the test score results were related to the results of the non-English intervention group average (baseline group or control group) for the same test:

   $corrected\ test\ score\ t_x = \text{test score } t_x - \emptyset\ control\ group\ (baseline)\ t_x;\ x \epsilon \{1; 2; 3; 4\}$

   These corrected test score can be interpreted as the difference between a student's result on a test to the test result of the students who did not have an English intervention on the same test. If the reading intervention is successful, you would expect this difference to be 0 for the test before the intervention and positive for the test after the intervention, because the student with the

reading intervention should perform better after the intervention compared to the students without a reading intervention. This would also be the case, if the second test is for example a bit more difficult, because this would hold for all students – the baseline group might perform worse than in the test before, but the students with the reading intervention should not perform as poorly as the student without the intervention and the difference would still be positive. Therefore, this measure is independent of potential differences in test difficulty, which is relevant in the next step.

3. To have a measure for the competence gain during the intervention period that the intervention took place (period p_x), the results for the test directly after the intervention and directly before the intervention have to be compared:

   $$competence\ score\ p_x = corrected\ test\ score\ t_{x+1} - corrected\ test\ score\ t_x;\ x\epsilon\{1;2;3\}$$

   Trivially, this scores is near 0 for the non-English intervention and you would expect the results to be positive for the other intervention groups engaging in English. If this is really the case and whether differences are significant or just coincidence or if other factors influence these differences, will be analyzed with the help of inferential statistics explained further in this chapter.

The students' grades in English are usually a combination of grades for class tests, vocabulary tests, oral participation, homework, etc. The combination and number of these grades differ from teacher to teacher, except for the number of class tests – two class tests are written each semester. The four class test grades in the year of the intervention were used to compute the independent variable grade. The average of the two class tests in the semester before the intervention were subtracted from the average of the two class tests in the semester of the intervention to obtain a measure for an improvement or decline in grade.

3.5.2 Bias Reduction

3.5.2.1 Outliers

To identify possible outliers, an analysis of the distribution of the z-scores for cloze test one showed that 94.8% of the values lie in a normal range of $|z| < 1.96$, 3.5% of the values lie in the potential outlier range $1.96 < |z| < 2.58$ and 1.2% of the values lie in the probable outlier range $2.58 < |z| < 3.29$. The number of values that lie in these ranges are consistent with a normal distribution – about max. 1% in the probable outlier and about max. 5% in the potential outlier range – and are therefore tolerable (Field, 2018). Nevertheless, there are two extreme z-scores (0.6 %) with $|z| > 3.29$. The results for cloze test two, three and four are similar and summarized in Table 3.12 below.

Table 3.12 Analysis of Outliers - Z-Score Results for Cloze Tests

	cloze test 1	cloze test 2	cloze test 3	cloze test 4
	n (%)	n (%)	n (%)	n (%)
extreme |z|>3.29	2 (0.6)	3 (0.8)	2 (0.6)	
probable outlier 2.58 < |z| < 3.29	4 (1.2)		3 (0.8)	
potential outliers 1.96 < |z| < 2.58	12 (3.5)	6 (1.7)	11 (3.1)	13 (3.8)
normal range |z| < 1.96	325 (94.8)	345 (97.5)	341 (95.5)	333 (96.2)

I analyzed the distribution of the z-scores, as I did for the cloze tests above in the same manner for the preposition tests and reading comprehensions (table 3.13 – table 3.15). Again, most of the values lie in a normal range of $|z| < 1.96$, a small, but acceptable number of the values lie in the potential outlier range $1.96 < |z| < 2.58$ and probable outlier range $2.58 < |z| < 3.29$ and are therefore unproblematic. Nevertheless, there are nine additional extreme z-scores (0.6 %) with $|z| > 3.29$.

Table 3.13 Analysis of Outliers - Z-Score Results for Non-Fictional Reading Comprehension

	NF-RC 1	NF-RC 2	NF-RC 3	NF-RC 4
	n (%)	n (%)	n (%)	n (%)
extreme |z|>3.29			1 (0.3)	
probable outlier 2.58 < |z| < 3.29	9 (2.6)	6 (1.7)		
potential outliers 1.96 < |z| < 2.58	10 (2.9)	12 (3.4)	5 (1.4)	6 (1.7)
normal range |z| < 1.96	324 (94.5)	336 (94.9)	351 (98.3)	340 (98.3)

Table 3.14 Analysis of Outliers - Z-Score Results for Preposition Test

	prep test 1	prep test 2	prep test 3	prep test 4
	n (%)	n (%)	n (%)	n (%)
extreme |z|>3.29	2 (0.6)	1 (0.3)		1 (0.3)
probable outlier 2.58 < |z| < 3.29	3 (0.9)	4 (1.1)	1 (0.3)	
potential outliers 1.96 < |z| < 2.58	6 (1.7)	9 (2.5)	10 (2.8)	15 (4.3)
normal range |z| < 1.96	332 (96.8)	340 (96.0)	346 (96.9)	330 (95.4)

Table 3.15 Analysis of Outliers - Z-Score Results for Fictional Reading Comprehension

	F-RC 1 n (%)	F-RC 2 n (%)	F-RC 3 n (%)	F-RC 4 n (%)
extreme |z|>3.29			4 (1.1)	
probable outlier 2.58 < |z| < 3.29		2 (0.6)	7 (2.0)	7 (1.9)
potential outliers 1.96 < |z| < 2.58	14 (4.1)	11 (3.1)	4 (1.1)	17 (4.9)
normal range |z| < 1.96	329 (95.9)	340 (96.3)	342 (95.8)	322 (93.1)

The same analyses were run for the SELLMO and SESSKO scales (see Appendix for results). Just as for the language tests, these results were all in all within the acceptable range of the normal distribution, except for some extreme z-score results. Data sets with extreme z-scores were excluded from further analyses: in total 34 from the non-English intervention group, 53 from the other English intervention, seven from the extensive reading intervention and 24 from the intensive reading intervention of grade 5 and grade 6. The following boxplots give an overview of the distribution of the data before and after the exclusion of the students.[22] Even though many data sets were excluded from future analyses, the boxplots in figure 3.15 show that these did not lead to severe changes in distribution.

The boxplots (figure 3.15) show the distribution of the language tests scores for the test taken directly before and after the intervention. The two diagrams in the first row show the uncorrected and the two diagrams in the second row the corrected test result differences (see section 2.3.1 above). The two diagrams to the left show the results before outlier analysis and elimination due to z-scores, the diagrams to the right the results after outliers were eliminated.

A comparison of the two diagrams on the left to those on the right shows that the exclusion apparently did not change the distribution of the results, but the number of outliers has decreased. Thus, the results will probably be more precise and unbiased. Additionally, a comparison of the diagrams at the top (uncorrected points) to those at the bottom (corrected points) also shows that these corrections did not lead to major changes in the distribution and did not distort the results.

[22] The choice between excluding the data or continuing, knowing that extreme outliers are not excluded, may seem like an option between bad and worse at first. I opted for bad instead of worse, because ignoring these extreme outliers was not an option for me: For one, the methods of analyses that I am using are very sensitive to extreme outliers, likely leading to distorted results. Secondly, having insight in how seriously students sometimes took the tests, outliers due to unwillingness are more likely than to extreme experimental effects. I do not want these effects of unwillingness to effect my analyses. Thirdly, effects that are now found, even though N is reduced, are even more reliable. Fourthly, working with fewer dependent variables would of course lead to less reduction, because the outliers for each scale sum up. The bigger number of outliers in total is therefore not a flaw of the internal consistency of this study per se, but a result of the design complexity. Fifthly and finally, I ran the main analyses without excluding outliers, just out of interest, and found very similar results, making the "cleaner" version even more attractive.

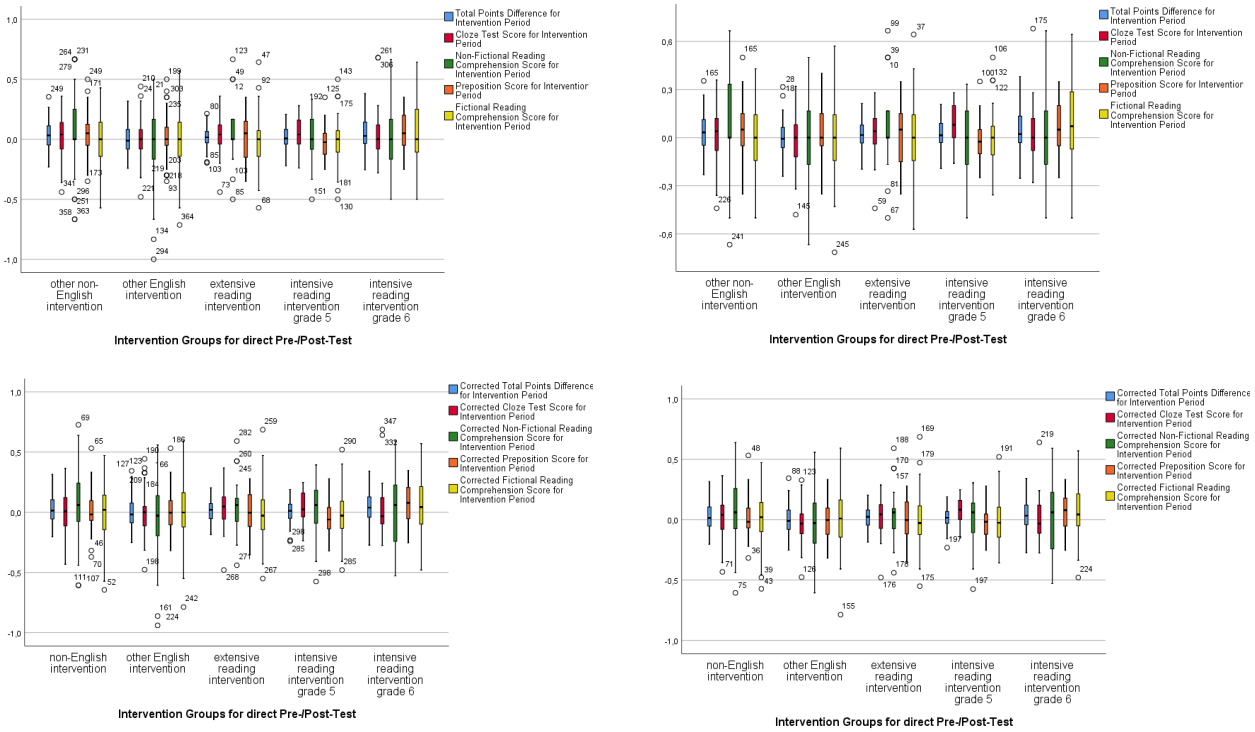


Figure 3.15 Direct Pre-Posttest Results for Language Tests Divided by Groups

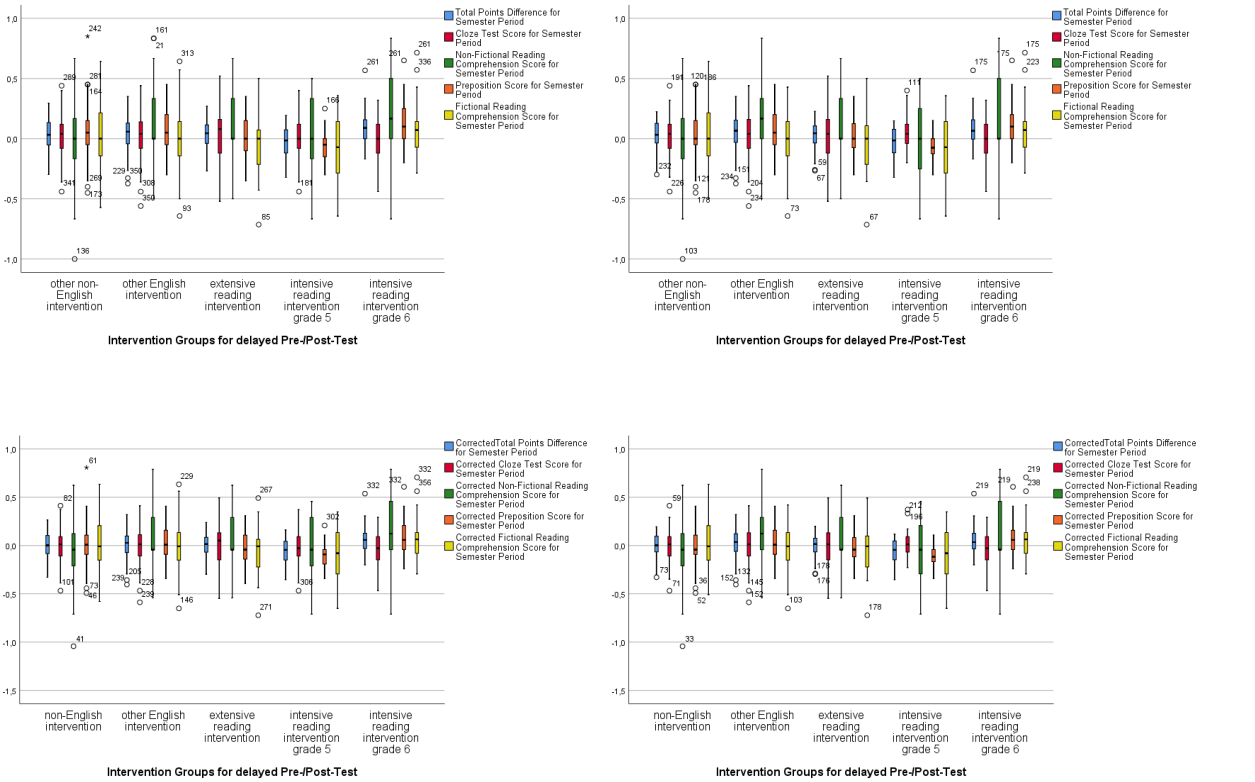


Figure 3.16 Delayed Pre-Posttest Results for Language Tests Divided by Groups

The same holds for the difference between the test points in T1 to T4 (semester period difference) depicted in the four diagrams below (figure 3.16 in the same manner as the result for the intervention period.

3.5.2.2 Checking Assumptions for Analyses

One condition that needs to be fulfilled when working with models that are based on t-statistics, is normality distribution. Of course, this does not refer to the entire sample; each score for each experimental group needs to be normally distributed. To assess the normal distribution of each of these scores, I looked at the frequency distributions and probability-probability plots (p-p-plots). In the p-p-plots all the data points fall very close to the ideal diagonal and the frequency histograms have the shape of a normal distribution (see Appendix). Therefore, it seems reasonable to assume normally distributed samples on the basis of these results. Additionally, I checked the data for heteroscedasticity[23] of variances, especially because the group sizes in my samples are not equal.

If the treatment variable and covariates are confounded, this would reduce the experimental effect. The distribution of the covariates used, are equally distributed across the intervention groups (see table 3.1), making it reasonable to assume that they are not confounded with the groups and can be used as covariates. Additional non-significant results of variance analyses with the covariates as independent and intervention as dependent variables underline this impression.

Scatterplots and regression lines of the covariates against the language competence gain for each intervention group already hint at the fact that the assumption of homogeneity of regression slopes, meaning that the relationship of the covariates and language competence gains (dependent variables) is the same for all intervention groups, cannot be made. Moderation analyses show significant results of the interaction term of some covariates and group contrasts on language test gains and grades (see tables on moderations analyses in the appendix and section 4.5 on moderation analyses results). This heterogeneity of regression slopes is actually an interesting result, supporting hypotheses that suggest interaction effects. ANCOVAs and MANCOVAs were therefore also run without covariates to see if this would lead to significant main effects and the need of more considerations. This was not the case. Also due to heterogeneity of regression slopes, moderation analyses were run separate for each covariate and contrast to achieve an overview of possible

[23] Checking heterogeneity of variances is important when testing groups to check that "these groups come from populations with the same variance" (Field, 2018, p. 237) and therefore confidence intervals and statistical assumptions based in these are unbiased. Homoscedasticity "means that the residuals at each level of the predictor(s) should have the same variance (Field, 2018, p. 387).

interactions. Moreover, hierarchical regression analyses were run separately for each intervention type. I will elaborate on this procedure in the following section.

3.5.3 Methods of Analysis

This study includes various variables and experimental groups, making it possible to look at various effects of intensive and extensive reading on different language competencies. In terms of analyses this of course also leads to large quantities of data that are tackled in the following.

To gain first insight into the performance development of the students in the four language tests, first descriptive statistics for the corrected test point gain are calculated for each experimental group. To gain some more insight into how the individual intervention groups performed over time, diagrams plotting the average points for each test period (T1-T4) are generated. These results are presented and described separately for each test type in section 4.1.

A second measure that is taken to gain an overview, but also to check for independence – an assumption needed to run the inferential statistical analyses - is looking at correlations of all experimental variables. The correlations are run separately for the experimental groups and are presented for the non-English, English, extensive and intensive reading intervention groups. Differences in correlation strength and direction between the different groups can already hint at possible effects expected in later analyses and are therefore very interesting. The correlations are presented in section 4.2 and divided by correlations among the psychological scales and correlations with language test gains.

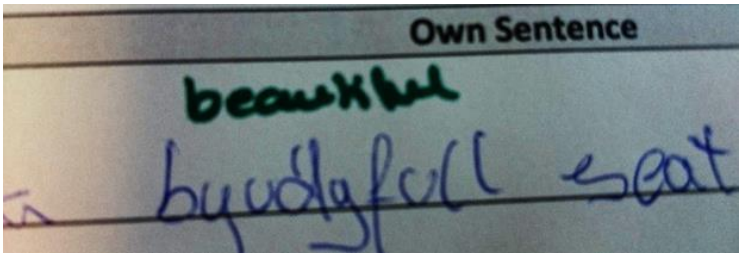
Third, main effects of intervention type on score gains in the language tests and psychological scales are analyzed using analyses of covariates (ANCOVAs) and multiple analyses of covariates (MANOVAs). In addition to a simple analysis of variance (ANCOVA), covariates are also entered into the analyses in ANCOVAs. A MANCOVA is used if more than one outcome variable is analyzed in one model. Total points and grades cannot be entered in the MANCOVA for the subtest results and are therefore run separately, because these variables are obviously not independent of each other. These results are presented in section 4.3. Moreover, goal orientation and self-concept are also analyzed in separate MANCOVAs and are presented in chapter 4.4.

These analyses only give insight into main effects so far, meaning that we only know if the intervention group or a demographic or psychological variable directly influence the English competence gain of the students. But some hypotheses of this study also assume moderation relationships. Therefore, in a fourth step, moderation analyses are run using the SPSS add in PROCESS.

Due to the various variables and heterogeneity of regression slopes issues, these were run as separate analyses. Significant results are presented and described in numbers and graphs in section 4.5.

Various main effects and interactions are now analyzed. We therefore know if different variables or interaction terms of variables influence language competence gain, but what is missing is information about the extent to which these variables influence language competence gain. The statistical question is, how much variance in the outcome variable (e.g. direct cloze test score gain) is explained by biographical variables. If you then add goal orientation variables, how much additional variance is then explained, how much more by adding variables on self-concept and how much more by adding variables on initial language competence. Hierarchical regression analyses can give insight into this question. Originally, I wanted to enter group (e.g. intensive vs. extensive reading) in a fourth step to see how much additional variance is explained, already having accounted for all other factors, but this plan was discarded after the various crossover effects showed in the moderation analyses. Adding this last step, would have meant integrating various interactions into the regression analyses and facing heterogeneity of regression slopes issues. I therefore opted for running the hierarchical regression analyses not as group differences (e.g. one analysis for intensive and extensive reading and then adding this difference as a variable), but as single analyses for each intervention group (non-English, other English, intensive, extensive). The advantages are obvious: Significant differences between these groups concerning the influence of various variables and interactions are already analyzed in the previous steps and are not of interest anymore. It is only of interest how big this effect is, which can be analyzed more straightforwardly and cleaner when looking at the groups separately. The results of the hierarchical regression analyses are presented in section 4.6.

4 Results



From a portfolio, student grade 5: A very creative form of spelling the word "beautiful"

The following sections present different effects and links between the dependent and independent variables of this study, by taking a descriptive and explorative look via diagrams, means and correlations and by using multivariate statistical approaches (see sections 3.5.3 for the methods of analysis). The descriptive results, which are organized by independent variable, are going to be presented first. The next part focuses on correlations within and between the psychological scales and language test results. This section is followed by the results of the analyses of variance for the language tests, grade and psychological scales and results of additional moderation analyses and hierarchical regression analyses for further insights on interaction effects and effect sizes. Chapter 4 closes with analyses concerning the reading behavior of the students, presenting and analyzing results of the questionnaire on reading preferences, analyses of what kind of books were borrowed from the library and the number of books borrowed.

When interpreting the reported numbers in the inferential statistics, it is important to keep in mind that the test scores are not absolute but relative scores, in other words, the percentage that the students increased from the pretest to the posttest. Moreover, these scores are corrected using the non-English intervention group as a baseline (see section 3.5.1 for details on procedure). These corrected values can be interpreted as the difference of increase between pre- and posttest compared to the control group (non-English intervention group). An average of for example -0.13 for an intervention group would mean that this group – on average - performed 13 percentage points worse in the posttest (compared to the pretest) than students in the non-English intervention control group.

4.1 Descriptive Statistics

4.1.1 Total Test Points

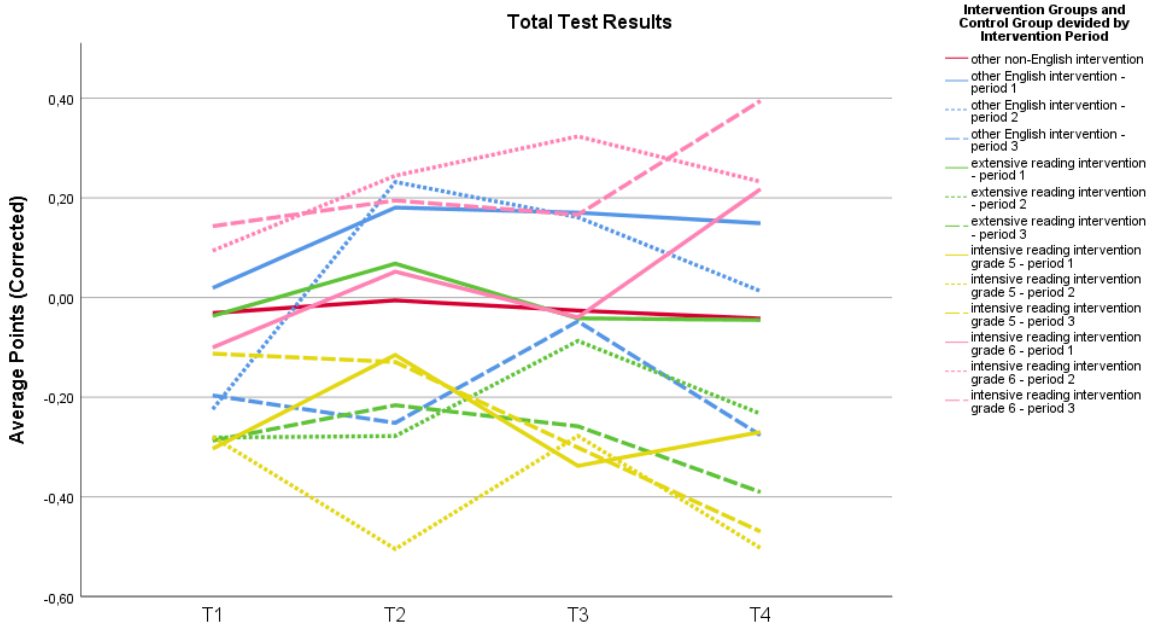


Figure 4.1 Average Total Points for Tests per Experimental Group

The diagram above shows the total points of the cloze test, the fictional and non-fictional reading comprehensions and the preposition test for each intervention period. As expected, and intended, the results for the non-English intervention stayed constant across all four tests. The results for the other English interventions (blue lines) are also as expected for the group with the intervention in period 1 (continuous line): There is an increase in points between T1 and T2 – this is where the intervention took place - and then only a small decrease between T2, T3 and T4 – no intervention between these tests anymore. An analogue pattern shows for the group with the intervention in period 3 (dotted line), but not for the intervention in period 2 (dashed line): Here you can see a large increase between T1 and T2 (no intervention at that time) that then even decreases between T2 and T3 (time of the intervention). The extensive reading and intensive reading intervention results are also partially surprising: Results increase for the interventions in period 1 and 2, but decrease for period 3. When comparing only T1 and T4 (delayed pre-/posttest results), the other English intervention groups (blue lines) and the intensive reading interventions in grade 6 (pink lines) performed better in T4 than in T1. In contrast, the intensive reading intervention group in grade 5 (yellow lines) performed worse in T4 than in T1 and the extensive reading intervention groups (blue lines) stay rather constant.

A look at the descriptive statistics shows that these are slightly contrary to what was expected as well. When looking at the tests taken directly before and after the interventions, students in the intensive reading interventions in grade 6 had the highest overall increase of points (*N* = 31, *M* = 0.04; *SD* = 0.13), followed by students in the non-English interventions (*N* = 74, *M* = 0.02; *SD* = 0.11), then the extensive reading interventions (*N* = 32, *M* = 0.01; *SD* = 0.10) and the intensive reading interventions grade 5 (*N* = 23, *M* = 0.01; *SD* = 0.11). Students in the other English interventions even slightly decreased in points (*N* = 76, *M* = -0.003; *SD* = 0.11). When looking at the results of T1 and T4 (delayed pre-/posttests), only students in the intensive reading interventions grade 6 (*N* = 31, *M* = 0.06; *SD* = 0.15) and students in the other English interventions (*N* = 71, *M* = 0.03; *SD* = 0.15) increased their results, the rest decreased (*N* = 74, *M* = -0.004; *SD* = 0.12 for the non-English intervention; *N* = 26, *M* = -0.003; *SD* = 0.13 for the extensive reading intervention and *N* = 23, *M* = -0.06; *SD* = 0.13 for the intensive reading intervention in grade 5).

4.1.2 Cloze Test

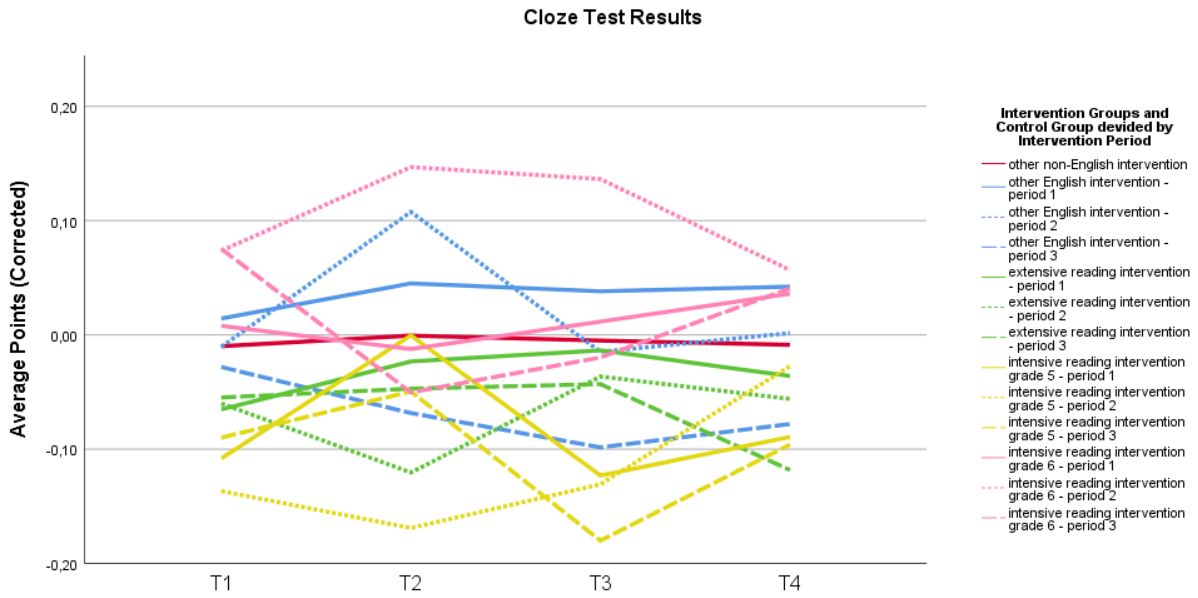


Figure 4.2 Average Cloze Test Results for Tests per Experimental Group

Just as the results for the total points in the previous sections showed some surprising results, the results for the cloze test (see figure 4.3) are also only partially in line with what one would expect. Students in the other English intervention in period 2 (dotted blue line) increased their results in the cloze tests from T1 to T2, where they did not have an intervention and then showed a decrease in results from T2 to T3, which is even lower than the results they started. In addition, students in the extensive reading intervention in period 3 (green dashed line) showed constant results in T1, T2 and T3, but then failed to hold that level or improve in the test after their intervention (T4). All the results

for the intensive reading interventions grade 6 (pink lines) are similarly odd. The rest of the groups achieved better results in the tests directly before and after their intervention, but partially show some unexpected results for the other tests, which I would have expected to stay constant. A look at the delayed pre-/posttest results (T1 compared to T4) shows that, even if the lines increase and decrease in between, the scores seem to be approximately at the same level. The only exceptions are the intensive reading intervention in period 2 in grade 5 (dotted yellow line) and the extensive reading intervention in period 3 (dashed green line).

A look at the descriptive statistics for the cloze tests are surprising as well, but also different from the results for the total points. In the direct pre-/posttests, students in the intensive reading interventions grade 5 had the highest overall increase (*N* = 23, *M* = 0.07; *SD* = 0.11), followed by students in the extensive reading interventions (*N* = 32, *M* = 0.01; *SD* = 0.15), then the non-English interventions (*N* = 74, *M* = 0.01; *SD* = 0.16) and the intensive reading interventions grade 6 (*N* = 31, *M* = 0.01; *SD* = 0.18). Students in the other English interventions slightly decreased in points (*N* = 76, *M* = -0.02; *SD* = 0.14). When looking at the delayed pre-/posttests, only students in the intensive reading interventions grade 5 (*N* = 23, *M* = 0.03; *SD* = 0.14) and students in the non-English interventions (*N* = 74, *M* = 0.01; *SD* = 0.16) improved their results, the rest decreased (*N* = 31, *M* = -0.001; *SD* = 0.20 for the intensive reading interventions grade 6; *N* = 26, *M* = -0.003; *SD* = 0.22 for the extensive reading interventions and *N* = 71, *M* = -0.004; *SD* = 0.19 for the other English interventions).

4.1.3 Preposition Test

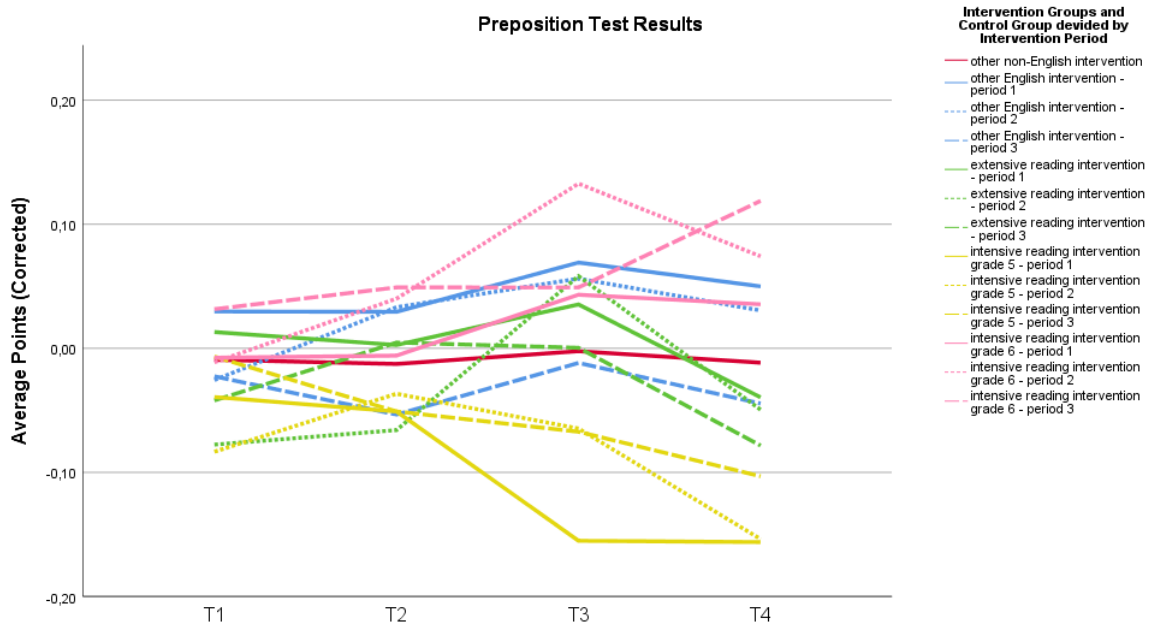


Figure 4.3 Average Preposition Test Results for Tests per Experimental Group

The preposition test results are even more of a mystery than the cloze test results. Only period 2 of the other English intervention group, the extensive reading intervention group and period 2 and 3 of the intensive reading intervention group show better results for the direct posttest compared to the direct pretest. The delayed pre- and posttest results (T1 compared to T4) show increases for the intensive reading interventions in grade 6 (pink lines) but decreases for grade 5 (yellow lines). The results for the extensive reading interventions (green lines) and the other English interventions (blue lines) seem to stay constant from T1 to T4.

The descriptive statistics for the preposition tests differ from the previous results. In the direct pre-/posttests, students in the intensive reading interventions in grade 6 had the highest overall increase (*N* = 31, *M* = 0.06; *SD* = 0.16), followed by students in the extensive reading interventions (*N* = 32, *M* = 0.01; *SD* = 0.17) and then the non-English interventions (*N* = 74, *M* = 0.01; *SD* = 0.14). The points of students in the intensive reading interventions in grade 5 and the other English interventions slightly decreased (*N* = 76, *M* = -0.002; *SD* = 0.14 and *N* = 23, *M* = -0.03; *SD* = 0.14). When looking at the delayed pre-/posttests, only students in the intensive reading interventions in grade 6 (*N* = 31, *M* = 0.07; *SD* = 0.19) and students in the other English interventions (*N* = 71, *M* = 0.02; *SD* = 0.17) increased their results, the rest decreased (*N* = 74, *M* = -0.002; *SD* = 0.19 for the non-English interventions; *N* = 26, *M* = -0.03; *SD* = 0.16 for the extensive reading interventions and *N* = 23, *M* = -0.10; *SD* = 0.12 intensive reading interventions grade 5).

4.1.4 Non-Fictional Reading Comprehension Test

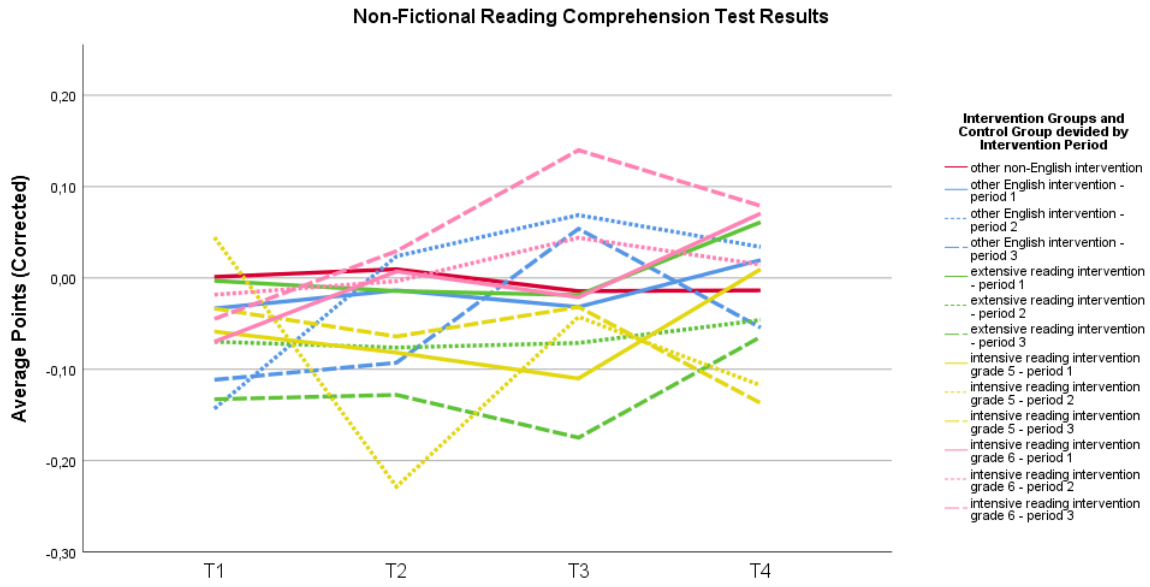


Figure 4.4 Average Non-Fictional Reading Comprehension Results for Tests per Experimental Group

The lines in figure 4.13 again show the test results for each intervention group, this time for the non-fictional reading comprehension groups. In contrast to the previous diagrams, it is striking in this diagram that almost all intervention groups performed worse in the different tests than the non-English intervention control groups (red line). Students in the intensive reading intervention groups (pink lines) managed to perform better than the control group in T4, as well as some other intervention groups, especially those that had their intervention in period 1. The results for the direct pre- and posttest directly before and after the interventions are again difficult to interpret, because there is no clear structure or trend visible. If one, for example, compares T3 and T4 for the dashed lines (period 3), you can see an improvement for the extensive reading intervention (green), and decreases for the other interventions. However, this pattern is different when looking at period 2 (dotted line; T2 and T3) and period 1 (straight lines T1 and T2).

Descriptive statistics for the direct pre-/post non-fictional reading comprehension tests show that students in the non-English interventions had the highest overall increase (*N* = 74, *M* = 0.06; *SD* = 0.25), followed by students in the extensive reading interventions (*N* = 32, *M* = 0.02; *SD* = 0.20), then the intensive reading interventions in grade 6 (*N* = 31, *M* = 0.02; *SD* = 0.27) and the intensive reading interventions in grade 5 (*N* = 23, *M* = 0.004; *SD* = 0.22).The scores of students in the other English interventions slightly decreased (*N* = 76, *M* = -0.01; *SD* = 0.26). When looking at the delayed pre-/posttests, students in the intensive reading interventions in grade 6 (*N* = 31, *M* = 0.11; *SD* = 0.35), students in the other English interventions (*N* = 71, *M* = 0.10; *SD* = 0.29) and students in the extensive reading interventions (*N* = 26, *M* = 0.06; *SD* = 0.28) increased their test results, the rest decreased (*N* = 74, *M* = -0.02; *SD* = 0.29 for the non-English interventions and *N* = 23, *M* = -0.08; *SD* = 0.35 for the intensive reading interventions in grade 5).

4.1.5 Fictional Reading Comprehension Test

Figure 4.5 shows that between T1 and T2 the straight lines (interventions in period 1) run almost parallel and all groups have a competence gain, however, this is not the same for period 2 (dotted lines; T2 vs. T3) and period 3 (dashed lines; T3 vs. T4). All the results decrease, except the scores of the students in the intensive reading interventions in grade 5 in period 2 (dotted yellow line) and the intensive reading interventions in grade 6 in period 3 (dashed pink line).

The descriptive statistics of the fictional reading comprehension are also slightly different to the other test parts. When looking at the tests taken directly before and after the interventions, students in the intensive reading interventions in grade 6 had the highest overall increase (*N* = 31, *M* = 0.06; *SD* = 0.23), followed by students in the other English interventions (*N* = 76, *M* = 0.02; *SD* = 0.21). The points of students in the non-English interventions, extensive reading interventions and intensive reading

interventions in grade 5 even decreased (*N* = 74, *M* = -0.002; *SD* = 0.21; *N* = 32, *M* = -0.007; *SD* = 0.32 and *N* = 23, *M* = -0.008; *SD* = 0.23). When looking at the results of T1 and T4 (delayed pre-/posttests), only students in the intensive reading interventions in grade 6 (*N* = 31, *M* = 0.07; *SD* = 0.23) and students in the non-English interventions (*N* = 74, *M* = 0.05; *SD* = 0.27) increased their results, the rest decreased (*N* = 71, *M* = -0.008; *SD* = 0.22 for the other English interventions; *N* = 26, *M* = -0.04; *SD* = 0.26 for the extensive reading interventions and *N* = 23, *M* = -0.08; *SD* = 0.29 for the intensive reading interventions in grade 5).

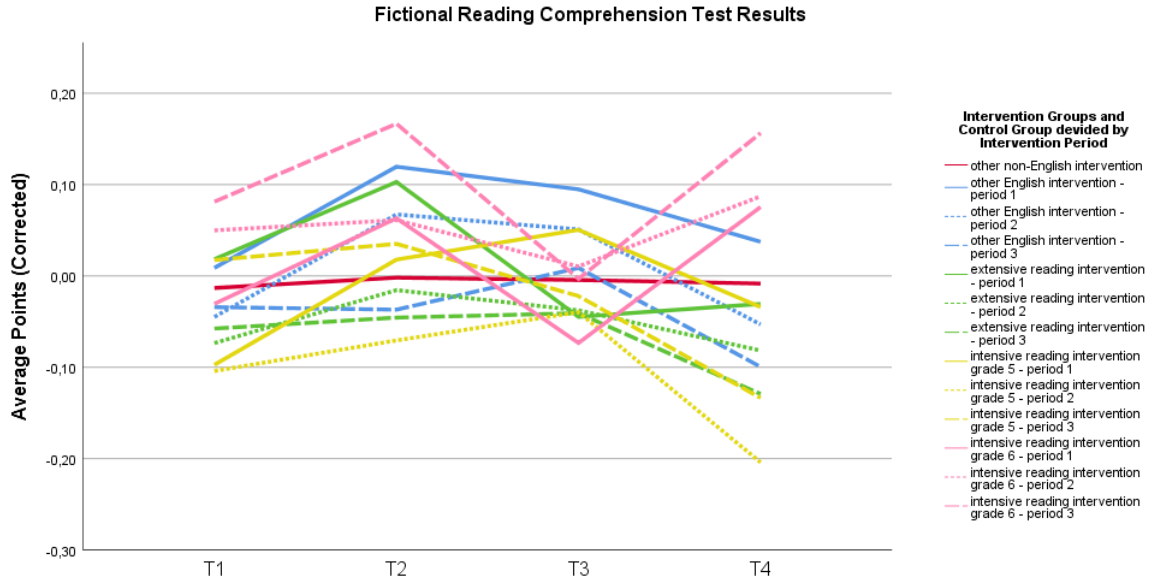


Figure 4.5 Average Fictional Reading Comprehension Results for Tests per Experimental Group

4.2 Correlations

Correlations between all dependent variables of this experiment, divided by groups (see Appendix, section 10.3.1 for correlation table) include numerous significant and highly significant correlations that I am going to report in the following. In order to give a detailed but still comprehensive and structured summary of the 2916 correlations, I am only going to report significant correlations and only those where a clear pattern is visible or results are contradictory. To support readability, I am also not going to report the correlations in numbers; all correlation results can be extracted from the table in the appendix.

4.2.1 Psychological Scales

First, the relationships within the goal orientation and self-concept scales are going to be presented: The four goal orientations and self-concepts measured in T1 highly, positively, correlate with those in T4. The correlations for the students participating in the extensive reading interventions were not significant for all pairs. Nevertheless, these psychological constructs seem to be very stable over time. The few non-significant correlations for the extensive reading comprehension group could hint at a change in learning goal orientation, performance-approach goal orientation and self-concept scores during the extensive reading interventions, but could also be due to internal consistency issues. The self-concept scales also correlate highly with each other pairwise, although this is, again, not always the case for the extensive reading comprehension group. These relationships are a bit more complex for the goal orientation dimensions: Learning goal orientation correlates positively with performance-approach goal orientation, and performance-avoid goal orientation correlates positively with work avoidance. Also, most groups showed a positive relationship between performance-approach and performance-avoid goal orientation. Additionally, learning goal orientation is negatively associated with performance avoid goal orientation and work avoidance.

Second, there are also significant relationships between goal orientation and self-concept: Learning goal orientation seems to be positively associated with all self-concept scores. This is not as clear for performance-approach goal orientation and self-concept: The results are significantly positive for the English (but non-reading) intervention groups, but not significant and negatively correlated for other groups. Performance-avoid goal orientation and work avoidance seems to be negatively associated with self-concept.

4.2.2 Correlations with Language Test Results

When looking at the correlations of the improvement (points gained) in the language test, the results for the intervention period points gain and semester period points gain correlate significantly positively for all language tests in all experimental groups. One exception are the extensive reading intervention groups that only significantly correlate for non-fictional reading comprehension results but not for the other test parts. Again, this could hint at experimental effects. The correlations between the different test parts logically reveal a strong positive relationship with the points gained in total due to the fact that the total points consist of the summation of the points gained in the test parts. Correlations between different test parts are also interesting. Even though there are some small significant correlations concerning some test parts, these only show for the other, non-reading, English intervention groups. Therefore, there is no hint at an intercorrelation between the different test parts, supporting the necessity, choice and construct of these language test parts.

Of special interest are of course the correlations between the scores on the psychological scales and the points gained on different test parts, since these correlations constitute the basis of the variance analysis performed and are directly linked to the hypotheses examined in this paper. Therefore, these correlations are reported in more detail in the following:

Learning goal orientation correlates positively with cloze test results for the extensive reading interventions, but negatively for the intensive reading interventions. Although these correlations were not significant for the psychological scales measured in T1 and T4 and the competence gain of the semester and intervention period, there is a visible pattern in these contrasting correlations. This could hint at a different effect of the two reading interventions on language proficiency. A similar effect shows between the gain in cloze test points and work avoidance, which correlate significantly negatively for the reading intervention groups, but significantly positively for the non-English intervention groups. This pattern is more visible for T1 and the intervention period, but can also be seen for the scores in T4 and the semester period.

There seems to be a negative relationship between self-concept scores and improvement in the language tests, which is quite surprising. Not all correlations are significant, but no matter whether significant or not, all correlations are negative or near zero. This negative association between self-concept and test performance is manifested when taking the improvement in grade into account: All significant correlations between self-concept and grade improvement are positive, meaning high self-concept scores are associated with a high positive grade difference and, therefore, a decrease in grade. There are only four exceptions to this picture: First, the absolute self-concept score in T1 of the non-reading English intervention groups has a positive significant correlation with non-fictional reading comprehension improvement (intervention period). Second, the absolute self-concept score T4 of the extensive reading intervention groups has a positive significant correlation with the preposition score improvement (semester period). Moreover, the criterial and individual self-concept score T4 of the extensive reading intervention groups has a positive significant correlation with cloze test improvement in the intervention period. These effects are interesting, because the change of the direction of the correlation from T1 (before the reading intervention) to T4 (after the reading intervention) hint at an effect of the extensive reading interventions on the impact of self-concept on language improvement.

4.3 Analyses of Variance

4.3.1 ANCOVA Results for Total Points

An ANCOVA analysis with the intervention groups as an independent variable, the difference in total points of the direct pre- and posttest as a dependent variable and the covariates gender,

migration background (here measured as another language as German as a mother language), time spent with L1 reading and points in the language subtests at T1 showed that there were no main effects of the intervention groups on the gain in total points in the language tests. However, the covariate of the points achieved in the non-fictional reading comprehension before the interventions (T1) was significantly related to the gain in total points, $F(1, 207) = 8.04$, $p < .01$, $r_{covariate\ non-fictional\ RC\ T1} = \sqrt{\frac{t^2}{t^2+df}} = \sqrt{\frac{(-2.84)^2}{(-2.84)^2+207}} = 0.19$. The coefficient β, signaling the direction of the effect, is negative in this case, therefore, lower scores in the non-fictional reading comprehension in T1 lead to higher competence gains on the overall test scores for the direct pre-/posttests.

These results also showed for the delayed pre-/posttest analyses $F(1, 197) = 18.52$, $p < .001$, $r_{covariate\ non-fictional\ RC\ T1} = \sqrt{\frac{t^2}{t^2+df}} = \sqrt{\frac{(-4.30)^2}{(-4.30)^2+197}} = 0.29$. Additionally, the covariates other L1 language, cloze test points in T1 and fictional reading comprehensions points at T1 had a significant effect on the total points ($F(1, 197) = 4.57$, $p = .04$, $r_{covariate\ other\ L1\ language} = \sqrt{\frac{t^2}{t^2+df}} = \sqrt{\frac{(-2.14)^2}{(-2.14)^2+207}} = 0.15$; $F(1, 197) = 8.06$, $p < .01$, $r_{covariate\ cloze\ test\ T1} = \sqrt{\frac{t^2}{t^2+df}} = \sqrt{\frac{(2.84)^2}{(2.84)^2+207}} = 0.20$ and $F(1, 207) = 8.04$, $p < .01$, $r_{covariate\ fictional\ RC\ T1} = \sqrt{\frac{t^2}{t^2+df}} = \sqrt{\frac{(-5.08)^2}{(-5.08)^2+207}} = 0.34$). The β coefficients were negative for all these effects, but for the cloze test points. Therefore, students who only speak German at home had higher gains in total points than students who speak another language (additional to German) at home. Moreover, students who had low scores in the first reading comprehension tests had higher total point gains, but students with high scores in the first cloze test also had high total point gains.

Planned contrasts revealed that these results did not differ significantly between the different intervention groups.

4.3.2 MANCOVA Results for Language Subtests

MANCOVA analyses of the results in the direct pre-/posttests, using Roy's largest root, showed that there was also no significant direct effect of the different intervention on the improvement in the different language tests, but several covariates were significant: the cloze test points in T1 (Θ = 0.63, F(4; 204) = 3.20, p = .014 for the direct pre-posttest and Θ = 0.72, F(4; 194) = 35.08, p = .001 for the delayed), the non-fictional reading comprehension points in T1 (Θ = 0.111, F(4; 204) = 5.66, p < .001 for the direct pre-posttest and Θ = 0.76, F(4; 194) = 36.63, p < .001 for the delayed), the fictional reading comprehension points in T1 (Θ = 0.68, F(4; 204) = 3.48, p = .009 for the direct pre-posttest and

Θ = 1.25, F(4; 194) = 60.48, p < .001 for the delayed) and the preposition test points in T1 (Θ = 0.52, F(4; 194) = 25.03, p < .001 only for the delayed pre-posttest).

Further separate univariate analyses on the outcome variables revealed that the covariates were only significant for the corresponding outcome variable: the non-fictional reading comprehension points in T1 were significantly related to the direct (F(1;207) = 22.54; p < 0.001) and delayed (F(1;197) = 128.52; p < 0.001) non-fictional reading competence gain, the fictional reading comprehension points in T1 to the direct (F(1;207) = 12.00; p < 0.001) and delayed (F(1;197) = 108.67; p < 0.001) fictional reading competence gain. The prepositions test score in T1, even though not showing an overall effect, was significantly related to the direct (F(1;207) = 4.57; p = 0.03) and delayed proposition test score gain (F(1;197) = 55.04; p < 0.001), but also to the non-fictional reading comprehension gain for the delayed pre-posttest (F(1;197) = 4.21; p = 0.04). Again, what is not in line with these findings is the cloze test score in T1, that was significantly related to the preposition test score gain in the direct pre-posttest (F(1;207) = 5.67; p = 0.02) and to all outcome variables in the delayed pre-posttest (F(1;197) = 50.92; p < 0.001 for cloze test gains; F(1;197) = 9.19; p = 0.003 for non-fictional reading comprehension points gains; F(1;197) = 20.35; p < 0.001 for preposition point gains and F(1;197) = 28.11; p < 0.001 for fictional reading comprehension score gains). Just as in the analyses of the total point gains, all β coefficients for these effects were negative, except for the cloze test effects, these were positive for the other language tests. High effects in the language tests in T1 are, therefore, associated with lower gains in the corresponding pre- and posttests. However, higher scores in the cloze test in T1 are associated with higher scores in delayed language test point gains from pre- to posttest.

Additionally, the L1 language was significantly associated with the delayed non-fictional reading comprehension score gain (F(1;197) = 4.78; p < 0.03). Another or additional L1 than German is associated with lower gains in the non-fictional reading comprehension scores for the delayed pre-posttests.

4.3.3 ANCOVA Results for Grades

ANCOVA results of the effects of the intervention groups on grades, with migration background (other or additional L1 language), gender and language competence test results at T1 as covariates, did not show any significant results. Therefore, the type of intervention did not have a statistically significant impact on grade improvement.

4.4 Intervention Effects on Psychological Scales

As reasoned in chapter 2, the reading interventions could not only have an impact on cognitive factors like language competence, but also on motivational factors like self-concept and goal orientation. Therefore, these psychological scales were not only measured at the beginning of the interventions (T1), but also at the end (T4). This section deals with the effects of the interventions on self-concept and goal orientation in their function as dependent variables.

4.4.1 Self-Concept

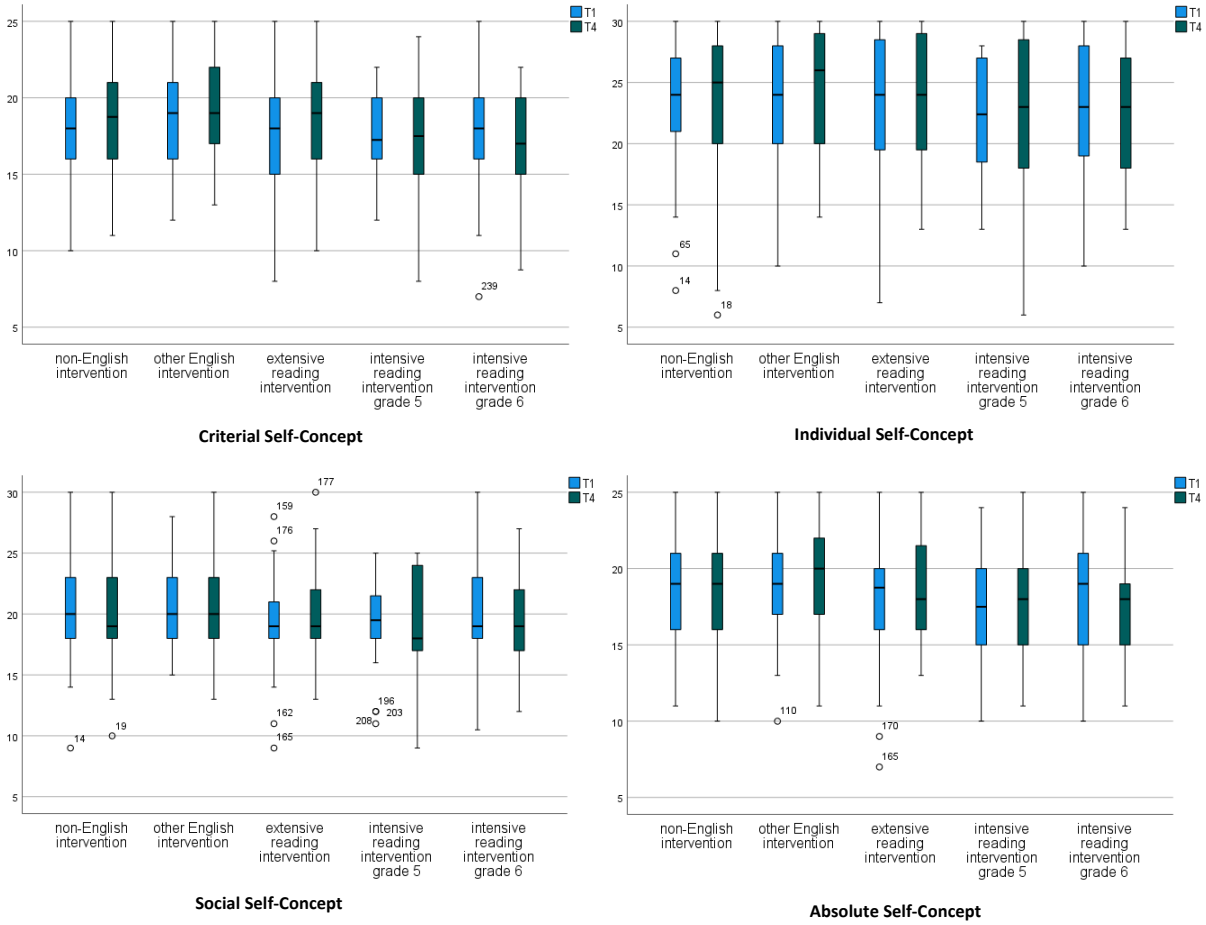


Figure 4.6 Self-concept Score Boxplots T1 vs. T4

The boxplots in figure 4.6 show the different self-concept scores before and after the interventions (T1 – before - blue; T4 – after - green). Criterial and individual self-concept increased in the non-English intervention groups, whereas social self-concept decreased and absolute self-concept stayed about the same from T1 to T4. The criterial and social self-concept of the other English interventions stayed at about the level and the individual and absolute self-concept increased. The extensive reading intervention groups increased in criterial self-concept, but decreased in absolute

self-concept, individual and social self-concept stayed about constant. Intensive reading interventions differed between grade 5 and grade 6: Whereas individual self-concept stayed nearly constant for both groups, absolute and criterial self-concept increased for groups in grade 5, but decreased for groups in grade 6; this was the other way around for social self-concept.

To investigate these possible differences using inferential statistics, a self-concept score gain (T4-T1) was calculated for each self-concept dimension in a first step and analyzed as independent variables in a MANCOVA in a second step. Roy's largest root showed a significant effect of the intervention type on the increase or decrease in self-concept; $\Theta = 0.05$, $F(4; 218) = 2.51$, $p = .04$. Separate univariate analyses on the outcome variables using contrasts revealed a highly significant difference of the students' criterial self-concept gain between the students participating in the extensive and those participating in the intensive reading interventions ($p = .01$). Whereas the criterial self-concept increased for the extensive reading intervention groups ($N = 74$; $M = 1.37$; $SD = 4.97$), it decreased for the intensive reading interventions ($N = 23$; $M = -0.67$; $SD = 4.59$ for grade 5 and $N = 31$; $M = -0.90$; $SD = 2.85$ for grade 6). Correlation results had already hinted at a possible contrary effect of these two reading intervention types on criterial self-concept (see section 4.2).

Additionally, analyses of the covariates showed that gender significantly influenced the increase in criterial academic self-concept ($F(1; 225) = 3.77$; $p = .05$): Girls showed significantly higher improvements than boys.

4.4.2 Goal Orientation Scales

Analogue to self-concept in the previous section, figure 4.7 shows boxplots of the goal orientation scores for each dimension before and after the interventions (T1 – before - blue; T4 – after - green). Learning goal orientation seems to stay constant for the non-English and the other English intervention groups from T1 to T4, slight increases in the extensive reading interventions and slight decreases in the intensive reading intervention groups. Performance-approach goal orientation slightly decreases for the non-English, other English and extensive reading intervention groups, but slightly increases for the intensive reading intervention groups. Performance-avoid goal orientation slightly increases for the non-English intervention groups, stays about the same for the extensive reading intervention groups and decreases for the other English and intensive intervention groups. Work avoidance, which will prove to be a strong moderator in the following section, stayed about the same for the non-English and other English interventions groups, increased in the extensive reading intervention groups and decreased in the intensive reading intervention groups.

Analogue to the self-concept scores, MANCOVA analyses were performed to analyze the effects of the interventions on goal orientation scores. Roy's largest root did not show a significant effect of

the intervention type on an increase or decrease in goal orientation scores, Θ = 0.03, F(4; 218) = 1.60, p = .18. Yet, separate univariate analyses on the outcome variables using contrasts revealed a significant difference of the students' work avoidance gain between the students participating in the extensive and those participating in the intensive reading interventions (p = .05). As the boxplots already showed, these scores were higher for T4 than T1 for the extensive reading intervention groups (N = 26; M = 1.83; SD = 9.96) and lower for the intensive reading interventions (N = 23; M = -0.35; SD = 5.93 for grade 5 and N = 31; M = -1.99; SD = 6.72 for grade 6).

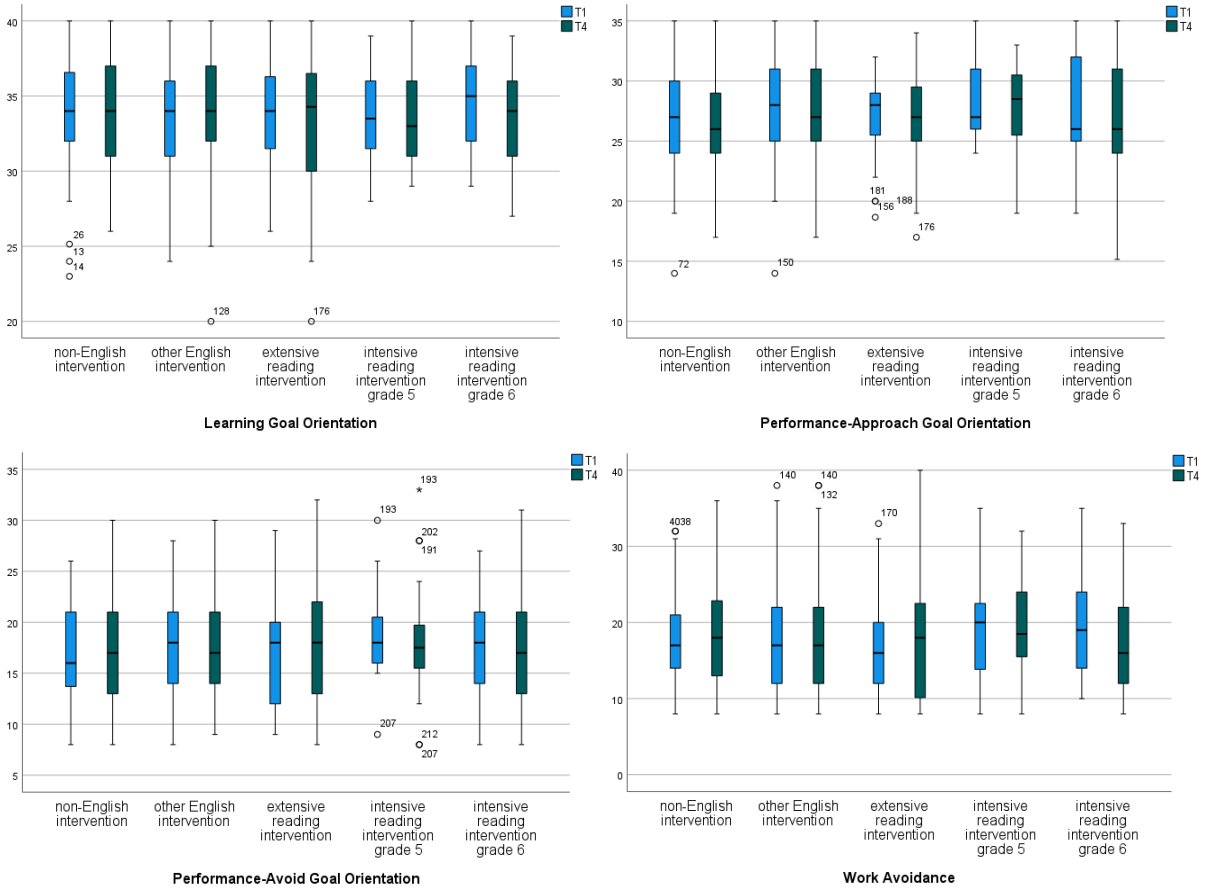


Figure 4.7 Goal Orientation Boxplots T1 vs. T4

4.5 Moderation Analyses

The previous analyses did not reveal any main effects of the type of reading intervention on the language test results, but various influences of covariates. Moreover, it was argued and hypothesized in chapter 2 that cognitive and motivational factors could influence the reading interventions differently and, therefore, function as moderators. If these moderation effects work in opposite directions, so called crossover effects occur, leading to non-significant effects of the independent variables, but to significant effects of the interaction term of the independent variable with the covariate on the dependent variable. Thus, additional moderation analyses are necessary to complete

the picture on the effects in this study. For a better reading, I will not include statistics here. All moderation analysis results are summarized in the tables in the appendix; significant statistics are highlighted in green. For an interpretation of the results, it is important to keep in mind that these are the corrected and relative point score gains (see description and example in the introduction to this chapter). I will demonstrate the interpretation of the diagrams and corrected cloze test gains again in detail for the first moderation.

Learning goal orientation moderated effects of intensive and extensive reading interventions on the direct cloze test results: Whereas students with a high learning goal orientation have a higher competence gain in their cloze tests results in the extensive reading interventions, the cloze test results of the students with a lower learning goal orientation did not increase as much (see figure 4.8). These effects are opposite for the intensive reading interventions.

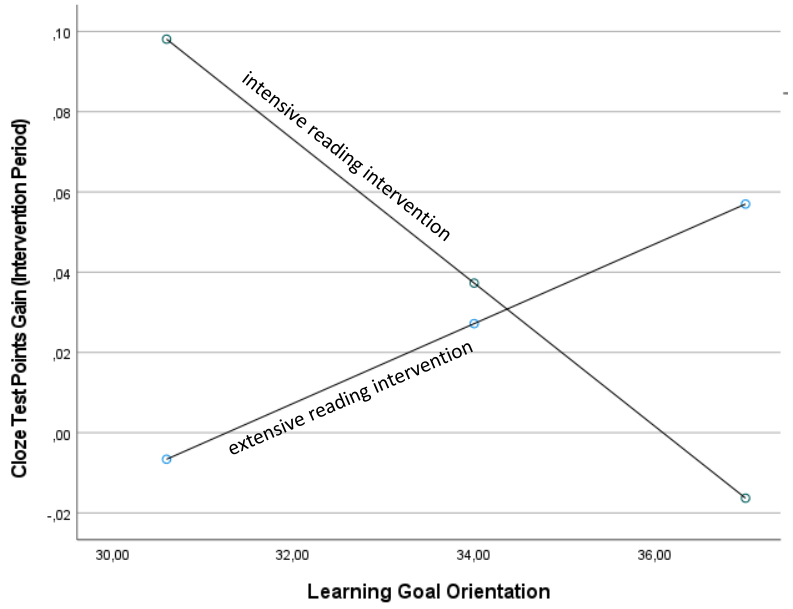


Figure 4.8 Interaction Effects of Learning Goal Orientation

These gains always need to be interpreted in relation to the non-English reading intervention groups (base line). In this case, students with a low learning goal orientation achieved an average 10 percent points more in their cloze test directly before and after participating in an intensive reading intervention than students in the non-English reading interventions. Students with a high learning goal orientation performed on average 2 percent points lower in the posttest than students in the non-English reading interventions when participating in the intensive reading interventions. Students who participated in an extensive reading intervention performed on average about 0.5 percent points less in the posttest than students in the non-English intervention groups when their learning goal orientation is low, but almost 6 percent points higher when their learning goal orientation is high. The

relationship to the non-English intervention groups derives from the use of the corrected points score gains and is important, because the non-English interventions are the control groups in this design. The corrected scores are therefore independent of normal learning effects and improvements that all students might have during the school year.

Moderation analyses also showed moderating effects of performance-approach goal orientation on the cloze tests results. Students in the non-English interventions with a low performance-approach goal orientation did not increase their results between the direct pre- and posttest, as much as all students in the non-English interventions, independently of their goal orientation scores. Yet, students of the non-English intervention groups with a high performance-approach goal orientation increased their results by 4 percent points. These results are contrary for the reading interventions (see figure 4.9).

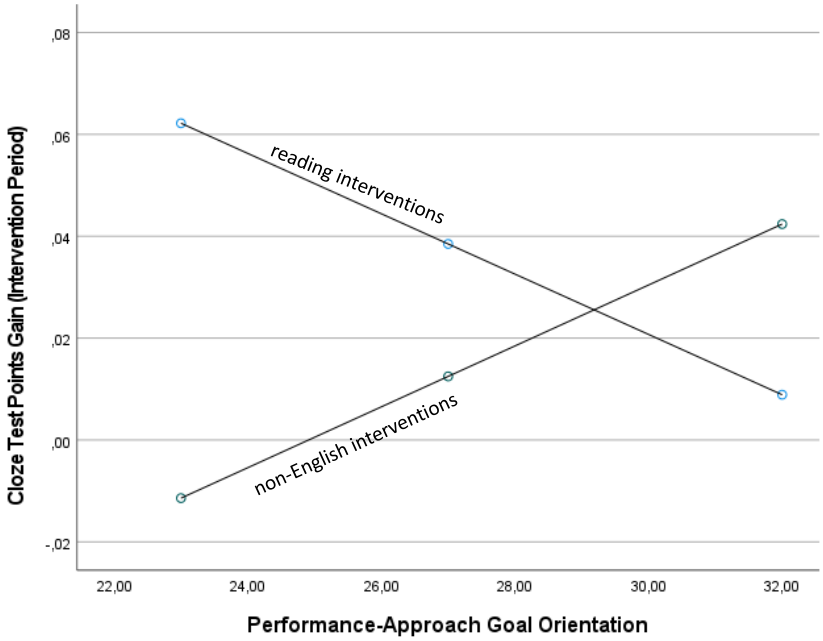


Figure 4.9 Interaction Effects of Performance-Approach Goal Orientation on Cloze Test Points

Interactions with performance-avoid goal orientations and intervention type did not significantly influence comprehension score gains, but work avoidance did. First, it moderated effects of the intervention group on the increase in cloze test results. Students in the non-English intervention groups had a higher increase in direct pre-/posttest points and the other way around when it was low. This effect was reversed for the reading interventions (see figure 4.10 left). This also shows when comparing other English interventions to the reading intervention groups, but only for low work avoidance scores (see figure 4.10 right).

The effect of the reading interventions on grade improvement was also significantly different to those of the non-English interventions when taking interactions with self-concept into account:

Whereas the grade improvement after the reading interventions was significantly higher than after other English and non-English interventions when absolute self-concept scores were low, these effects are reversed for high self-concept scores (see figure 4.11).

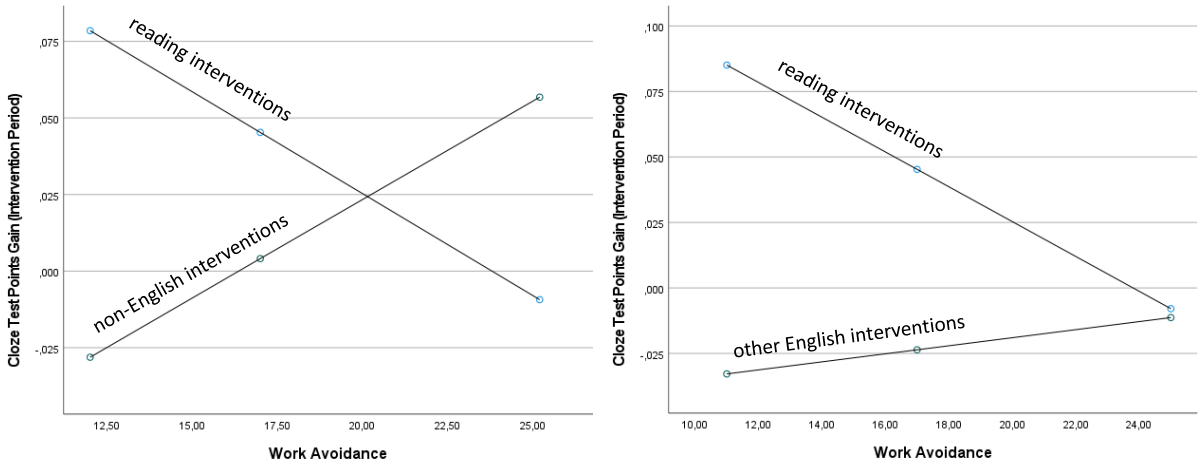


Figure 4.10 Interaction Effects of Work Avoidance

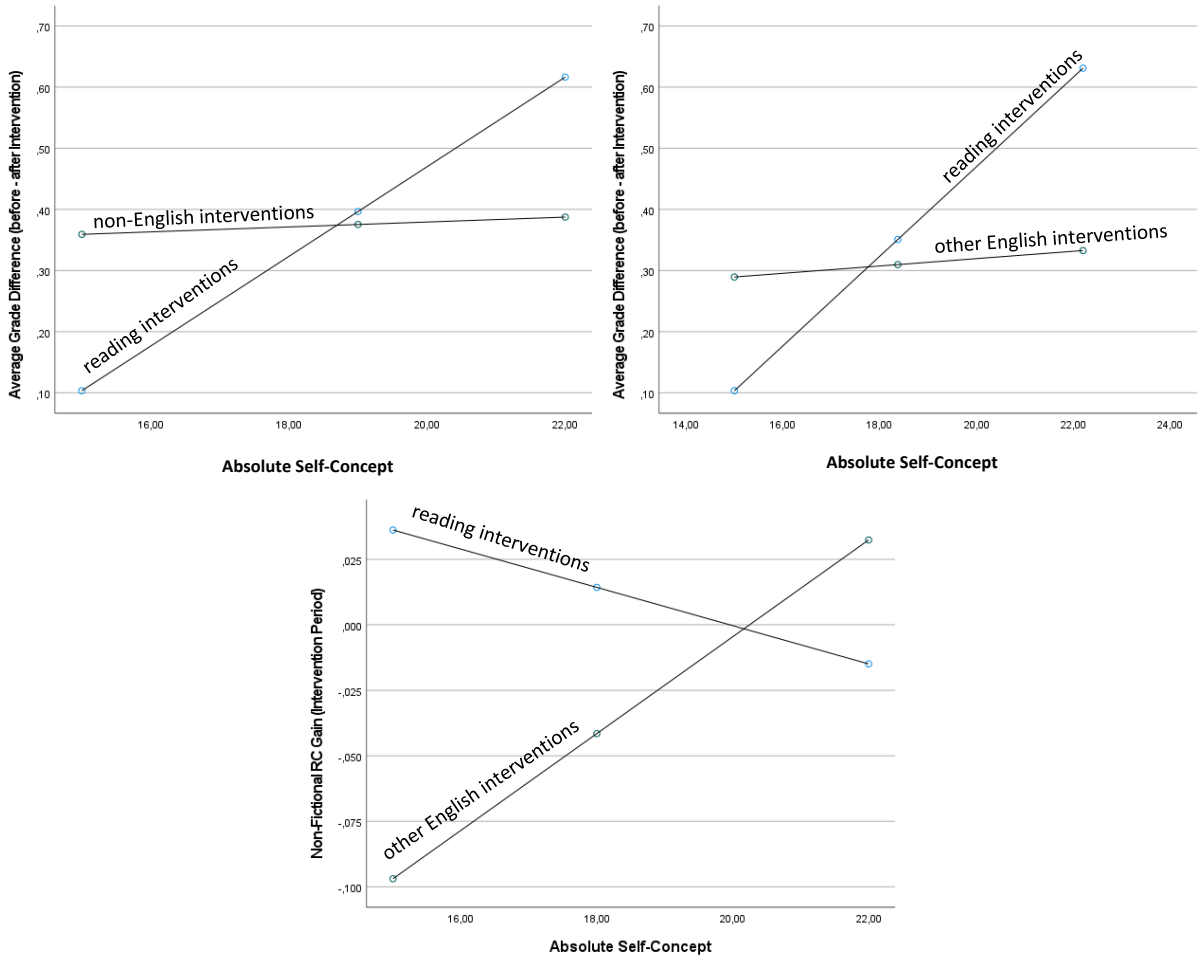


Figure 4.11 Interaction Effects of Absolute Self-Concept Scores

Moreover, absolute self-concept moderated effects on the non-fictional reading intervention points gain for students participating in other non-reading English interventions compared to the reading interventions (see figure 4.11). Students with a low self-concept had a lower point gain when they were in the other English interventions than when they had taken part in a reading intervention. Additionally, individual self-concept moderated effects on the non-fictional reading intervention points gain, in a similar way (see figure 4.12).

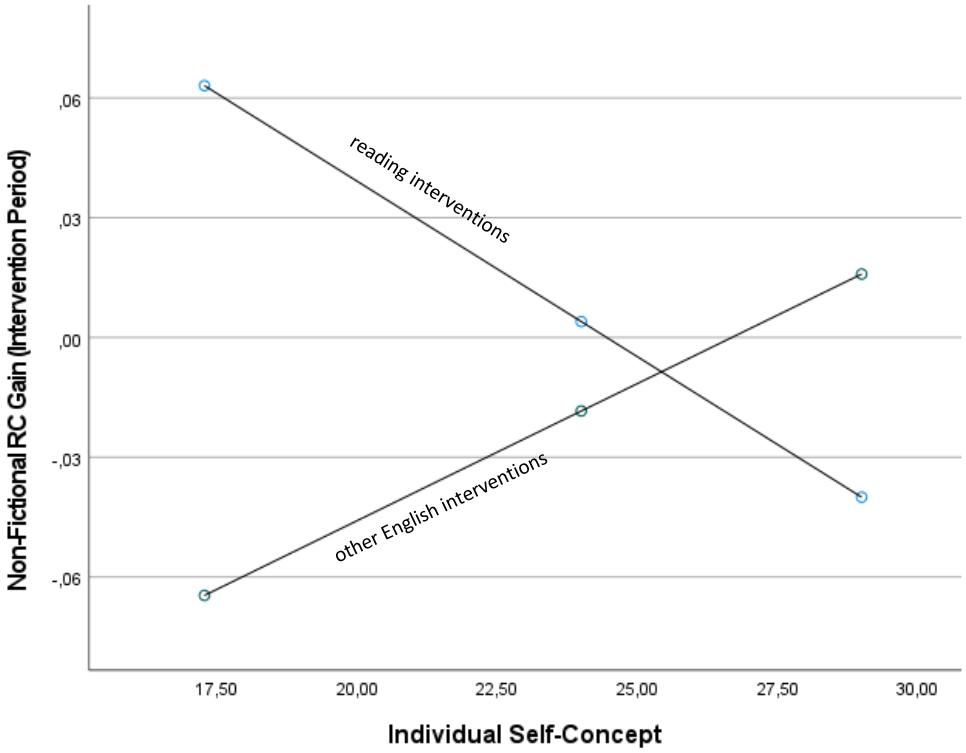


Figure 4.12 Interaction Effect of Individual Self-Concept Scores

The effects of intensive and extensive reading interventions on the fictional reading comprehension points gain, which are moderated by criterial self-concept, are interesting as well (figure 4.13): Students with low self-concept scores achieved a higher improvement (compared to the control groups) in fictional reading comprehension scores than students in the extensive reading interventions, who had lower competence gains than the control groups. For high self-concept scores, in contrast, the improvements of the students in the extensive reading interventions were higher than the control group, whereas the students in the intensive reading interventions improved about as much as the control groups did.

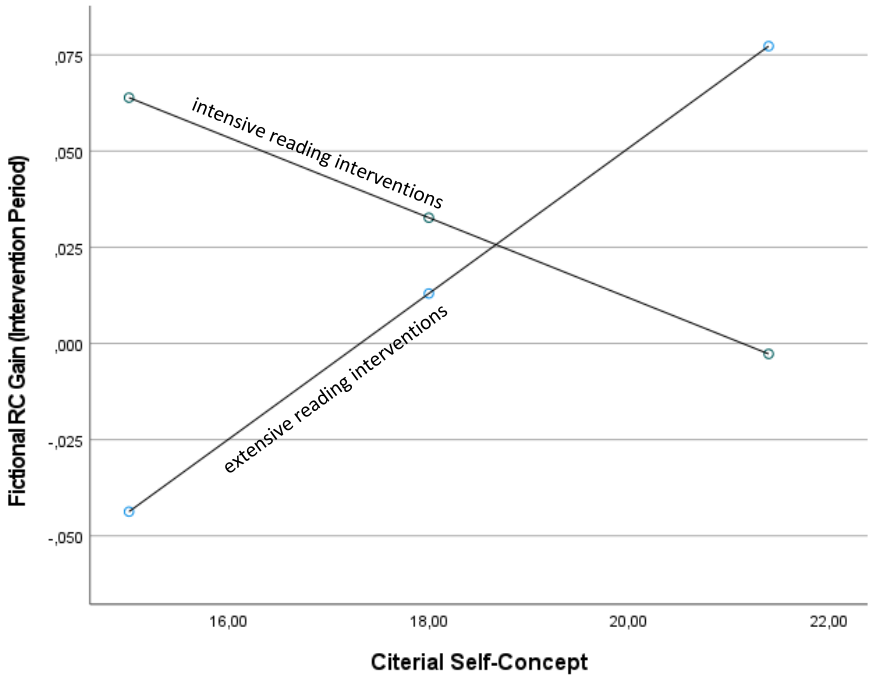


Figure 4.13 Interaction Effect of Self-concept Scores on Fictional Reading Comprehension Points

Not only psychological scales moderated intervention effects, cognitive effects showed significant interactions with the language competence gains as well. Moderation analyses of effects of reading behavior on non-fictional reading comprehension score gains revealed that students with a low private L1 reading behavior profited more from the extensive reading interventions than students with a high private reading behavior, who did not profit as much. The effects were vice versa for the intensive reading interventions: Compared to the control groups, the intensive reading interventions did not lead to a competence gain for students with a low private L1 reading behavior, but improved the competence of students with a high reading behavior (see figure 4.14 left).

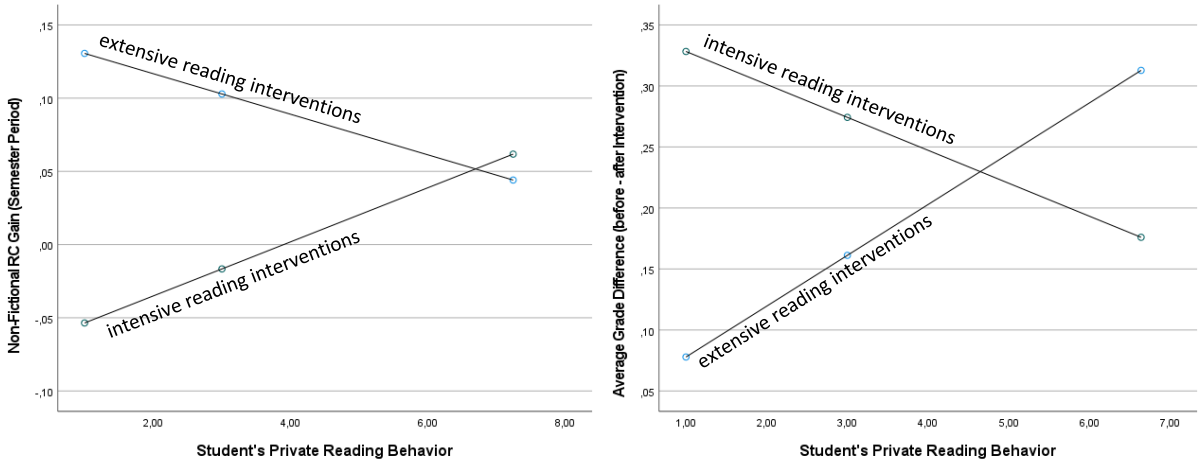


Figure 4.14 Interaction Effects of Student's Private Reading Behavior

Also, the effect of the reading interventions on the improvement of the English grades was significantly moderated by the students' private reading behavior. Due to the grade scale from 1 to 6 and the calculation of grade difference, student's grades with a high average grade difference got worse and with a low difference improved. Therefore, students with less private reading had lower grade improvements than the control groups when they were in the intensive reading interventions and higher grade improvements than the control groups in the extensive reading interventions. This was vice versa for students with high private reading (see figure 4.14 right).

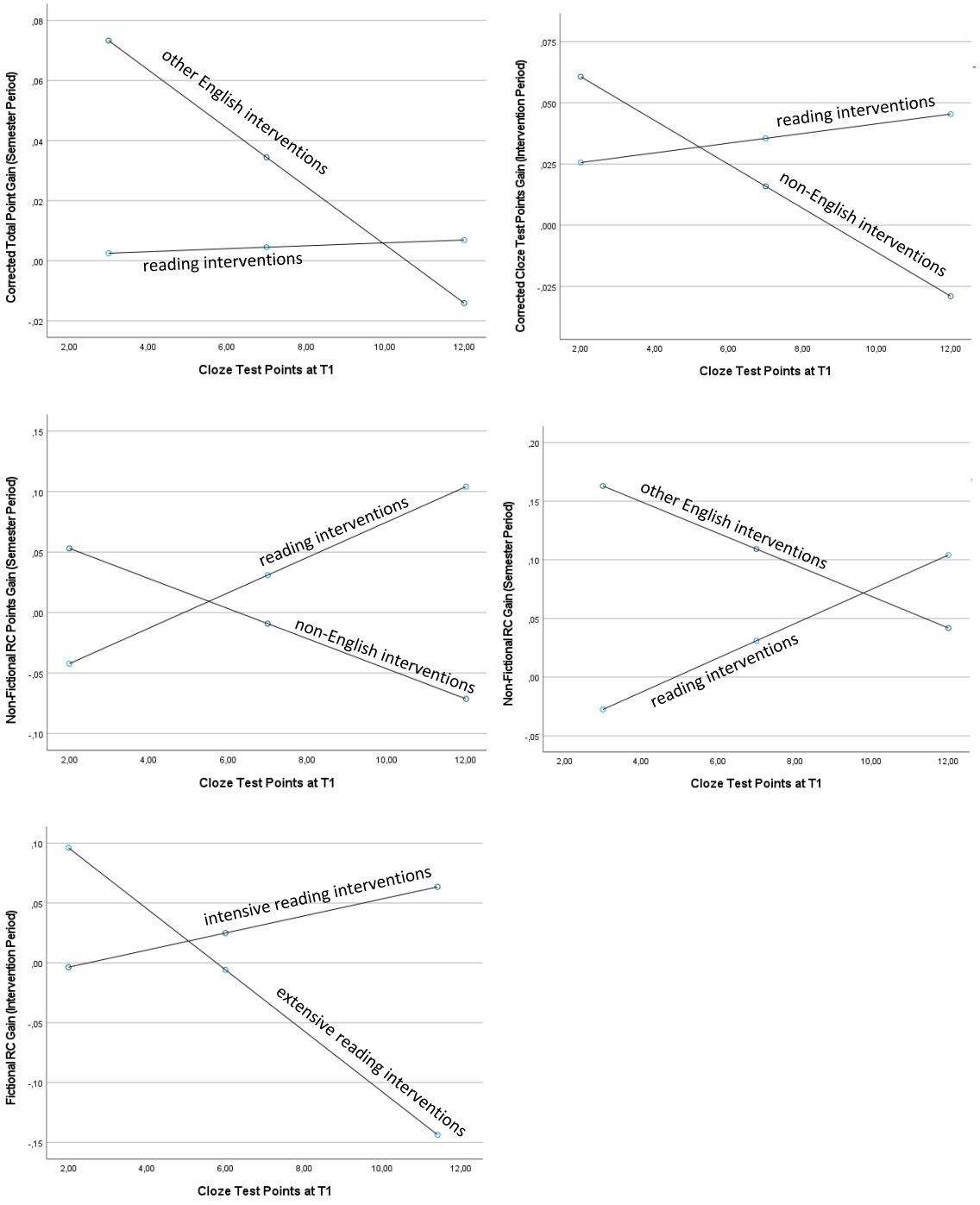


Figure 4.15 Interaction Effects of Student's Cloze Test Results at T1

Initial language competence before the interventions (T1) also moderated the competence gain of the students. Especially cloze test results in T1 moderated intervention effects (see figures 4.15). The results can be summarized as that students with high cloze test scores in T1 have higher competence score gains when they take part in the reading interventions, whereas students in other or non-English interventions have low or even negative score gains. These results are reversed for low cloze test results in T1.

The cloze test results in T1 also moderate the effect of the intensive and extensive reading interventions on fictional reading comprehension improvement (see figure 4.15 bottom). Students with low initial cloze test scores improve more than the control groups (baseline) in the extensive reading interventions, but fall below the baseline when they have high initial cloze test scores. This effect is reversed for the intensive reading interventions.

Non-fictional reading comprehension test scores at T1 also moderated the effect of the reading interventions on the fictional and non-fictional competence gain (see figure 4.16). Students with high initial non-fictional reading comprehension scores perform worse than the baseline (control groups) when they were in the reading intervention groups, but if their non-fictional reading score is low, reading interventions lead to higher competence gains than the baseline. When looking at the non-fictional reading competence gain, this result is slightly stronger for the intensive reading interventions than for the extensive reading interventions.

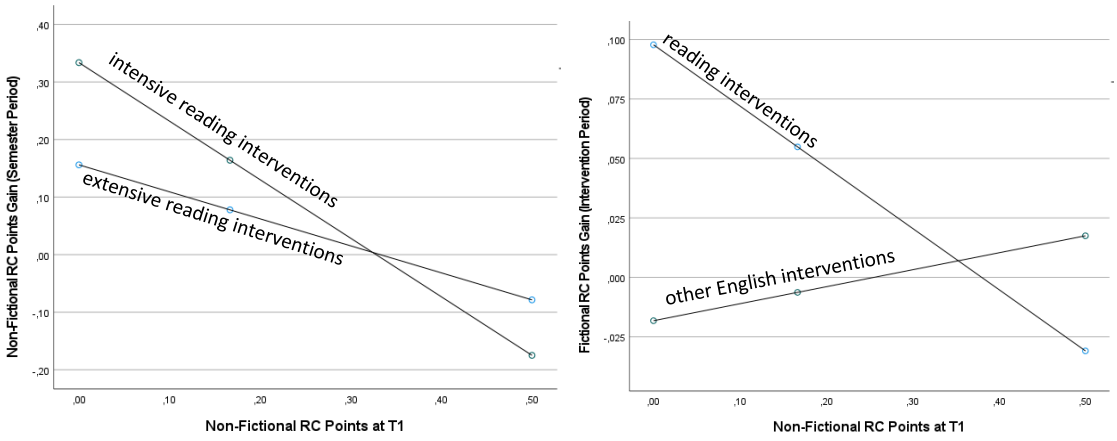


Figure 4.16 Interaction Effects of Student's Non-Fictional RC Results at T1

Only one moderation analysis is significant for the time spent learning English at primary school. It moderates the effect on the direct fictional reading comprehension gain score; diagram 4.17 depicts these results. The years spent learning English at primary school did not moderate the effects of the intensive reading interventions, but those of the extensive reading intervention: Students who only had three years of primary school English lessons did not profit as much from the extensive reading

interventions (lower competence score gains than control group) than those who had more than three years of English lessons (higher competence score gains than control group).

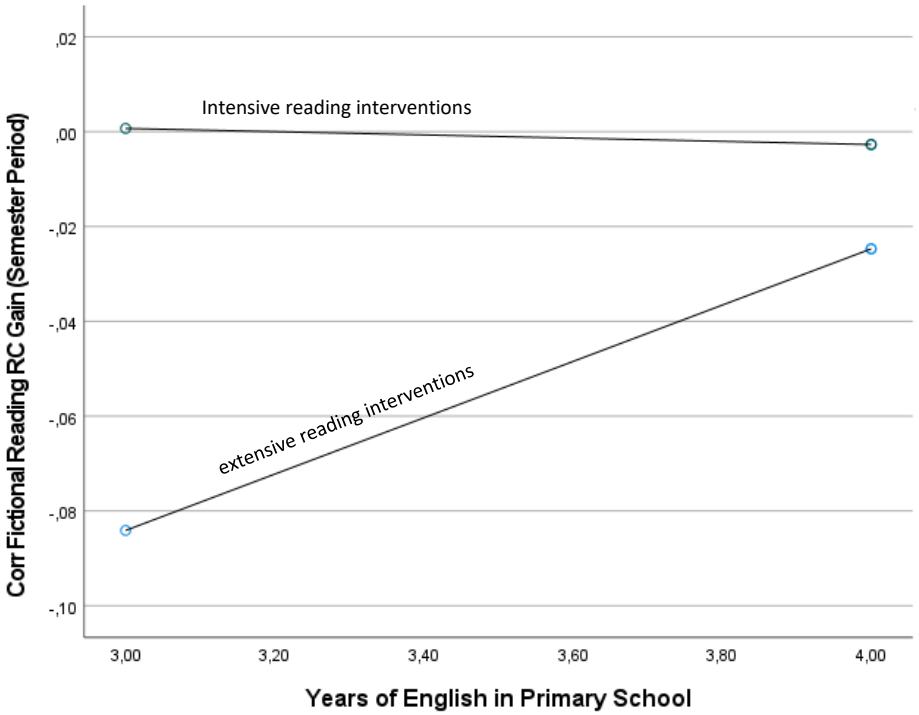


Figure 4.17 Interaction Effect of Years Learning English at Primary School on Fictional RC

4.6 Regression Analyses

The previous sections have presented results of significant direct and moderation effects of various variables on L2 language competence test gains. What is also of interest is which of these predictors are the most salient, i.e. explain most of the variance of the outcome variables and how much percent is explained. To answer this question, regression analyses were performed for each outcome variable (language test gains for direct and delayed pre-/posttests). Since previous results have shown that these effects differ between the intervention groups, these analyses were run separately for each of these four groups and are presented by intervention group in the following sections.

When running regression analyses, there are various ways[24] of entering the predictors into the regression model. Since the aim of these analyses is not to verify a certain model and enter all variables

[24] Field (2018) argues for entering the variables hierarchically according to a fixed order deducted by hypotheses and previous research and the expected effects or by forced entry, entering all variables at once. He is rather distant about stepwise procedures: "The stepwise fashion bases decisions about the order in which predictors enter the model on a purely mathematical criterion." (p. 398).

hierarchically according to hypothesized effects, but just to deliver some additional information on effect sizes, I opted for the *forward* method[25]. In this method the predictor that best predicts the outcome variable, i.e. that it has the highest correlation with the outcome variable, is entered first. This procedure is continued as long as the predictors significantly explain more variance and, therefore, deliver a better fit of the model. It is important to keep in mind that the effect sizes can vary according to the order that the predictors are entered in the model, therefore, I also tested other procedures that did not strongly differ in their results.[26] The following sections and tables present the results of these regression analyses.

4.6.1 Non-English Interventions

Regression analyses results for the intervention groups that did not have English classes show that in most cases the language test results in T1 were the main predictors for competence gain in the corresponding test part: The cloze test results in T1 explain 6.1% of the variance in competence gains in the direct pre-/posttests and 26% in the delayed pre-/posttests. The same applies to the fictional and non-fictional reading comprehension points in the first test (T1) and the competence gains in the direct (9.5% for non-fiction and 11.3% for fiction) and delayed (37.5% for non-fiction and 40.4% for fiction) pre-/posttest of the corresponding test types. For the preposition test, these results only show for the delayed pre-/posttests and only explain 9.5% of the variance. The effect is negative in all cases, meaning that students who initially have low scores on one of the test parts have higher improvements (higher competence gain) in those test parts.

Cloze test results in the first test (T1) proved to be a significant predictor for other competence gains as well: in the delayed non-fictional reading comprehension test (additional explanation of 6.5% of the variance), preposition test (additional explanation of 17.4% of the variance) and fictional reading comprehension test (additional explanation of 17.1% of the variance). Differing from the effects of the corresponding test results in T1 and their gains presented in the first paragraph, the effects of the cloze test results in T1 are all positive, meaning that students with high points in the first cloze test have higher competence gains in the reading comprehension and preposition tests at the end of the

[25] Field (2018) is critical about these models due to the danger of overfitting and having too many variables that do not make a contribution or underfitting and overlooking important predictors. Since main effect and moderations effects have already been analyzed in the previous sections and significant predictors have already been identified, this is not a problem here.

[26] I ran hierarchical analyses in a fixed fashion first, entering biographical, then psychological and then language proficiency at T1 with a forced entry. The significant predictors were the same. I opted for this version first, because the different models (and predictors) would have been the same across all the groups and therefore the R^2-change from model to model could have been compared directly. I decided against this presentation of the results though, because the presentation and structure of so many variables that were entered and models tested that did not turn out to be significant did not seem effective to me. Moreover, since different predictors were significant for different interventions, the R^2-change between the models was based on different predictors and therefore not comparable anymore. Moreover, entering the predictors in the same fashion for each intervention group would also not meet the picture that developed in the previous analyses that predictors seem to influence the outcome variables very differently in the different groups.

semester than students with lower points. Further test results in T1 were also significant predictors and can be found in table 4.1.

A further significant explanation of variance is delivered by self-concept scores (5.8% of the direct and 13.9% of the delayed pre-/posttests and 3.6% of the non-fictional reading comprehension). Whereas this effect was positive for the social and individual self-concept scores, it was negative for the criterial self-concept scores: On the one hand, students who believe that they are better in English when comparing themselves to others or to previous situations, show higher improvement in the cloze tests. On the other hand, students who have a high criterial English self-concept, measured as feeling to be able to meet the requirements at school, have a lower competence gain in the cloze tests than students who feel that they have trouble meeting the requirements.

Moreover, German L1 explains 6.5% of the variance of the delayed non-fictional reading comprehension gains. The effect is negative and speaking another language as an L1 was coded with 1, not speaking it with 0. Therefore, students who speak another language at home have lower competence gains in the non-fictional reading comprehension tests than students whose L1 is German. All these results and further statistics are summarized in the following table 4.1.

Table 4.1 Regression Analyses Results for Non-English Interventions

Variable	Model 1			Model 2			Model 3			Model 4			Model 5		
	b	*SE B*	*β*	*b*	*SE B*	*β*	*b*	*SE B*	*β*	*b*	*SE B*	*β*	*b*	*SE B*	*β*
Corrected Total Points Gain (direct = intervention period)															
Constant	.08	.03													
Fictional RC T1	-.13	-.26	-.28												
	R^2=.076; corrR^2=.062 F(1;70)=5.73*														
Corrected Total Points Gain (delayed = semester period)															
Constant	.09	.03													
Fictional RC T1	-.22	.06	-.41												
	R^2=.169; corrR^2=.157 F(1;70)=14.23**														
Corrected Cloze Test Points Gain (direct = intervention period)															
Constant	.07	.03		.06	.03		-.12	.09							
Cloze Test T1	-.21	.10	-.25	-.39	.13	.46	-.50	.14	-.60						
Non-Fictional RC T1				.19	.09	.32	.23	.09	.37						
Individual Self-Concept							.01	.01	.27						
	R^2=.061; corrR^2=.048 *ΔR^2=.061 ;F(1;70)=4.57**			R^2=.116; corrR^2=.090 *ΔR^2=.054 ;F(1;69)=4.25**			R^2=.174; corrR^2=.137 *ΔR^2=.058 ;F(1;69)=4.80**								
Corrected Cloze Test Points Gain (delayed = semester period)															
Constant	.13	.03		.09	.03		.05	.04		.23	.09		-.13	.09	
Cloze Test T1	-.44	.09	-.51	-.56	.09	-.65	-.66	.10	-.77	-.60	.11	-.70	-.60	.10	-.71
Preposition Test T1				.34	.12	.32	.33	.12	.30	.32	.11	.30	.35	.10	.32
Fictional TC T1							.16	.08	.23	.20	.08	.28	.17	.08	.24
Criterial Self-Concept										-.01	.01	-.22	-.03	.01	-.55
Social Self-Concept													.02	.01	.47
	R^2=.260; corrR^2=.249 *ΔR^2=.260;F(1;70)=24.59***			R^2=.340; corrR^2=.321 *ΔR^2=.080;F(1;69)=17.77***			R^2=.377; corrR^2=.349 *ΔR^2=.037;F(1;68)=13.70***			R^2=.413; corrR^2=.378 *ΔR^2=.037;F(1;67)=11.80***			R^2=.515; corrR^2=.479 *ΔR^2=.102;F(1;66)=14.04***		
Corrected Non-Fictional Reading Comprehension Points Gain (direct = intervention period)															
Constant	.16	.05													
Non-Fictional RC T1	-.30	.11	-.31												
	R^2=.095; corrR^2=.083 F(1;70)=7.39**														
Corrected Non-Fictional Reading Comprehension Points Gain (delayed = semester period)															
Constant	.21	.05		.14	.05		.17	.05		.50	.16				
Non-Fictional RC T1	-.70	.11	-.61	-.97	.14	-.85	-.98	.14	-.85	-1.02	.13	-.89			
Cloze Test T1				.54	.19	.35	.58	.19	.37	.76	.20	.49			
German L1							-.15	.07	-.21	-.15	.06	-.20			
Criterial Self-Concept										-.02	.01	-.21			
	R^2=.375; corrR^2=.366 *ΔR^2=.375;F(1;70)=24.59***			R^2=.440; corrR^2=.424 *ΔR^2=.065;F(1;69)=27.11***			R^2=.482; corrR^2=.459 *ΔR^2=.042;F(1;68)=21.10***			R^2=.518; corrR^2=.489 *ΔR^2=.036;F(1;67)=17.99***					
Corrected Preposition Test Points Gain (direct = intervention period)															
no significant predictors															
Corrected Preposition Test Points Gain (delayed = semester period)															
Constant	.09	.04		.02	.04										
Preposition Test T1	-.38	.14	-.31	-.65	.14	-.52									
Cloze Test T1				.46	.12	.47									
	R^2=.094; corrR^2=.08 *ΔR^2=.094;F(1;70)=7.23***			R^2=.268; corrR^2=.247 *ΔR^2=.174;F(1;69)=12.62***											
Corrected Fictional Reading Comprehension Points Gain (direct = intervention period)															
Constant	.13	.05													
Fictional RC T1	.31	.11	-.34												
	R^2=.113; corrR^2=.100 F(1;70)=8.88**														
Corrected Fictional Reading Comprehension Points Gain (delayed = semester period)															
Constant	.32	.05		.25	.05										
Fictional RC T1	-.76	.11	-.64	-1.08	.11	-.90									
Cloze Test T1				.70	.13	.49									
	R^2=.404; corrR^2=.395 *ΔR^2=.404;F(1;70)=47.45***			R^2=.575; corrR^2=.563 *ΔR^2=.171;F(1;69)=46.66***											
Grade Difference before-after the Semester															
Constant	.09	.03		.10	.03										
Fictional RC T1	-.23	.06	-.42	-.16	.07	-.29									
Non-Fictional RC T1				-.12	.06	-.27									
	R^2=.174; corrR^2=.163 *ΔR^2=.174;F(1;71)=14.98***			R^2=.229; corrR^2=.206 *ΔR^2=.054;F(1;70)=10.36**											

*p < .05. **p < .01.

4.6.2 Other English Interventions

Table 4.2 Regression Analyses Results for other (non-reading) English Interventions

	Model 1			Model 2		
Variable	*b*	*SE B*	*β*	*b*	*SE B*	*β*
Corrected Total Points Gain (direct = intervention period)						
Constant	.03	.02				
Non-Fictional RC T1	-.13	.06	-.24			
	R^2=.057; corrR^2=.043 F(1;69)=4.18*					
Corrected Total Points Gain (delayed = semester period)						
Constant	.13	.04		.18	.04	
Fictional RC T1	-.23	.08	-.35	-.27	.08	-.42
German L1				-.08	.03	-.28
	R^2=.121; corrR^2=.107 *ΔR^2=.121; F(1;64)=8.80***			R^2=.192; corrR^2=.167 *ΔR^2=.071; F(1;63)=7.50***		
Corrected Cloze Test Points Gain (direct = intervention period)						
no significant predictors						
Corrected Cloze Test Points Gain (delayed = semester period)						
Constant	.13	.04		.18	.04	
Cloze Test T1	-.38	.11	-.39	-.44	.11	-.45
German L1				-.11	.04	-.30
	R^2=.151; corrR^2=.137 *ΔR^2=.151;F(1;64)=11.34***			R^2=.236; corrR^2=.212 *ΔR^2=.085;F(1;63)=9.73***		
Corrected Non-Fictional RC Points Gain (direct = intervention period)						
Constant	.13	.05		-.43	.18	
Non-Fictional RC T1	-.53	.14	-.41	-.65	.14	-.50
Absolute Self-Concept				.03	.01	.34
	R^2=.171; corrR^2=.15 *ΔR^2=.171;F(1;69)=14.22***			R^2=.281; corrR^2=.260 *ΔR^2=.110;F(1;68)=13.28***		
Corrected Non-Fictional RC Points Gain (delayed = semester period)						
Constant	.27	.05		.32	.05	
Non-Fictional RC T1	-.64	.14	-.50	-.68	.14	-.53
German L1				-.14	.06	-.23
	R^2=.250; corrR^2=.238 *ΔR^2=.250;F(1;64)=21.36***			R^2=.302; corrR^2=.280 *ΔR^2=.052;F(1;64)=21.36***		
Corrected Preposition Test Points Gain (direct = intervention period)						
Constant	.07	.04		.22	.08	
Preposition Test T1	-.31	.13	-.27	-.30	.13	-.26
Primary School English				-.04	.02	-.23
	R^2=.074; corrR^2=.061 *ΔR^2=.074;F(1;69)=5.52***			R^2=.129; corrR^2=.103 *ΔR^2=.055;F(1;68)=5.03***		
Corrected Preposition Test Points Gain (delayed = semester period)						
no significant predictors						
Corrected Fictional RC Points Gain (direct = intervention period)						
Constant	-.07	.03				
German Reading	-.01	.004	-.26			
	R^2=.068; corrR^2=.05 F(1;69)=5.01**					
Corrected Fictional RC Points Gain (delayed = semester period)						
Constant	.24	.05		.22	.05	
Fictional RC T1	-.58	.11	-.54	-.69	.12	-.64
Non-Fictional RC T1				.26	.11	.27
	R^2=.296; corrR^2=.285 *ΔR^2=.296;F(1;64)=26.97***			R^2=.357; corrR^2=.337 *ΔR^2=.061;F(1;63)=17.50***		
Grade Difference before – after Semester						
Constant	.15	.04				
Fictional RC T1	-.29	.08	-.41			
	R^2=.171; corrR^2=.159 F(1;69)=14.24**					

*p < .05. **p < .01.

Comparable to the non-English classes, the main predictors explaining the largest amount of variance in competence gain of the students in the other English classes were the results of the corresponding tests at the beginning of the intervention. I will, therefore, not go into detail on these; they can be found in table 4.2 on the left. There are also some effects of other language tests on the total points and fictional reading comprehension (RC) points gain (see table 4.2 for results).

What is interesting is that German L1 seems to play a larger role in this group than in the non-English classes, explaining an additional 7.1% of the variance in the delayed total points gain, 8.5% in the delayed cloze test gain and 5.2% in the delayed fictional RC gain. The direction of the effect is also negative, as in the non-English intervention group, meaning that students with an L1 other than German had lower competence gains in these tests.

German L1 reading is the strongest – and only significant – predictor of the direct fictional RC gain, explaining 6.8% of the variance. Moreover, this effect is negative – the more students read in their L1, the lower their competence gain in the fictional reading RC. This effect does not show for the delayed pre-/posttest.

Primary school English explained an additional 5.5% of the variance in the direct preposition points gain and absolute self-concept 11% in the direct non-fictional RC gain. However, these results do not show for the delayed pre-/posttests.

4.6.3 Extensive Reading Interventions

Table 4.3 Regression Analyses Results for Extensive Reading Interventions

	Model 1			Model 2		
Variable	*b*	*SE B*	*β*	*b*	*SE B*	*β*
Corrected Total Points Gain (direct = intervention period)						
Constant	.06	.02				
Fictional RC T1	-.14	.06	-.42			
	R^2=.173; corrR^2=.142 $F(1;26)=5.45^{**}$					
Corrected Total Points Gain (delayed = semester period)						
Constant	.13	.04		.21	.05	
Fictional RC T1	-.31	.08	-.68	-.30	.07	-.65
Work Avoidance				-.01	.003	-.33
	R^2=.461; corrR^2=.434 $\Delta R^2=.461; F(1;20)=17.10^{**}$			R^2=.566; corrR^2=.521 $\Delta R^2=.105; F(1;19)=12.41^{**}$		
Corrected Cloze Test Points Gain (direct = intervention period)						
Constant	.14	.05				
Work Avoidance	-.01	.003	-.42			
	R^2=..174; corrR^2=.142 $F(1;26)=5.48^{**}$					
Corrected Cloze Test Points Gain (delayed = semester period)						
Constant	-1.05	.37				
Learning GO	.03	.01	.55			
	R^2=..301; corrR^2=.266 $F(1;20)=8.61^{**}$					
Corrected Non-Fictional RC Points Gain (direct = intervention period)						
no significant predictors						
Corrected Non-Fictional RC Points Gain (delayed = semester period)						
no significant predictors						
Corrected Preposition Test Points Gain (direct = intervention period)						
no significant predictors						
Corrected Preposition Test Points Gain (delayed = semester period)						
Constant	.12	.04				
Preposition Test T1	-.69	.16	-.70			
	R^2=.490; corrR^2=.464 $F(1;20)=19.20^{**}$					
Corrected Fictional RC Points Gain (direct = intervention period)						
Constant	-.44	.14		-.34	.14	
Primary School English	.12	.04	.54	.12	.03	.54
Cloze Test T1				-.54	.19	-.36
	R^2=.287; corrR^2=.260 $\Delta R^2=.287; F(1;26)=10.46^{**}$			R^2=.415; corrR^2=.368 $\Delta R^2=.128; F(1;25)=8.85^{**}$		
Corrected Fictional RC Points Gain (delayed = semester period)						
Constant	.36	.05		.31	.05	
Fictional RC T1	-1.01	.11	-.89	-1.14	.11	-1.01
Cloze Test T1				.47	.16	.28
	R^2=.798; corrR^2=.788 $\Delta R^2=.798; F(1;20)=78.94^{**}$			R^2=.861; corrR^2=.846 $\Delta R^2=.063; F(1;24)=58.68^{**}$		
Grade Difference before – after Semester						
Constant	.11	.04		.24	.06	
Fictional RC T1	-.31	.09	-.57	-.31	.08	-.58
Work Avoidance				-.01	.003	-.40
	R^2=.324; corrR^2=.295 $\Delta R^2=.324; F(1;24)=11.48^{**}$			R^2=.483; corrR^2=.438 $\Delta R^2=.160; F(1;24)=10.75^{*}$		

$^{*}p < .05$. $^{**}p < .01$.

Other than in the previous two groups, the corresponding language test results in T1 do not turn out to be the strongest predictor in most of the regression analyses run for the extensive reading intervention classes. Yet, this was still the case for the delayed preposition and fictional RC test gains (see table 4.3). What is baffling are the cloze test results in T1 as a significant predictor for fictional RC gain, because the effect is negative for the direct pre-/posttest but positive for the delayed one.

Contrary to the other two previously discussed classes, goal orientation, especially work avoidance, is a significant predictor in the extensive reading classes, explaining an additional 10.5% of the variance in the delayed total point gains and 17.5% of the variance in the direct cloze test score gains. Students with a high work avoidance have lower competence gains. Learning goal orientation, in contrast, has a positive effect on student's delayed cloze test score gains.

Moreover, the years students had learned English in primary school turns out to be the strongest predictor of the direct fictional RC gain, explaining 28.7% of the variance: The more years students had English at primary school, the higher their competence gain in the direct fictional RC.

4.6.4 Intensive Reading Interventions

Table 4.4 Regression Analyses Results for Intensive Reading Interventions

Variable	Model 1			Model 2			Model 3		
	b	*SE B*	*β*	*b*	*SE B*	*β*	*b*	*SE B*	*β*
Corrected Total Points Gain (direct = intervention period)									
Constant	.09	.03							
Non-Fictional RC T1	-.20	.07	-.38						
	R^2=.142; corrR^2=.122 F(1;45)=7.42**								
Corrected Total Points Gain (delayed = semester period)									
Constant	.11	.03		.24	.07				
Fictional RC T1	-.33	.09	-.50	-.38	.09	-.57			
Work Avoidance				-.01	.003	-.27			
	R^2=.247; corrR^2=.230 ΔR^2=.247 ;F(1;46)=15.06*			R^2=.312; corrR^2=.282 ΔR^2=.066 ;F(1;45)=4.29*					
Corrected Cloze Test Points Gain (direct = intervention period)									
Constant	.58	.23							
Learning GO	-.02	.01	-.34						
	R^2=.113; corrR^2=.093 F(1;45)=5.74**								
Corrected Cloze Test Points Gain (delayed = semester period)									
Constant	.12	.04							
Cloze Test T1	-.39	.12	-.45						
	R^2=.198; corrR^2=.181 F(1;46)=11.37**								
Corrected Non-Fictional Reading Comprehension Points Gain (direct = intervention period)									
Constant	.16	.05							
Non-Fictional RC T1	-.42	.14	-.40						
	R^2=.159; corrR^2=.141 F(1;45)=8.54**								
Corrected Non-Fictional Reading Comprehension Points Gain (delayed = semester period)									
Constant	.30	.07		.13	.08		-.03	.08	
Non-Fictional RC T1	-.95	.18	-.61	-1.07	.17	-.69	-1.02	.15	-.66
Cloze Test T1				.72	.21	.38	.87	.19	.46
German Reading							.02	.0	.35
	R^2=.372; corrR^2=.359 ΔR^2=.372;F(1;46)=27.28**			R^2=.508; corrR^2=.486 ΔR^2=.136;F(1;45)=23.26**			R^2=.617; corrR^2=.591 ΔR^2=.109;F(1;44)=23.64**		
Corrected Preposition Test Points Gain (direct = intervention period)									
no significant predictors									
Corrected Preposition Test Points Gain (delayed = semester period)									
Constant	.11	.04		.01	.04		-.16	.08	
Preposition Test T1	-.47	.15	-.41	-.84	.15	-.73	-.87	.15	-.76
Cloze Test T1				.61	.13	.60	.60	.13	.59
Primary School English							.05	.02	.25
	R^2=.171; corrR^2=.153 ΔR^2=.171;F(1;46)=9.49**			R^2=.430; corrR^2=.405 ΔR^2=.259;F(1;45)=17.01**			R^2=.492; corrR^2=.457 ΔR^2=.061;F(1;45)=14.18**		
Corrected Fictional Reading Comprehension Points Gain (direct = intervention period)									
no significant predictors									
Corrected Fictional Reading Comprehension Points Gain (delayed = semester period)									
Constant	.27	.07		.20	.06				
Fictional RC T1	-.62	.13	-.57	-.96	.15	-.88			
Cloze Test T1				.74	.20	.50			
	R^2=.320; corrR^2=.305 ΔR^2=.320;F(1;46)=21.63**			R^2=.473; corrR^2=.450 ΔR^2=.154;F(1;45)=13.12**					
Grade Difference before – after Semester									
Constant	.11	.03		.24	.07				
Non-Fictional RC T1	-.35	.08	-.53	-.39	.08	-.59			
Work Avoidance				-.01	.003	-.26			
	R^2=.278; corrR^2=.264 ΔR^2=.278;F(1;51)=19.62**			R^2=.339; corrR^2=.313 ΔR^2=.062;F(1;50)=12.84**					

*p < .05. **p < .01.

In line with the results in the previous groups, reading test results in T1, again, are the strongest predictors of point gains in the corresponding delayed language tests. Also, the cloze test result in T1 is a significant positive predictor of a high gain in other language tests (see table 4.4 on the left for detailed results).

Just as in the extensive reading intervention group, work avoidance is a negative predictor of the delayed total points gain. More interesting is the fact that learning goal orientation is the strongest predictor of the cloze test points gain as well, explaining 11.3% of the variance, but in this case the effect is negative.

Further predictors and the amount of explained variance are summarized in table 4.4 on the left.

4.6.5 Summarizing Regression Analyses Results

All in all, the points a student achieves in the first test seems to be the strongest predictor of the increase of points from first to last test. This is not surprising: Students who have fewer points in the first test have the largest potential to improve and seem to do so.

Moreover, the cloze test points in T1 seem to predict the improvements students make in the other language tests (prepositions and fictional and non-fictional reading comprehension). The direction of the effect is different to the cloze test points gain: The higher the results of the cloze test are, the higher the improvement in the other language areas. This effect also seems to be relatively stable across the different groups.

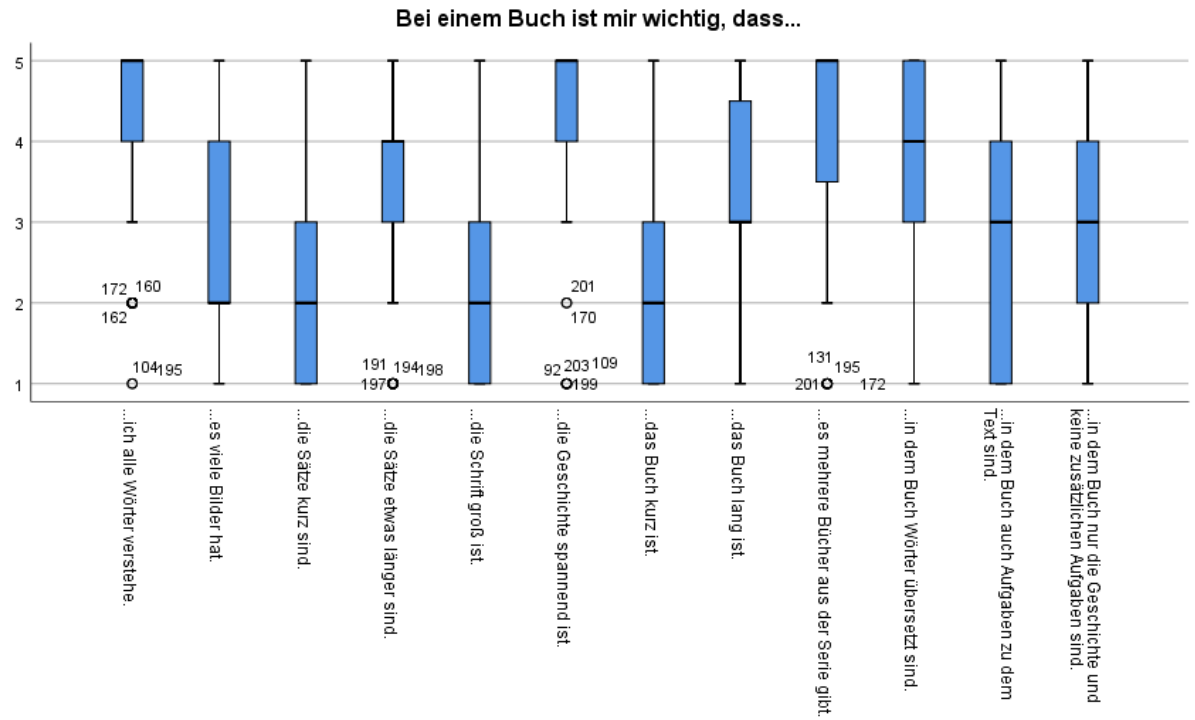
English at primary school is only a significant predictor in the English classes, but not in the non-English classes. Additionally, the direction of the effect is different: In the non-reading classes the effect is negative – students who had learned English for a shorter amount of time show more improvement - whereas it is positive in the reading classes – student who had learned English for a longer time show a higher improvement. These results seem logical because students who have a better knowledge of English can profit more from the reading interventions. However, students who have learned English for a shorter amount of time could feel overwhelmed.

Moreover, work avoidance is only a significant predictor in reading classes. Students with a low work avoidance profit less. The results for learning goal orientation differ for the reading intervention types – students in the extensive reading intervention seem to profit from a high learning goal orientation, whereas this effect is reversed for the intensive reading intervention.

All in all, the predictors explain a respectable amount of variance. In some cases, the significant predictors together explain more than 50% of the variance of the outcome variable. Considering that this study was not carried out in a totally controlled experimental setting, but at school in the field with many influences leading to possible variance, this is quite a high number. The explained variance is higher for the delayed pre-/posttest than for the direct pre-posttest results.

4.7 Reading Preferences

4.7.1 Questionnaire



1: stimmt gar nicht; 2: stimmt eher nicht; 3: weder noch; 4: stimmt eher; 5: stimmt genau

Figure 4.18 Preferred Book Features

To explore students' preferences when choosing books, students were asked what was important for them when making that choice. The answers are presented in figure 4.18 and are phrased in the original German phrasing, but will be discussed in English in this paragraph. Most important to students with an average of 5 from 5 points was that they understand all the words in a book (Q1), that the story is gripping (Q6) and that there are more books from the series (Q9). Most of the students also want longer sentences (Q4 vs. Q3 that the sentences are shorter), longer books (Q8 vs. Q7 that the books are shorter) and a translation of the words (Q10). Answers varied concerning the importance of pictures (Q2) and tasks for text comprehension (Q11 and Q12 that there are no tasks and only text). Big letters are not an important asset of a book for most students (Q5).

Students were also asked what they read in their free time (see figure 4.19). The main medium for reading were books, followed by comics, magazines and online content and ended with newspapers. Correlations (see table 4.5)

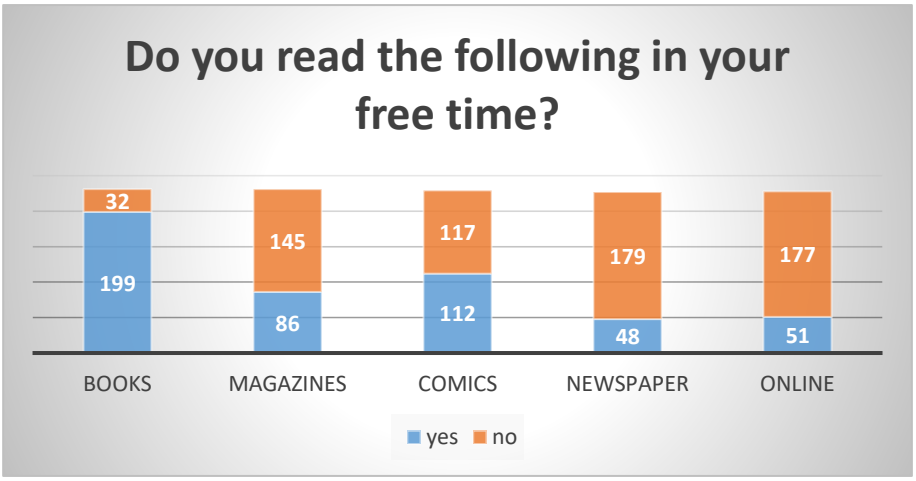


Figure 4.19 Text Types Read in Free Time

Table 4.5 Correlation Between Preferred Text Type and Gender / L1 Language

	gender	language
Bücher (books)	,112	-,078
Zeitschriften (magazines)	,027	**-,141***
Comics (comics)	**-,372****	,108
Zeitungen (newspaper)	-,006	-,078
im Internet (on the internet)	**-,172****	**,203****

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

The correlation results in table 4.5 show that children speaking German as L1 read more magazines than students with an additional or other L1 than German. Additionally, children with another L1 read more internet content. Moreover, the results show that boys read more comics and internet content than girls.

4.7.2 Books Borrowed

To establish a notion of the students' reading preferences on the basis of the most popular books of the library, I ranked the books according to the number of times they were borrowed (see section 10.1 in the Appendix for a full version). I am only going to present the results of the top 113 ranks [27] in more detail in the following; additionally, the results are summarized in the pie chart below (figure 4.20).

[27] I am not presenting the top 100, but the top 113 because the books on ranks 94 to 113 were all borrowed the same amount of times.

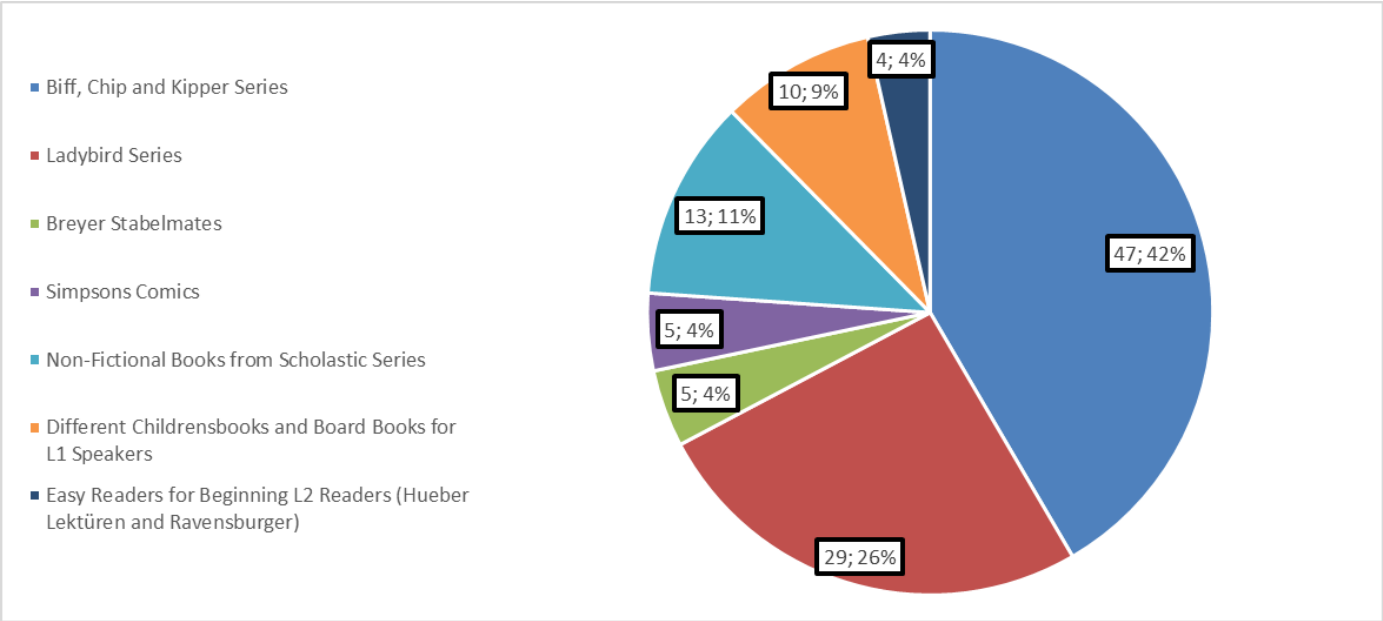


Figure 4.20 Types of Books Borrowed

The top 18 ranks are dominated by books from the *Biff, Chip and Kipper Series* (level 1-4) written for beginning readers of L1 English. Books from this series make up almost 50 percent of the top 113 ranks. Moreover, 47 of the 58 *Biff, Chip and Kipper* books in the library were in these top ranks. This series is followed by the *Ladybird* books, a similar series also aimed at beginning L1 English readers. With 29 out of 51 *Ladybird* books in the library in the top ranks, these make up almost another 30 percent of the top 113.

Additionally, further books aimed at beginning readers or children in their English L1 are represented in these ranks: all of the five *Breyer Stablemates* books of the library (approx. 4%), different board and children's books (approx. 9%) and different non-fictional books (approx. 11%). Only four (approx. 4%) of the most popular books are easy readers for beginning L2 readers: three from the Hueber Verlag and the only book in the library from the Ravensburger Verlag. When looking at these numbers in reference to the number of books of one specific kind available in the library, it becomes obvious that easy readers which are aimed and designed exactly for the investigated learner group were not the preferred reading material.

Additionally, to see if there were differences in reading preferences between intensive and extensive readers, the following table 4.6 shows how many books were borrowed by which group. These numbers refer to all books borrowed and not only the top 113. The absolute score is the total number of books borrowed from that category and the relative score in % is the total score divided by the number of books borrowed in total. The relative score is still not comparable, because there were different numbers of books of each category in the library, therefore, this score was again divided by

the number of books from this category (see # in the table below for amount), leading to a weighted relative score.

Table 4.6 Intensive vs. Extensive Books Borrowed

	intensive reading			extensive reading		
	absolute	relative in %	weighted realtive	absolute	relative in %	weighted relative
L1 Board Books (#18)	20	2,36	0,13	77	5,07	0,28
L1 Children's Books (#20)	17	2,01	0,10	65	4,28	0,21
Breyer Stablemates (#5)	12	1,42	0,28	43	2,83	0,57
Dirty Bertie Series (#14)	9	1,06	0,08	38	2,50	0,18
Advanced or Adult Books (#8)	11	1,30	0,16	13	0,86	0,11
Simpsons Comics (#18)	40	4,73	0,26	75	4,94	0,27
Biff, Chip and Kipper Series (#58)	524	61,94	1,07	510	33,60	0,58
Ladybird Series (#51)	81	9,57	0,19	421	27,73	0,54
Scholastic Non-Fiction (#36)	50	5,91	0,16	133	8,76	0,24
National Geographic Kids (#8)	11	1,30	0,12	19	1,25	0,11
Huber Lektüren (L2 Easy Readers) (#23)	28	3,31	0,14	35	2,31	0,08
Other L2 Easy Readers (#128)	43	5,08	0,12	89	5,86	0,14

The results show that books from the *Biff, Chip and Kipper* series were by far the most popular books in the intensive reading intervention. These are followed by the *Breyer Stable Mates* books and the *Simpsons Comics*. In the extensive reading intervention, *Biff, Chip and Kipper* and the *Breyer Stablemates* series were also very popular, but do not show such a difference in popularity as in the intensive reading intervention. Along with these two series, the books from the *Ladybird* series proved to be equally popular in the extensive reading interventions.

When extracting these results, I had to go through the digital library card, listing all the borrowed books one by one to code each book. I had the notion that the way books were borrowed

differed between the extensive and the intensive reading intervention groups. Both groups borrowed books of different kinds and series in the beginning, probably trying out what they liked and understood, but once they read a *Biff, Chip and Kipper*, a *Breyer Stablemates* or a *Ladybird* series book, they usually stayed with this series for some time. The extensive groups seemed to read broader, trying out new things in between reading books from these series. I also noticed that some students in the extensive reading group borrowed books they had read in one of the first lessons again in one of the later lessons, maybe rereading favorites or books they did not quite understand the first time.

4.7.3 Lending Behavior

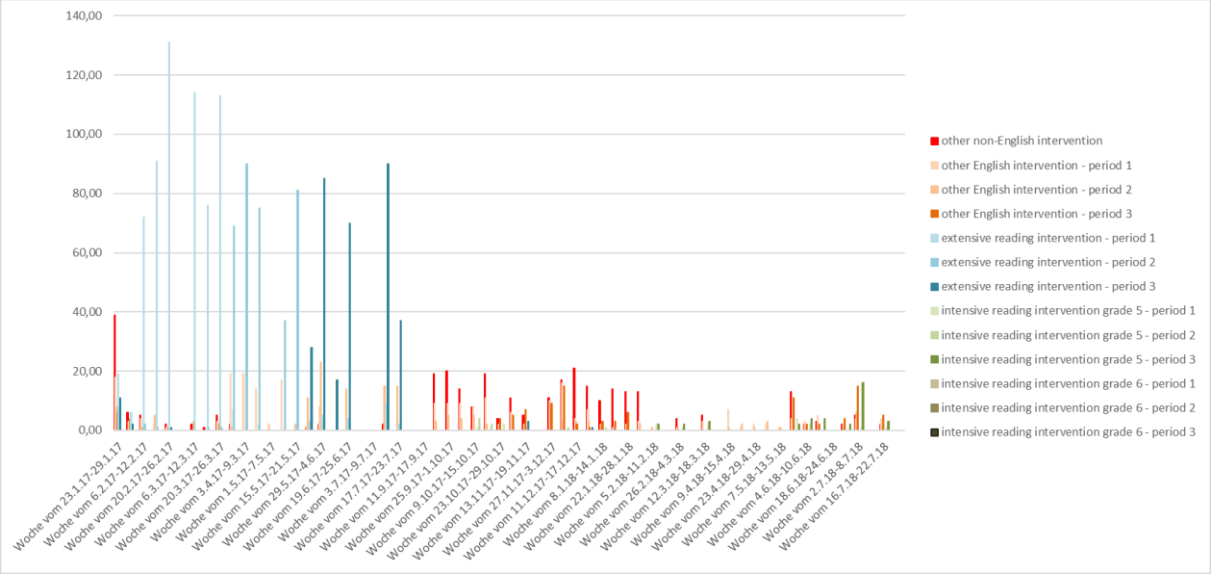


Figure 4.21 Books Borrowed per Week

Diagram 4.21 shows the books borrowed per week per intervention group. The results show that an initial interest in week one, when the library was introduced, quickly flattened. Additionally, students participating in the extensive reading interventions (blue bars) read multiple times the number of books that students in the intensive reading intervention (green and grey bars) read.

My initial anticipation, when designing the study, was that students participating in the reading interventions would continue reading after the end of the intervention and, therefore, tracking the number of books makes it possible to see when these effects fade and students need to be remotivated. The numbers - and also my experiences when running the library - unfortunately - show that there is no such effect of the intervention. The number of books borrowed immediately drops to zero when the last week of the intervention is reached.

5 Discussion

Once we accept our limits, we go beyond them.

- Albert Einstein -

5.1 Impact of Extensive and Intensive Reading Interventions on L2 Competence Gain

ANCOVA and MANCOVA analyses showed that there was no significant difference in the effects of the intervention types on the students' improvement on the language tests and their English grade. There are several possible explanations for these findings that I would like to shortly discuss:

First, a common explanation when discussing non-significant effects is that the group size might not have been large enough to produce statistically significant effects of actual underlying group differences. Power analyses using G*Power version 3.1.9.7 (Faul et al., 2007) show that the minimum sample size required to test effects of the four intervention groups on the competence gain variables was adequate. The results of the power analysis for the ANCOVAS, for example, indicated that the required sample size to achieve 95% power for detecting a medium effect (f = .25), at a significance criterion of α = .05, was N = 210 (denominator degrees of freedom = 199). If the power is reduced to 80% even a sample of N = 128 (denominator degrees of freedom = 117) would have been sufficient. Therefore, it should have been possible to detect medium effects with the sample size in this study, making it sound to assume that there are no differences in the effects of the interventions on point or grade improvement.

A second explanation for the non-significant effects could thus be that the effects were too small to be detected. Indeed, a sample of over 800 students would have been necessary to detect small (f = 0.10) effects. Since implicit language learning needs extensive reading input and multiple repetitions (Grabe, 2009), the amount of extensive reading and length of the interventions in this study might not have been long enough to produce large enough competence gain advantages, but would have produced stronger effects if practiced over a longer period of time. Further research with longer intervention periods or an even larger sample size would be helpful.

A third explanation could be that the interventions did not produce statistically significantly different effects because there are no differences. One explanation for this could be that students at this beginning level of English do not have enough L2 language competence yet to fully profit from reading lessons. However, if this had been the case, you would expect the reading intervention groups and non-Englisch control groups to have the same effects, but not the other English intervention groups with more explicit language training. Another explanation could be that students profited alike

from the reading interventions and other English interventions. Especially since it was very important for me to keep as many factors as possible constant between the experimental groups, this could be a result of reduced bias in this study compared to previous studies. In the light of previous research non-significant effects between different intervention types are also neither new nor surprising (Smith (2006), Mason (2004), partially Yamashita (2008), see section 2.3.3.) and were explained this way – reading interventions have the same effects as other language interventions. Yet, if this were the case, the effects of the English intervention groups should differ from the non-English intervention groups. Yet, none of the intervention groups significantly differed in their effects on the outcome variables.

This leads to the conclusion that either the effect sizes of the interventions are too small to show statistically significant effects or that the English interventions are not effective. But, as already mapped out in chapter two, the nature of intensive, extensive and other explicitly taught English interventions differs and therefore presumably also their effects on the competence gain of the students in different test types. Therefore, there could be underlying factors that moderate the results in such a way that there are no significant main effects, but significant crossover effects. This was indeed the case in this study, as the moderation results presented in the previous chapter show. These crossover effects could also explain why some previous studies have found differences in the effects of their reading interventions and others have not. Maybe the studies that found differences had more homogeneous intervention groups concerning variables that moderate the effects of extensive and intensive reading, compared to the studies that did not find main effects. These moderation effects are discussed in more detail in the following sections.

To move away from inferential statistics and necessary effect sizes, the following table 5.1 provides an overview on which intervention group the 10 students with the highest gain in points belonged to. These results are presented separately for each subtest.

Table 5.1 Intervention Group of Students with Top 10 Competence Gain Scores

	Non-English (N = 111)	Other-English (N = 131)	Intensive (N = 82)	Extensive (N = 40)
Total Points	0	**6**	**4**	0
	(0.00)	**(0.05)**	**(0.05)**	(0.00)
Cloze Test	2	4	2	**2**
	(0.02)	(0.03)	(0.02)	**(0.05)**
Fictional RC	**4**	2	**3**	1
	(0.04)	(0.02)	**(0.04)**	(0.03)
Prepositions	**4**	3	**3**	0
	(0.04)	(0.02)	**(0.04)**	(0.00)
Non-Fictional RC	0	4	**4**	**2**
	(0.00)	(0.03)	**(0.05)**	**(0.05)**

absolute scores, relative scores in brackets; highest relative scores printed in green

The total points gain (all test parts combined) shows that six of the ten students with the highest competence gain had taken part in the other English interventions and four in the intensive reading interventions (see first row); similar results show for the non-fictional reading comprehension (see last row). The results are more diverse for the other subtests, but again, the students in the extensive reading comprehension do not form the majority in the ranking of the top ten; an exception are the results of the cloze test and the non-fictional reading comprehension. Overall best results, taking all subtests into account, are achieved by students participating in the intensive reading interventions. These solely descriptive results could lead to the impression that intensive reading interventions have a higher potential for improving students' language competence than extensive reading. The question of which factors play a role in enabling students to profit from intensive and extensive reading interventions is discussed in the next sections.

5.2 Impact of L1 Reading Experience and L2 Language Competence (H1)

Section 2.2.4 presented theories on the impact of L1 reading experience and L2 language competence on reading and reading comprehension. It was hypothesized that:

Both, L1 reading experience (measured as a self-report of L1 books read per week) and L2 language competence (measured as scores in language test at T1) will be significantly associated with competence gain for students in the English language tests and grade improvement.

a. For the students in the non-English intervention group, I expect L1 reading experience and L2 language proficiency both to have a positive impact on the competence gain.

b. For students that receive additional training, especially for the reading interventions, I do not expect these effects to be as high or significant, or even negative.

5.2.1 Impact of L2 Language Competence

MANCOVA and regression analysis results show that points achieved in language tests at T1 (initial L2 language competence) play a significant role for the competence gain in all test parts and across all intervention groups. The initial points in the language tests are the strongest predictor for the competence gain in the corresponding following language tests. For example, the points students achieved in the first fictional reading comprehension at T1 explain at least 30% (40.4% in the non-English intervention, 29.6% in the other English intervention, 78.8% in the extensive reading intervention and 32% in the intensive reading intervention) of the variance in competence gain from T1 to T4 in the fictional reading comprehension tests. There is a similarly high explanation of variance for non-fictional reading comprehensions, cloze tests and preposition tests (see section 4.6).

Bernhardt (2005) suggested that L2 knowledge explains around 30% of the variance in reading comprehension, thus, these numbers would fit, but the effect in this study is negative. Students with low points in a test part at T1 have a higher increase in their scores in that test part from pre- to posttest. Taking into account that students with lower scores have the highest potential to develop, and seem to do so, these results are reasonable. Moreover, the estimated 30% in previous research describe the contribution of L2 competence to L2 reading comprehension at one point in time, but do not give any prediction of the effect of L2 competence on L2 competence gain over a period of time, as it is the case in this study. The effects are, thus, not comparable. These results contribute to the discussion in that they show that students with low L2 competencies strongly develop their language competence over time – more than students with a high initial language competence - and that more longitudinal research is necessary to see how the effects of L2 language competence develop over time.

Positive effects of language competence on long term competence gains show for non-corresponding test results, especially the cloze test results in T1 for delayed pre-/posttest gains of other tests. In the non-English interventions these explain an additional 6.4% of the variance of the non-fictional reading comprehension gain, 17.4% of variance in the preposition test points gain and 17.1% of the fictional reading comprehension gain. This effect does not show in the non-reading English interventions. However, the fictional reading comprehension gain in the extensive reading intervention shows a smaller but comparable effect (12.8% additional explained variance in the direct and 6.3 in the delayed pre-/posttests), as well as strongly again for the intensive reading interventions (13.6% for the non-fictional reading comprehension gain, 25.9% for the preposition test gain and 15.4% for the fictional reading comprehension gain). Assuming that cloze tests measure a somehow general language ability (see discussion in section 3.3.3.1), these results give evidence for a significant impact of language competence on competence gain in the non-English intervention groups and reading intervention groups. That this effect does not show for all test parts in the extensive reading intervention could be due to statistical reasons – the extensive reading intervention has a smaller number of participants and, therefore, a smaller *N* than the other groups. Further research with larger groups will be necessary to see if these differences originate in the nature of intensive and extensive reading interventions or are just statistical products. Cloze test results in T1 did not significantly explain additional variance in the other English interventions. The other English intervention groups train grammar, writing and reading competence more explicitly and directly than in the reading interventions and there are two groups for strong and weak students. This could be one reason why initial language competence does not have a significant effect on the competence gain in the other English intervention groups, because students are supported at their level and improve accordingly. It would be desirable to have more research in the field using various language tests to see which aspects

of “language competence” influence which aspects of “competence gain” or “comprehension” in what way, because there seem to be differences in development and effect depending on which competences are measured.

Cloze test results also significantly moderated language test results, as the differences in effects between the different groups already show: Students in the non-English interventions (= control group, baseline) who have low scores in the cloze test, show higher improvement in cloze test scores (intervention period), whereas students with high initial cloze test scores show lower improvement than the entire control group (baseline) in the posttest. This has already been discussed in the previous section on the regression analyses results – students with low initial scores in the language tests have a higher gain in points from pre- to posttest. This is very likely due to their initially low scores, because they have more potential to improve. Therefore, it is interesting that this is not the case for students in the reading interventions: Students with low cloze test scores also show a higher improvement than all non-English intervention students (baseline), even though this is not as high as for the non-English intervention group with equally low cloze test scores. However, the students who take part in the reading interventions improve more strongly, when their cloze test scores are high. Similar effects show for the non-fictional reading comprehension (semester period). Students with a high initial language competence, consequently, seem to profit more from reading interventions. That this shows for non-fictional reading comprehension, but not for fictional reading comprehension and cloze test scores, could be due to the circumstance that non-fictional texts are something students do not encounter as much in their English class books as in the reading interventions. Moreover, non-fictional reading involves coping with unknown, technical vocabulary and a more academic form of writing. Therefore, reading interventions at this early stage of language learning might be a boost, especially for more competent students, for developing academic English reading skills.

Cloze test results also moderated the impact of extensive and intensive reading interventions on the fictional reading comprehension gain (intervention period). Whereas students in the extensive reading intervention outperform students in the intensive reading interventions when their initial cloze test scores were low, this effect is reversed for medium or high initial cloze test points. Students who already have a high language competence (= high cloze test scores) profit much more from intensive reading interventions concerning their fictional reading comprehension points gain and perform worse than the non-English control group (baseline) in the extensive reading intervention. Therefore, at least for the improvement in fictional reading comprehensions, L2 competence is decisive for whether a student profits from extensive or intensive reading interventions.

Moreover, non-fictional and fictional reading comprehension scores at T1 are significant predictors of English grade improvement, explaining 22.9% of the variance in grade improvement in

the non-English intervention group, 17.1% in the other English intervention group, 32.4% in the extensive reading intervention group and 27.8% in the intensive reading intervention group. Students who have a high reading competence, thus, improve their grade much more. These results support the idea that reading competence predicts academic success.

5.2.2 Impact of L1 Reading Experience

The impact of L1 language competence that was measured here as a self-report of the hours spent reading in the L1 is also part of this first hypothesis. Regression analyses only showed small impacts of German L1 reading: It was the strongest predictor of direct fictional reading comprehension gain in the other English interventions, but with only 6.8% explained variance. Moreover, this did not show for delayed results. In the intensive reading intervention, German L1 reading explained another 10.9% variance in non-fictional reading comprehension gain. Consequently, these results do not support the Developmental Interdependence Hypothesis, because there is no evidence for a strong impact of L1 reading experience on language competence. Yet, L1 reading experience seems to be a small but decisive factor for profiting from reading interventions.

Indeed, moderation analyses show that students' private L1 reading behavior seems to moderate the effect of intensive and extensive reading interventions. This shows for non-fictional reading comprehension results as well as for the grade improvement: Students who read fewer hours in their L1 at home have positive competence gains and, therefore, seem to profit more when having participated in the extensive reading interventions, whereas students with more hours of L1 reading have higher competence gains and, therefore, seem to profit more when having participated in the intensive reading intervention. The origin of this difference cannot be identified on the basis of the data collected in this study, but one explanation could be that students with low L1 reading experience, and, as a consequence, not as developed reading skills, profit more from extensive reading input and are overchallenged in intensive reading interventions where additional tasks demand a more intensive analysis and different processing of the reading (see section 2.2). Students who read more in their L1, and have consequently developed stronger reading skills, profit more when engaging in additional tasks in the intensive reading intervention that help them to develop and transfer their reading skills. In what way these students really profit from an interdependent transfer of reading skills can only be assumed but not answered in this study. An interesting design for further research could be to see if students who are poor readers in their L1 show a development in their L1 reading skills when they participate in extensive L2 reading interventions and vice versa.

5.2.3 Developmental Interdependence or Language Threshold Hypothesis

All in all, the results on L1 reading experience in this study show some support of the Developmental Interdependence Hypothesis in that students who read a lot profit more from intensive reading interventions due to a possible skill transfer of highly automated L1 reading skills. Yet, the more dominant factor in predicting students' competence gain is L2 language competence before the intervention, especially the points they achieved in the cloze test. This rather supports the Language Threshold Hypothesis - the idea that students have to reach a certain amount of L2 language competence in order to be able to transfer L1 reading skills.

In this study students profited from reading interventions (cloze test and non-fictional reading comprehension) when their initial L2 language competence is high (cloze test T1). These results could be explained and seen as additional support for the Language Threshold Hypothesis, because students with higher L2 competence are able to transfer more L1 reading skills and, thus, profit more from reading interventions.

5.2.4 Evaluating Hypothesis 1

To summarize, this first hypothesis can only partially be accepted. L1 reading experience (measured as a self-report of L1 books read per week) is marginally significantly associated with competence gain and solely in the reading interventions. Intensive and extensive reading also differ in the way students profit from their L1 reading experience.

L2 language experience (measured as scores in language tests at T1) is significantly associated with the students' L2 competence gain. Students' language test results in T1 are the strongest predictor for competence gain in that particular test type – students with low scores improve significantly more than students with high test scores. Moreover, cloze test scores are positively associated with competence gain. Other than hypothesized, this does not only show for students in the non-English intervention group, but also for the reading interventions.

5.3 Impact of English in Primary School on Language Competence Gains (H2)

Section 2.2.4.1 stresses the importance of rapid word recognition for successful reading and traces important similarities and differences between German and English that could enhance but also hinder transferring L1 reading skills concerning word recognition. English and German have very similar orthographies; students will recognize letters in English. But English has, other than German, a deep orthography making it difficult to match letters or graphemes with sounds. Additionally, students learning English as an L2 have not had as extensive auditory English input as in their L1 when learning

to read, making it even more difficult to cope with intransparent phonology-orthography correspondences. It is, therefore, hypothesized that students who had more English lessons in primary school and, thus, have more oral language experience in English, are at an advantage in the reading interventions:

> *Independently of the intensive or extensive reading intervention, the students' competence gains in the pre- and posttests is positively related to the number of years of English they have had at primary school.*

As hypothesized, years of English at primary school significantly predicted the competence gain achieved in the fictional reading comprehension (intervention period in regression analyses and semester period in moderation analyses) of students in the extensive reading intervention. It even was the strongest predictor with 28.7% explained variance. Primary school English also significantly predicted the competence gain in preposition tests (semester period) in the intensive reading intervention, explaining an additional 6.1% of the variance in competence gains. These effects were, as predicted, positive – more years of English in primary school lead to higher competence gains in the reading interventions.

Primary School English was not a significant predictor in the non-English intervention, but in the non-reading English one, also for preposition score gain, explaining an additional 5.5% of the variance. However, in this case, the direction of the effect is negative – fewer years of primary school in English lead to higher competence gains. This effect can, again, be explained by the nature of the other English interventions: Students who have not yet had many English lessons in primary school profit from the extra training in the other English interventions, catch up and improve more than students who have had English in primary school.

Consequently, there is strong evidence that supports the second hypothesis. Students who have had fewer English lessons in primary school profit more from additional English lessons that focus on grammar and skills training, whereas students who have had more English in primary school profit more from reading lessons. In line with these results and reasoning, the effect of primary school English seems to be stronger in the extensive reading intervention than in the intensive reading intervention. These results are alarming, taking into account that English is not taught from first grade on anymore, but has been delayed to grade three in primary schools since 2017. In this study 43.1% of the students report that they started learning English in their first year at primary school, because they visited primary school before 2017. This would not be the case anymore. Moreover, even though English had been taught at primary school from first grade on at that point, 11.9% of the students report that they started learning English in primary school in grade three or four and 1.5% of the students even report to not have had any English lessons at all.

5.4 Text and Situation Model during Extensive and Intensive Reading (H3 and H4)

5.4.1 Differences in Test Results (H3)

Section 2.2.4 and the examples from the book *The fair haired Saumrai* illustrated that syntactic information is essential for establishing meaning and, therefore, a text model of the text. It was reasoned that, on the one hand, students in the intensive reading interventions might develop linguistic awareness through the additional tasks. The ability to make use of semantic and syntactic cues to achieve understanding would most likely show in cloze test and preposition test results. On the other hand, students in the extensive reading intervention receive more L2 input

Moreover, students in the extensive reading intervention are advised to read for pure pleasure, understanding the gist, whereas students in the intensive reading interventions have to engage in various tasks afterwards that require a more precise understanding of the text and its structures and call attention to vocabulary and syntactic structures. Therefore, extensive reading might emphasize the use of a situation model and intensive reading building the use of a text model (see section 2.2.4.5).

Taking these differences between intensive and extensive reading into account, I hypothesized that:

> *Students participating in the extensive reading intervention will perform better in the fictional and non-fictional reading comprehension tests than students in the intensive reading intervention, who will perform better in the cloze tests and the preposition tests.*

Descriptive statistics give some evidence for this third hypothesis, but the mean differences proved to be non-significant in inferential statistics. The mean improvement in the cloze tests is higher for the intensive reading interventions (but only in grade 6 for the direct and grade 5 for the delayed pre-/posttests) than for the extensive reading interventions (see results in section 4.1.1 for details), the results for the preposition tests are similar (see results in section 4.1.2 for detail). Therefore, even though there are some differences between grade 5 and grade 6, there is some support for this hypothesis. Also, as hypothesized, the means in the non-fictional reading comprehension are different: The extensive reading group has a higher average improvement than the intensive reading intervention groups (see results in section 4.1.3 for detail). What does not fit are the results of the fictional reading comprehension. Here the intensive reading intervention groups in grade 6 had the highest increase and not the extensive reading intervention. These mean differences were, as the MANCOVA analyses presented in section 4.3.2 show, all insignificant.

Since non-significant results do not mean that there is no impact, this third hypothesis can neither be rejected, nor accepted. Intensive reading interventions could lead to better results in cloze tests and preposition tests, whereas extensive reading interventions could lead to better results in

non-fictional reading comprehensions, as means interestingly show exactly as hypothesized. Statistical issues (too small group size and different group sizes, differences between grade 5 and grade 6 for intensive reading interventions, influences of covariates and moderation effects) might have led to too small effect sizes concerning this question. For now, this hypothesis can only be confirmed with a certain and large amount of reserve and maybe not even that.

5.4.2 Differences in Reading Preferences (H4)

In section 2.2.4.5 it was reasoned that students in the extensive reading intervention might be tempted to draw more strongly on situation model building. I hypothesized that this would also reflect in their choice of books:

> *In the extensive reading intervention students will prefer reading books where they can draw on background knowledge to understand them, whereas this will not show as strongly for intensive reading.*

Section 4.7.2 shows the results of students' borrowing behavior divided by intensive and extensive reading. In both classes the *Biff, Chip and Kipper Series* is the most popular, followed by *Breyer Stablemates*. Other than expected, the *Simpson Comics* are the third most borrowed books in the intensive reading intervention. These are definitely far beyond the level of 5th and 6th graders and it is quite unlikely that the students in the intensive reading interventions understand what they are reading. They also did not fill in any task on the Simpson Comics in their portfolios. Therefore, they probably did not finish the comics, realized that they were too difficult and continued with something else. Thus, there is no evidence that students in the extensive reading intervention relied more strongly on books with picture support than in the intensive reading intervention. Also, non-fictional books were read to about the same amount, but were a little more popular in extensive reading.

The popularity of the *Ladybird Series* differs between the interventions. The books were as popular as the *Biff, Chip and Kipper Series* and the *Breyer Stablemates* in the extensive reading interventions, but far behind in the intensive reading interventions. The *Ladybird Series* contain stories the children already know, like Classic Fairy Tales or *Peppa Pig* stories. These are stories where students can draw upon their background and genre knowledge for understanding. In contrast, L2 easy readers were much more popular in intensive reading than in extensive reading. Observations of the borrowing behavior and the distributions of the books and genres borrowed show that students in the extensive reading interventions read more braodly, which means that they tried out different books and genres, whereas reading in the intensive reading interventions was narrower.

All in all, the results support the fourth hypothesis: The results show that students in the extensive reading intervention do seem to prefer books where they can draw on background knowledge more strongly, like fairy tales or TV series, than students in intensive reading interventions. Taking the discussion of hypotheses three and four together, students in the extensive reading interventions might in fact draw more strongly on building a situation model when reading, whereas students in the intensive reading interventions might focus more strongly on building a text model. Even though these results are only based on descriptive statistics, further research in this direction could be promising.

5.5 Goal Orientation

5.5.1 Learning Goal Orientation (H5 and H6)

In section 2.3.5.1 it was reasoned that goal orientation moderates the effect of the reading interventions:

> *Students who have a stronger learning goal orientation profit more from the extensive reading intervention (higher competence gain) than students with a higher performance-approach goal orientation, who profit more from the intensive reading intervention (higher competence gain). These effects are presumed due to a better fit of the task to the students' motives.*

Indeed, learning goal orientation moderated the proficiency gain of the cloze test results (intervention period) of the intensive and extensive reading interventions as hypothesized. Performance-approach goal orientation also moderated proficiency gain of the cloze test results (intervention period), however, not as expected for the intensive and extensive reading interventions, but for reading interventions in general compared to non-English interventions. Students with a low performance-approach orientation show stronger improvements in the reading interventions than students in the control group (non-English intervention). This is reversed for high performance-approach goal orientations. A possible explanation is that even though intensive reading interventions involve filling in tasks and receiving feedback, these are not graded and still involve an amount of self-engagement and intrinsic motivation to engage with the tasks. Moreover, the tasks are performed quietly without interaction with others. Students with a high performance-approach goal orientation might still not receive enough positive feedback from their teachers and peers to be extrinsically motivated to read.

All in all, these results show that learning goal orientation is a decisive moderator and therefore factor for the success of extensive and intensive reading interventions and support the fifth hypothesis. The assumption that performance-approach goal orientation is also a decisive factor, could

not be confirmed. Yet, these results are evidence that intensive and extensive reading do not only differ concerning cognitive but also motivational factors and that the students' fit to these factors needs to be taken into account when planning and conducting reading interventions.

Since learning goal orientation has shown to positively influence skills and strategies that are associated with higher academic achievement (see section 2.3.5.1) it was further hypothesized that:

> *Learning goal orientation has an overall positive effect on students' competence gains in the different language tests. This effect is independent of the intervention group type.*

Learning goal orientation in T1 does not correlate significantly with any of the language test or grade gains. Thus, this relationship of the sixth hypothesis cannot be confirmed with the results of this study.

5.5.2 Work Avoidance

Apart from the hypothesized impacts concerning goal orientation, work avoidance proved to be a very significant factor. Work avoidance was a strong predictive power of competence gain and strongly moderated the effect of the reading interventions on cloze test improvement (intervention period).

First, work avoidance moderated the competence gain in cloze tests (intervention period). Students only profited from the reading interventions when their work avoidance score was low. These effects do not only show in comparison to the non-English reading interventions but also to the other English interventions. These results are quite comprehensible, since the reading interventions are self-paced and students have to take responsibility for their learning process. Therefore, the students' attitude to work is very likely to influence the way in which they profit from these reading lessons and their English competence gain, here reflected in higher or lower cloze test results.

Second, in the extensive reading intervention, work avoidance is the strongest predictor of cloze test points gain (intervention period) with 17.4% explained variance and the second strongest of grade improvement, explaining an additional 16% of the variance. What is striking is that in both cases the relationship is negative: Students with high work avoidance scores have a lower cloze test points gain, but (due to the grading scale) have a higher grade improvement. The direction of moderation for the cloze test results is in line with the argument why students with low work avoidance scores seem to profit more from the reading interventions than with high work avoidance scores. Just sitting in the lesson and simply looking at pictures in a Simpson comic - because this is the most fun and least work - does not improve English skills, but engaging in books at the right level and taking responsibility for one's own learning process will. A higher grade improvement for students with high work avoidance

scores does not fit this picture. A far-fetched and very speculative explanation could be that these students are clever ones who know how to maximize their success with the least effort.

Third, MANCOVA analyses with work avoidance as a dependent variable showed that intensive and extensive reading interventions also had an impact on work avoidance scores. Students participating in the extensive reading intervention have higher work avoidance scores after the interventions (T4) than before the interventions (T1). In contrast, students participating in the intensive reading intervention have lower work avoidance scores after the interventions (T4) than before the interventions (T1). A possible, but again speculative, explanation could be that students in the extensive reading intervention, due to the voluntary way of working, are confirmed in their work avoiding attitude, whereas students in the intensive reading intervention experience that engaging in a book and tasks without the pressure of a graded classroom setting can be fun and develop a more positive attitude towards work. Again, this is pure speculation.

Even though work avoidance, at first, did not seem to be a likely predictor in any way for this study, which is why I did not develop any hypothesis around this construct, it seems to play a more central role for the success and nature of extensive and intensive reading interventions than I assumed. Yet, the reasons for these relationships do not become clear in this study. It might be interesting for further research to also take personal statements of students into account on how they liked the intervention, whether they feel they have profited and in what way. These statements could give additional qualitative insight into reasons for some of these results.

5.6 Self-Concept

5.6.1 Impact of Comparison Frames (H7)

Section 2.3.5.2 describes how an internal and external frame of reference and the big-fish-little-pond-effect effect self-concept development. Since students of various classes come together in the reading interventions and form a totally new group, they might experience themselves differently and more competently in the English reading interventions than in their regular English classes - a big-fish-little-pond effect might reverse and lead to a more positive English self-concept. Moreover, students in the non-English interventions take part in math or science classes, enhance their mathematical skills and thus might develop a more positive self-concept in these areas. According to the I/E-Model, this might effect their English self-concept in a negative manner (see section 2.3.5.2). I hypothesized that:

English (reading) interventions will have a positive impact on the students' self-concept (self-concept gain), whereas non-English interventions will have a negative impact on the students' self-concept.

MANCOVA analyses of the effect of reading intervention types on self-concept gains show a significant impact of intervention type. Post-hoc analyses revealed that intensive reading has a negative impact and extensive reading a positive impact on students' criterial self-concept. Since age could be relevant in this context, results of the intensive reading intervention were divided into grade 5 and grade 6, but did not prove to be significantly different. Therefore, the hypothesis can only be accepted for extensive reading interventions and there is also no evidence of a negative impact of non-English interventions on students' English self-concept.

5.6.2 Skill-Development vs. Self-Enhancement (H8)

Positive or negative self-concepts of students concerning their English abilities can influence their achievements in test situations and their language development (self-enhancement view, see section 2.3.5.2 for details and research). On this basis, I hypothesized that:

Students' self-concepts at the beginning of the interventions (T1) will correlate significantly positively with their test score gains and significantly negatively with their average grade difference (meaning a positive relationship of self-concept and grade improvement due to lower grades being better than higher grades). This effect is independent of the experimental group the students are in.

Correlations of the different self-concept scores show that this hypothesis cannot be confirmed. In the non-English intervention group (control group) criterial self-concept, individual self-concept and absolute self-concept correlated negatively with the cloze test points gain (semester period). These correlations were highly significant. There are further significant negative correlations in other groups and for other test parts and, correspondingly, positive correlations of grade difference (see section 4.2.2 for detailed results). Carefully generalizing these results, self-concept is negatively associated with competence gain, which is not in line with the expected.

Moreover, moderation analyses show that absolute self-concept moderates the effect of reading interventions on grade improvement: Whereas the absolute self-concept does not influence grade improvement in the non-English and other English reading interventions, students in the reading interventions with low absolute self-concepts improve significantly compared to students with high absolute self-concept scores. The same effect shows for non-fictional reading comprehension gains (intervention period), not only for absolute self-concept scores but also for individual self-concept

scores. Reading interventions can, therefore, help students with a low absolute or individual English self-concept to improve their English skills.

Concerning criterial self-concept, the effects on the fictional reading comprehension gain (intervention period) differ concerning intensive and extensive reading interventions. Students with a low criterial self-concept seem to profit more from intensive reading interventions, whereas students with a high criterial self-concept seem to profit more from extensive reading interventions. Moreover, MANCOVA results with self-concept as a dependent variable showed that the criterial self-concept of students significantly increased in the extensive reading intervention and decreased in the intensive reading intervention. Thus, extensive reading seems to have a positive reciprocal effect on students: Students with a low criterial self-concept develop a higher criterial self-concept through the participation in the extensive reading intervention (skill-development view). Moreover, these students then have a higher competence gain than students with low self-concepts (self-enhancement view).

5.7 Contributions to the Extensive Reading Research Field

In section 2.3.4.4 I stated that this study will contribute to the following question posed by Nation and Waring (2020) to enhance the extensive reading research field. I can now answer these for my sample. Since this is a repetition of what has already been discussed in the previous sections, I will only briefly answer the questions and give reference to other sections for further detail.

- **What are the favorite books and genres students read?**

  The students in this study preferred reading easy reader series for beginning L1 readers, like *Biff, Chip and Kipper, Breyer Stablemates* and the *Lady Bird Series*. See the results in sections 4.7.2 and 5.4.2 and the discussion in section 5.4.2, also concerning differences between intensive and extensive reading preferences.

- **Do students prefer 'native texts' or graded readers?**

  Borrowing behavior shows that students preferred texts that are simplified but authentic in that they are written for beginning L1 readers, but not easy readers for L2 readers. An assumption based on journal entries and talks with students is that these books have a much more attractive layout and that students like the series characters. Non-fictional books were equally popular to fictional books. See the results in section 4.7.2 and the discussion in section 5.4.2 for further detail. When asked about their reading preferences students answered very differently to what they actually borrowed from the library, saying that they preferred longer texts, sentences and books with vocabulary help (see section 4.7.1 and the following section 6.1.1).

- **How do learners from different language groups / abilities levels benefit from extensive reading?**

  Students with an above average L2 English proficiency profited more from extensive and intensive reading interventions concerning their improvement in further cloze tests and non-fictional reading comprehensions (see section 5.2.1). Students with a below average L2 English proficiency profited more from extensive reading interventions concerning their fictional reading comprehension skills.

- **Can all learners benefit from extensive reading?**

  On the basis of this research, this answer needs to be answered with a clear 'no'. Students with a

  - high initial language competence
  - few English lessons at primary school
  - high L1 reading experience
  - low learning goal orientation
  - high performance-approach goal orientation
  - high work avoidance
  - low criterial self-concept
  - high absolute or individual self-concept

  do not seem to benefit or benefit as much from extensive reading interventions as from other intervention types.

- **How do gains in extensive reading compare with those from other 'rich' forms of input?**

  Students with the traits stated in the previous question profit more from other non-reading English interventions or intensive reading interventions.

- **How does extensive reading impact other language skills? Which skills? Why? Are there more effective methods than extensive reading affecting these skills?**

  Especially for low proficient students, non-reading English interventions are more effective. If intensive and extensive reading interventions differ concerning their impact on certain skills could not be answered for sure on the basis of the present results.

  Extensive reading, however, does not only impact language skills but also psychological concepts like self-concept and work avoidance.

6 Implications for Language Teaching

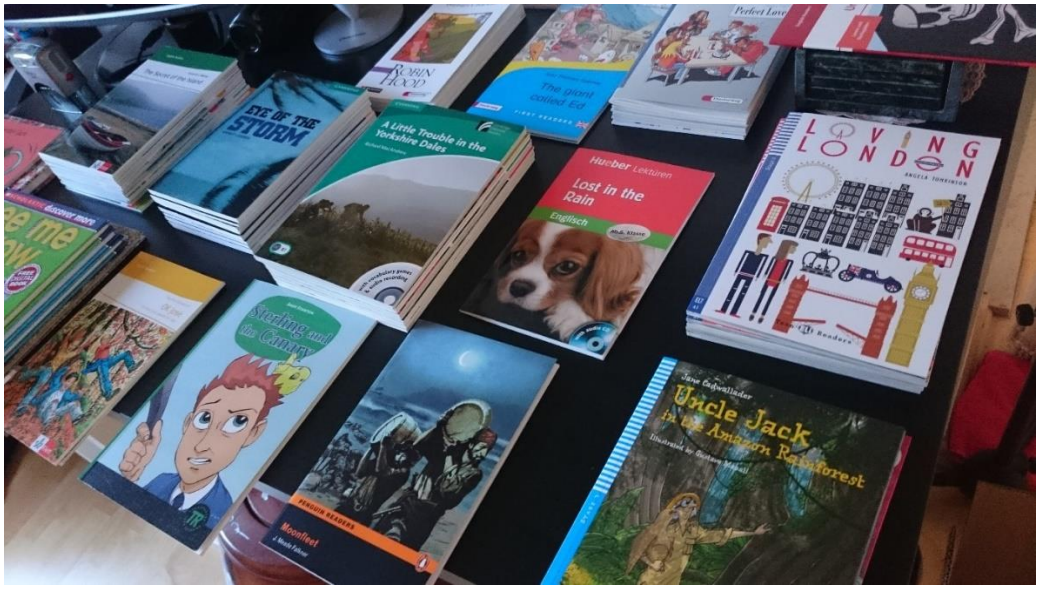


Figure 6.1 Preparing Newly Arrived Books for the Library

Not only the theoretical analyses of intensive and extensive reading literature and studies, but also the practical implication of this study have led to a great deal of experience in setting up reading interventions at school. I would like to share, summarize and elaborate further on some aspects that are relevant when wanting to implement such a program at school in the following sections (see also section 2.3.4 on the essentials of extensive reading programs).

6.1 Implementing an Extensive or Intensive Reading Course

6.1.1 Building up a Library

Choosing motivating books at the right level for the students is a very difficult but also crucial task for a successful reading program. The references of the books I chose for this study and their ranking on how often they were borrowed in the appendix, might give some guidance which books to choose, when setting up an own library.

In chapter 2.2.4.1 and 2.2.4.2 I reasoned that students are likely to encounter many words they do not know, either in that they might recognize the word but do not have a concept for it yet, or that they know the concept, but not yet the English word. Therefore, books with picture support can be very helpful to support the connection of the orthographic symbol of a word and its lemma, supporting lexical access. I speculated that easy readers aimed at English L2 learners do not implement as much short sentences, large font and picture support, because publishers might assume that students have

already established basic reading skills in their L1 and do not need support in this direction, especially because unknown words can be found in the glossaries. Students were questioned on these aspects; the results were presented in section 4.7.1. The results are very heterogeneous concerning picture support, but it is rather important to students that words are translated. Also, students reported to prefer longer books and texts with longer sentences to shorter books and texts with shorter sentences. Interestingly, the actual borrowing behavior does not reflect these self-reported preferences. Section 4.7.2 named the most popular books. Only 4.4% of the top 113 books are aimed at L2 readers, are longer and contain glossaries. 90.2% of the top 113 books are books aimed at beginning L1 readers with many pictures, large front and short stories. Therefore, when choosing books for an English library, I would suggest not only buying L2 easy readers, but also L1 beginning readers or children's books with short stories and many pictures.

As the type of books borrowed, presented in section 4.7.2, show, the *Biff, Chip and Kipper Series* and fairy tales as in the *Ladybird Series* were very popular in the reading interventions. One reason could be that students were able to develop a mental network around the series or could draw on background information, instead of having to constantly cope with new plots and characters. This helps them build a text model representation and integrate the new information in the text into an already existing mental network (see 2.2.4.5). Therefore, I also suggest integrating series with the same characters and similar plots into an English library. This is also something students named as being important to them in the survey of reading preferences (see 4.7.1).

6.1.2 Additional Tasks: Reading and Listening

This study did not implement additional listening tasks into the reading courses due to various organizational issues. Most of the books were not provided with an audio which would have biased the results of the reading preferences. Additionally, measures, for example using eye-tracking, would have been necessary, to ensure that the students actually read or look at the texts and not only listen to the audio. In reading interventions that do not need to fulfill certain experimental standards, I would consider providing the students with audios. This can help students overcome difficulties of English having a deep orthography (see orthographic depth hypothesis in section 2.2.4.1) and misleading German-English grapheme-phoneme correspondences.

6.1.3 Further Ideas for Implementation

Grabe (2009) states that "the development of L2 processing efficiency is brought about through consistent practice, extended exposure to L2 reading, the development of sight-word reading

of high-frequency words, and the growth of receptive L2 vocabulary knowledge" (p. 128). Especially because English has a deep orthography, it is not possible for teachers to rely on students quickly building an awareness for English grapheme-phoneme correspondences. Therefore, it is important to establish a sight vocabulary of high frequency words (Grabe, 2009). An idea how this could be done is to integrate quizzes at the beginning of English school lessons or reading intervention courses where students have to choose the correct picture to a given word (word – concept correspondence) and to practice grapheme-phoneme correspondences by presenting them with spoken words and the students choose the correct written word. Additionally, the teacher could present written words that the students then have to read aloud. If this is done regularly at the beginning of each lesson, integrating new words week by week, the students will quickly develop a considerable sight vocabulary.

6.2 Who Should Participate in Extensive or Intensive Reading Interventions?

6.2.1 Age

What age and, therefore, what English experience should students have to profit best or even be able to profit at all from reading interventions? This study has looked at the reading process and reading interventions not only from a cognitive but also from a motivational point of view. Both of these views will be taken into account in the following to discuss the question of age and intervention success.

From a cognitive point of view, students have had very little experience and input in English when expected to practice intensive or extensive reading. They can, therefore, neither fall back on a rich lexicon nor on experiences with the complex phoneme-grapheme correspondence in English. This makes it difficult to overcome a certain linguistic threshold to understand the texts and perceive them as rich input to enhance their English skills. Results in this study (section 5.2.3) also suggest that students' language competence strongly predicts the students' further improvement in language competence – also when participating in reading interventions. This is an argument for adding reading interventions to the language curriculum after a certain amount of language competence is established.

From a psychological point of view, age is also an important factor in the development of self-concept: Students with initially uncorrelated self-concepts are more likely to develop a negative self-concept than a positive self-concept. Moreover, with increasing age, the self-concept becomes more stable (see differentiation-distinctiveness hypothesis in section 2.3.5.2 for further reference). This speaks for integrating interventions that boost students' self-concept as early as possible. In grade 5

and 6, where students just begin to extensively learn English and develop an academic English self-concepts would be just the right time. In this study, extensive reading had a positive and intensive reading a negative effect on the students' self-concept (see section 5.6). Therefore, extensive reading interventions could help students to develop a more positive English self-concept or prevent them from developing a negative English self-concept. Moreover, students need a frame of reference to develop an English self-concept. If they do not have individual or criterial frames of references available, they will rely on social comparisons which are strongly biased by their social environment (see Big-Fish-Little-Pond-Effect in section 2.3.5.2). Reading interventions at an early age, especially when practiced in other groups than in the English class, as in this study, can give students these additional frames of reference.

Summarizing, especially extensive reading can be important for beginning learners to develop a positive self-concept, even if learners might not have the optimal language competence to profit from these interventions yet. Intensive reading, on the contrary, might lead to a negative self-concept. This is especially interesting taking into account that this is the commonly practiced reading type in the German English language classroom – the teacher chooses one book that is then thoroughly read with the entire class. Substituting these phases with extensive reading could be worth considering.

6.2.2 Proficiency

In this study, students of all proficiency levels participated in the reading interventions. At the beginning of this dissertation, I summarized the Matthews effect of reading – the rich get richer - as presented by Stanovich (2000). One important factor that enforces this effect are organism-environment correlations. Children with poor reading skills usually have not experienced reading facilitation at an early age at home and will not seek reading facilitating environments (Stanovich, 2000). As a consequence, this would mean that poor readers would not be reached by facultative and extracurricular reading activities. In this study students were not able to choose the reading courses, but were randomly assigned because in mandatory but motivating reading environments students can experience a positive reading environment. The teacher or fellow classmates can function as positive reading role models.

Whether students develop better reading skills in homogenous or heterogenous groups is a difficult question. On the one hand students might be motivated by positive reading role models in the group (enhancement of a positive Mathew effects), on the other hand they might develop a low reading self-concept due to negative comparisons (big-fish little pond) that decreases their reading motivation. To which extent these could be salient effects is probably very group specific and dependent on various other factors, like social competence and is something teachers should take into

account for their class individually. The reading groups in this study were very heterogeneous and this did not turn out to be a problem since the students worked individually.

The question is whether good and poor readers profit from reading interventions in the same way. As discussed in section 5.2.2., the results of this study show that students with a low L1 reading experience profit more from extensive reading interventions and students with a high L1 reading experience from intensive reading interventions. A possibility to deal with this aspect in the language classroom is to integrate additional language tasks, as in the intensive reading interventions in this study, for more advanced and experienced students.

6.3 Motivational Issues: Getting Students to Continue Reading

Hudson (2007) writes: "Our long-term goal is to have students who do not stop reading when the reading class is over. In order to get there, they have to be allowed to read in a meaningful way" (p. 29). This study shows that even though students were acquainted with the books, the lending procedure and had the possibility to borrow books in the breaks, independently of the reading interventions, this did not happen. When talking to parents about this, they told me that their children actually enjoyed the reading interventions but just did not have enough time to engage in reading after school due to other extra-curricular activities. Combined with Hudson's statement, reading English books did not seem meaningful nor fulfilling enough for them to continue in their free time.

A solution to this could be to combine reading time at school with reading time at home, for example, encouraging students to read a short English book on at least one evening in the week and then to talk to them about their reading experience at school. In this way they will get used to reading English books not only in the classroom but also develop a reading routine at home.

Apart from reading interventions, teachers could also try to integrate purposeful reading classes, so that students are able to experience reading English texts differently – not just for language learning purposes. A possibility is to establish a context were reading an English text is necessary to get relevant information on a topic, for example, planning a school trip, getting information for a project topic in another subject or corresponding with an English-speaking community. In my psychology courses for grade 11 and 12, I always teach one unit with original scientific studies where students have to use their English and math skills to learn something about interesting psychological experiments and their results. Student feedback shows that these units are experienced as difficult but also as the most interesting. Reading in English in this case is a purposeful method to reach a higher aim.

7 Summary and Conclusion

The longitudinal, experimental study presented in this dissertation, analyzes what effects intensive and extensive English reading interventions have on students' English abilities and on their goal orientation and self-beliefs concerning the school subject English. The reading interventions were conducted with 5th and 6th grade students of a German grammar school in a multiple control group design. All experimental classes were taught for 90 minutes per week over a six week period. Tests were conducted directly before and after the intervention (direct pre/posttest) and after a half year period (delayed pre/posttest). The language tests consisted of reading comprehensions tests, cloze tests and prepositions tests. Additionally, the students' English grades and their goal orientation and self-concept scores were surveyed before and after the interventions.

In the summary at the end of chapter two, I presented a reading model that derived from the theoretical assumptions made in that chapter. In general, the results of this study support the theoretical assumptions made and can be summarized as the following (see figure 7.1):

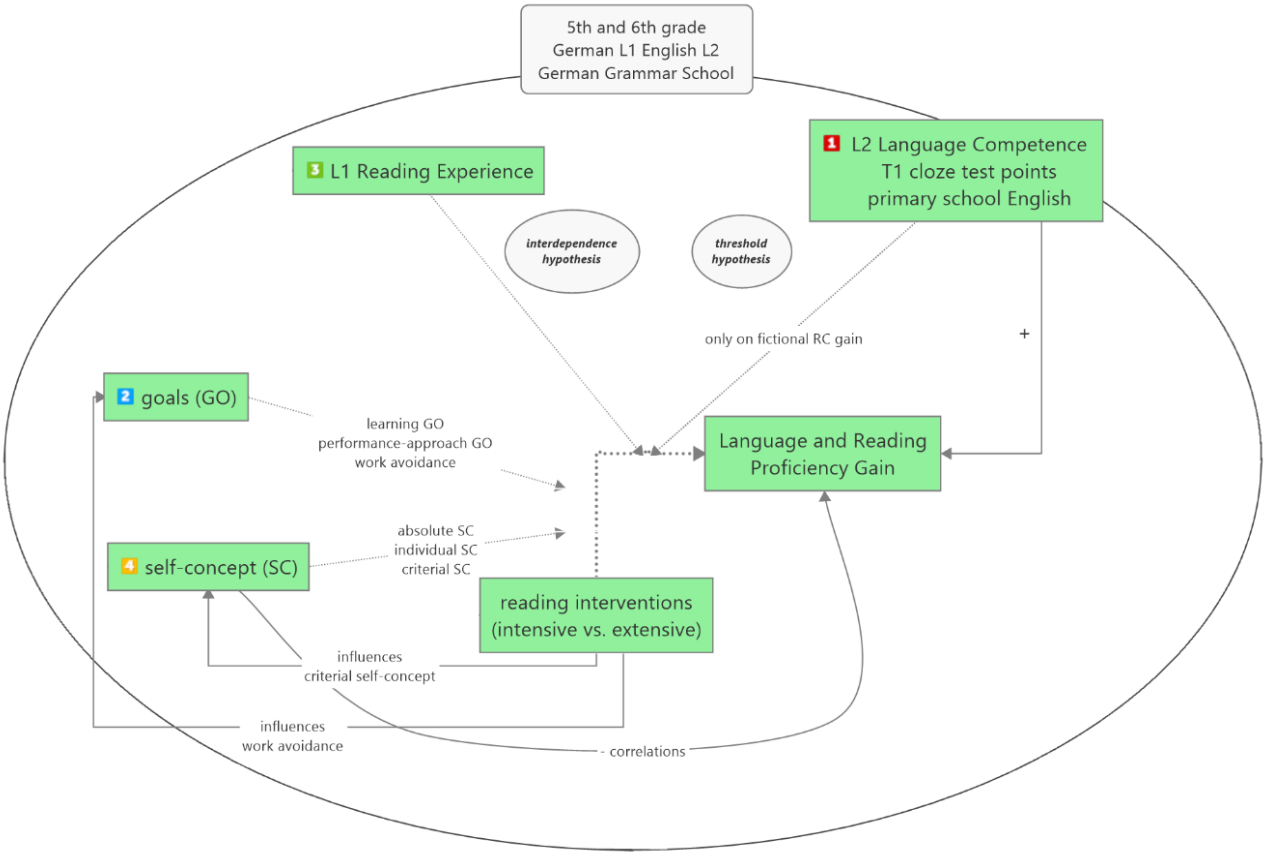


Figure 7.1 Summary Results - Reading Model

L2 language competence influences language proficiency gains directly and is the strongest predictor of competence gains in this study. Students who have high initial cloze test scores and students who have spent more years learning English in primary school have higher competence gains in English language and reading comprehension tests than students with lower cloze tests scores and fewer years of English in primary schools. These effects show for the reading interventions groups, as well as for the non-English interventions (control groups). Thus, there is strong evidence for the assumption that a certain level of L2 competence is necessary to profit from reading interventions (Threshold Hypothesis).

This study did not find any direct effects of the intervention type on improvement in the language tests, but on goal orientation and self-concept. Extensive reading interventions had a positive impact on students criterial self-concept, but also on their work avoidance, whereas intensive reading had a negative impact on students criterial self-concept, but a positive impact on their work avoidance.

As presented and summarized in chapter two, intensive and extensive reading differ significantly on a conceptual level. It was reasoned that intensive or extensive reading, thus, is not an either/or issue. It seemed reasonable to assume that student profit differently from these two intervention types depending on their reading and language experience but also depending on psychological factors like self-beliefs and goals. Indeed, the results of this study suggest – at a very generalized level - that students with a high initial language competence, high L1 reading experience, low learning goal orientation and/or low criterial self-concept profit more from intensive reading interventions and students with a low initial language competence, low L1 reading experience, high learning goal orientation and/or high criterial self-concept from extensive reading interventions. Students who have had more English lessons at their primary school, who have a low performance-approach goal orientation, low work avoidance scores and/or have a low absolute or individual English self-concept, profit more from reading interventions – no matter if intensive or extensive – whereas students with fewer English lessons, a high performance-approach goal orientation, high work avoidance scores and/or a high absolute or individual self-concept profit more from other, non-reading, English interventions or non-English interventions. These moderation effects are also represented in figure 7.1 and discussed in more detail in chapter five.

All in all, these results support the theoretical assumption that, due to the different nature of these two intervention types, students profit very differently depending on their demographic background, motivation, reading experience and language competence. In total the variables in this study and therefore the model presented above explain over 50% of the variance (see section 4.6 for details). The initial language competence – either measured by points achieved in the first test or the years spent learning English in primary school – are the strongest predictors. The next strongest

predictor for the competence gains of the reading intervention classes is work avoidance, followed by self-concept scores and L1 reading experience. These factors should be taken into consideration when implementing reading interventions at school, but also when doing research on the benefit of reading programs.

8 References

Aitken, K. G. (1977). Using Cloze Procedure as an Overall Language Proficiency Test. *TESOL Quarterly*, *11*(1), 59-67. https://doi.org/10.2307/3585592

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs (N.J.): Prentice-Hall.

Alderson, J. C. (1979). The Cloze Procedure and Proficiency in English as a Foreign Language. *TESOL Quarterly*, *13*(2), 219-227. https://doi.org/10.2307/3586211

Al-Homoud, F., & Schmitt, N. (2009). Extensive reading in a challenging environment: a comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, *13*(4), 383–401. https://doi.org/10.1177/1362168809341508

Allan, R. (2008). Can a graded reader corpus provide 'authentic' input? *ELT Journal*, *63*(1), 23–32. https://doi.org/10.1093/elt/ccn011

Balik, K. (1980). *Der Cloze-Test als Messinstrument der Fremdsprachenkompetenz von Schülern mit Deutsch als Muttersprache: eine empirische Studie*. Universität München.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Beile, W. (1996). *Learning English - Green Line: Unterrichtswerk für Gymnasien*. Stuttgart: Klett

Bernhardt, E. (2005). Progress and Procrastination in Second Language Reading. *Annual Review of Applied Linguistics*, *25*, 133–150. https://doi.org/10.1017/S0267190505000073

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press.

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Bong, M., & Skaalvik, E. M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Bredel, U., Fuhrhop, N., & Noack, C. (2017). *Wie Kinder lesen und schreiben lernen* (2nd ed.). Tübingen: Narr Francke Attempto Verlag. Retrieved from http://nbn-resolving.org/urn:nbn:de:bsz:24-epflicht-1436600

Burrows, L. P. (2012). *The effects of extensive reading and reading strategies on reading self-efficacy* (Dissertation). Temple University, Tokyo, Japan. https://doi.org/10.34944/DSPACE/868

Carrell, P., & Carson, J. (1997). Extensive and Intensive Reading in an EAP Setting. *English for Specific Purposes*, *16*(1), 47–60.

Clarke, M. A. (1980). The Short Circuit Hypothesis of ESL Reading-or When Language Competence Interferes with Reading Performance. *The Modern Language Journal*, *64*(2), 203–209. https://doi.org/10.1111/j.1540-4781.1980.tb05186.x

Cummins, J. (1979). Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research*, *49*(2), 222–251. https://doi.org/10.3102/00346543049002222

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). New York: Cambridge University Press.

Cummins, J. (2008). BICS and CALP: Empirical and Theoretical Status of the Distinction. In Street, B. V., & Hornberger, N. (Eds.). *Encyclopedia of Language and Education*, Vol. 2: Literacy. (pp. 71.83). New York: Springer

D. K. Darnell (1970). Clozentropy: A Procedure for Testing English Language Proficiency of Foreign Students. *Speech Monographs, 37(1)*, 36–46. https://doi.org/10.1080/03637757009375647

Daumiller, M., Dickhäuser, O., & Dresel, M. (2019). University instructors' achievement goals for teaching. *Journal of Educational Psychology*, *111*(1), 131–148. https://doi.org/10.1037/edu0000271

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom.* Cambridge, Stuttgart: Cambridge Univ. Press; Klett.

Day, R. R., & Bamford, J. (2002). Top Ten Principles for Teaching Extensive Reading. *Reading in a Foreign Language*, *14*(2), 136–141.

Day, R. R., & Robb, T. (2015). Extensive Reading. In D. Nunan (Ed.), *Language Learning Beyond the Classroom* (pp. 3–12). New York: Routledge.

Derkow Disselbeck, B., & Schwarz, H. (Eds.) (1997). *English G 2000 A1*. Berlin: Cornelsen.

Dolch, E. W., Dolch, M. P., & Patterson, R. (1961). *Once there was a rabbit*. Allen, TX: DLM Teaching Resources.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. https://doi.org/10.1037/0003-066X.41.10.1040

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*. https://doi.org/10.1016/j.cedpsych.2020.101859

Edwards, L. (2014). *Timesaver common European framework assessment tests (A1 - B2)*. *Timesaver photocopiable intermediate*. London, Berlin: Scholastic; Cornelsen.

Elley, W. B. (2000). The potential of book floods for raising literacy levels. *International Review of Education*, *46*(3/4), 233–255. https://doi.org/10.1023/A:1004086618679

Elley, W. B., & Mangubhai, F. (1981). The long-term effects of a book flood on children's language growth. *Directions*, 15–24.

Elley, W. B., & Mangubhai, F. (1983). The Impact of Reading on Second Language Learning. *Reading Research Quarterly*, *19*(1), 53-67. https://doi.org/10.2307/747337

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, *34*(3), 169–189. https://doi.org/10.1207/s15326985ep3403_3

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*(1), 218–232. https://doi.org/10.1037/0022-3514.72.1.218

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*(3), 461–475. https://doi.org/10.1037/0022-3514.70.3.461

Elliot, A. J., & McGregor, H. A. (2001). A 2 X 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3 × 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632–648. https://doi.org/10.1037/a0023952

Evers, A. (2001). The Revised Dutch Rating System for Test Quality. *International Journal of Testing*, *1*(2), 155–182. https://doi.org/10.1207/S15327574IJT0102_4

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/BF03193146

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.

Gellert, A. S., & Elbro, C. (2012). Cloze Tests May be Quick, But Are They Dirty? Development and Preliminary Validation of a Cloze Test of Reading Comprehension. *Journal of Psychoeducational Assessment*, *31*(1), 16–28. https://doi.org/10.1177/0734282912451971

Gillis-Furutaka, A. J. (2015). Graded reader readability. *Journal of Extensive Reading*, *3*.

Goodman, K. S. (1967). Reading: A psycholinquistic quessing game. *Journal of the Reading Specialist*, *6*(1), 126–135.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139150484

Grabe, W. (2014). Key issues in L2 reading development. *Proceedings of the 4th CELC Symposium for English Language Teachers-Selected Papers*, 8–18. Retrieved from https://www.nus.edu.sg/celc/research/books/4th%20symposium%20proceedings/2).%20william%20grabe.pdf

Grabe, W., & Stoller, F. L. (2020). *Teaching and researching reading* (3rd ed.). *Applied linguistics in action*. New York, NY: Routledge.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and Motivation in Reading. In M. L. Kamil (Ed.), *Handbook of reading research*. Mahwah, N.J.: L. Erlbaum Associates.

Hamada, A. (2020). Using Meta-Analysis and Propensity Score Methods to Assess Treatment Effects Toward Evidence-Based Practice in Extensive Reading. *Frontiers in Psychology*, *11*, Article 617. https://doi.org/10.3389/fpsyg.2020.00617

Hardy, J. E. (2016). The effects of a short-term extensive reading course in Spanish. *Journal of Extensive Reading*.

Harger, L., & Schwarz, H. (2013). *English G21 A1* (A1). Berlin: Cornelsen.

Hedgcock, J., & Ferris, D. (2018). *Teaching readers of English: Students, texts, and contexts* (2nd ed.). New York, London: Routledge.

Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion*, *16*(3), 231–247. https://doi.org/10.1007/BF00991653

Hudson, T. (2007). *Teaching Second Language Reading*. *Oxford handbooks for language teachers*. Oxford: Oxford University Press.

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422–449. https://doi.org/10.1037/a0018947

Hunt, R., & Brychta, A. (2015). *The fair-haired samurai*. *Read with Biff, Chip and Kipper stories. Decode and develop: Level 9*. Oxford: Oxford University Press.

Interesting facts about the English language (2020, October 29). *Word Connection*. Retrieved from https://word-connection.com/interesting-facts-about-the-english-language/

International Literacy Association (2021, December 19). Retrieved from https://www.literacyworldwide.org/about-us

James, W. (1890). The consciousness of self. In W. James (Ed.), *The principles of psychology, Vol I* (pp. 291–401). New York: Henry Holt and Co. https://doi.org/10.1037/10538-010

Jeon, E.-Y., & Day, R. R. (2016). The effectiveness of ER on reading proficiency: A meta-analysis. *Reading in a Foreign Language,* 28(2), 256–265.

Jong, M. T. de, & Bus, A. G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. *Journal of Educational Psychology*, *94*(1), 145–155. https://doi.org/10.1037/0022-0663.94.1.145

Khalifa, H., & Weir, C. J. (2009). *Examining reading: Research and practice in assessing second language reading*. *Studies in language testing: Vol. 29*. New York, NY: Cambridge Univ. Press.

Kim, J.-R. (2012). A Meta-study of Extensive English Reading Researches. *English Language & Literature Teaching*, *18*(2), 85–106.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. *Cambridge applied linguistics*. Cambridge: Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9781139524841

Koda, K. (2007). Reading and Language Learning: Crosslinguistic Constraints on Second Language Reading Development. *Language Learning*, *57*, 1–44. https://doi.org/10.1111/0023-8333.101997010-i1

Koolstra, C. M., & van der Voort, T. H. (1996). Longitudinal Effects of Television on Children's Leisure-Time Reading: A Test of Three Explanatory Models. *Human Communication Research*, *23*(1), 4–35. https://doi.org/10.1111/j.1468-2958.1996.tb00385.x

Krashen, S. (2007). Extensive reading in English as a foreign language by adolescents and young adults: A meta-analysis. *The International Journal of Foreign Languages*, *7*(2), 23–29.

Lake, J. W. (2014). Curious Readers and Interesting Reads. *Journal of Extensive Reading*, *2*.

Lieber, R. (2022). *Introducing morphology* (3rd ed.). *Cambridge introductions to language and linguistics*. Cambridge, New York, Port Melbourne, New Delhi, Singapore: Cambridge University Press.

Liu, J., & Zhang, J. (2018). The Effects of Extensive Reading on English Vocabulary Learning: A Meta-analysis. *English Language Teaching*, *11*(6). https://doi.org/10.5539/elt.v11n6p1

Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology*, *111*(2), 331–353. https://doi.org/10.1037/edu0000281

Mason, B. (2004). The effect of adding supplementary writing to an extensive reading program. *International Journal of Foreign Language Teaching*. (1), 2–16.

Maxim, H. H. (1999). *The effects of extensive authentic reading on first-semester German student's reading comprehension, cultural horizon, and language proficiency* (Dissertation). UMI, Ann Arbor, MI, USA.

McLean, S., & Poulshock, J. (2018). Increasing Reading Self-Efficacy and Reading Amount in EFL Learners with Word-Targets. *Reading in a Foreign Language*, *30*(1), 76–91.

Medienpädagogischer Forschungsverbund Südwest (2018). *KIM-Studie 2018: Kindheit, Internet, Medien : Basisuntersuchung zum Medienumgang 6- bis 13-jähriger*. Stuttgart, Baden-Baden: Medienpädagogischer Forschungsverbund Südwest

Medienpädagogischer Forschungsverbund Südwest (2020). *JIM-Studie 2020: Jugend, Information, Medien : Basisuntersuchung zum Medienumgang 12- bis 19-jähriger*. Stuttgart, Baden-Baden: Medienpädagogischer Forschungsverbund Südwest

Metsala, J. L., & Ehri, L. C. (1998). *Word recognition in beginning literacy*. New York, London: Routledge.

Möller, J., & Trautwein, U. (2015). Selbstkonzept. In E. Wild & J. Möller (Eds.), *Pädagogische Psychologie. Springer-Lehrbuch.* (2nd ed., pp. 177–199). Berlin, Heidelberg: Springer

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A Meta-Analysis of Relations Between Achievement and Self-Concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Moosbrugger, H., & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion. Springer-Lehrbuch*. (2nd ed.). Berlin, Heidelberg: Springer-Verlag

Nakanishi, T. (2015). A Meta-Analysis of Extensive Reading Research. *TESOL Quarterly*, *49*(1), 6–37. https://doi.org/10.1002/tesq.157

Nation, P. (2007). The Four Strands. *Innovation in Language Learning and Teaching*, *1*(1), 2–13. https://doi.org/10.2167/illt039.0

Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136-145 Wellington: Victoria University of Wellington Library https://doi.org/10.26686/wgtn.12543404.v1

Nation, P., & Waring, R. (2020). *Teaching extensive reading in another language*. *ESL & applied linguistics professional series*. New York, London: Routledge Taylor & Francis Group.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*(3), 328–346. https://doi.org/10.1037/0033-295X.91.3.328

Oller, J. W. (1973). Cloze test of second language proficiency and what they measure. *Language Learning*, *23*(1), 105–118. https://doi.org/10.1111/j.1467-1770.1973.tb00100.x

Oxford University Press (2020). *Oxford Advanced Learner's Dictionary*. *10*. Oxford: Oxford University Press.

Park, A. Y. (2020). A comparison of the impact of extensive and intensive reading approaches on the reading attitudes of secondary EFL learners. *Studies in Second Language Learning and Teaching*, *10*(2), 337–358. https://doi.org/10.14746/ssllt.2020.10.2.6

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *The Journal of Applied Psychology*, *92*(1), 128–150. https://doi.org/10.1037/0021-9010.92.1.128

Quinault, R. J., & Baumgärtner, A. (1991). *30 Stunden Englisch für Anfänger* (10th ed.). Berlin: Langenscheidt.

Renandya, W. A., Krashen, S., & Jacobs, G. M. (2018). The potential of series books: How narrow reading leads to advanced L2 proficiency. *LEARN Journal*, *11*(2), 148–154.

Renner, E.-M. (2010). *Extensive reading in the lower grades* (Thesis). University of Vienna. https://doi.org/10.25365/THESIS.10133

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2003). Das Fahigkeisselbstkonzept und seine Erfassung. In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), *Tests und Trends: Vol. 2. Diagnostik von Motivation und Selbstkonzept*. Göttingen, Bern: Hogrefe.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2012). *Skalen zur Erfassung des schulischen Selbstkonzepts: SESSKO ; Manual*. Göttingen: Hogrefe.

Schutte, N. S., & Malouff, J. M. (2004). Univerity student reading preferences in realation to the big five inventory dimensions. *Reading Psychology*, *25*(4), 273–295. https://doi.org/10.1080/02702710490522630

Shapiro, J., & Whitney, P. (1997). Factors involved in the leisure reading of upper elementary school students. *Reading Psychology*, *18*(4), 343–370. https://doi.org/10.1080/0270271970180402

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Singer, L. M., & Alexander, P. A. (2017). Reading Across Mediums: Effects of Reading Digital and Print Texts on Comprehension and Calibration. *The Journal of Experimental Education*, *85*(1), 155–172. https://doi.org/10.1080/00220973.2016.1143794

Smith, K. (2006). A comparison of "pure" extensive reading with intensive reading and extensive reading with supplementary activities. *The International Journal of Foreign Language Teaching*. (2).

Spinath, B., & Schöne, C. (2003). Ziele als Bedingung von Motivation am Beispiel der Skalen zur Erfassung der Lern- und Leistungsmotivation (SELLMO). In J. Stiensmeier-Pelster & F. Rheinberg (Eds.), *Tests und Trends: Vol. 2. Diagnostik von Motivation und Selbstkonzept*. Göttingen, Bern: Hogrefe.

Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2012). *Skalen zur Erfassung der Lern- und Leistungsmotivation: SELLMO* (2nd ed.). Göttingen, Bern, Wien: Hogrefe.

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Eds.) (2022). *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (1st ed.). Münster: Waxmann.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.

Taylor, W. L. (1953). "Cloze Procedure": A New Tool for Measuring Readability. *Journalism Quarterly*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

TeeVee (2016). *Test of reading comprehension with answer key*. Kindl eBook.

Thonis, E. W. (1980). *Teaching reading to non-English speakers* (5th ed.). *Collier-MacMillan teacher's library*. New York: Collier MacMillan.

Urguhart, A. H., & Weir, C. J. (2014). *Reading in a Second Language: Process, Product and Practice*. *Applied Linguistics and Language Study*. Hoboken: Taylor and Francis.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Van Schooten, E., & Glopper, K. de (2002). The relation between attitude toward reading adolescent literature and literary reading behavior. *Poetics*, *30*(3), 169–194. https://doi.org/10.1016/S0304-422X(02)00010-4

Van Yperen, N. W., Blaga, M., & Postmes, T. (2014). A meta-analysis of self-reported achievement goals and nonself-report performance across three achievement domains (work, sports, and education). *PloS One*, *9*(4). https://doi.org/10.1371/journal.pone.0093594

Vandewalle, D. (1997). Development and Validation of a Work Domain Goal Orientation Instrument. *Educational and Psychological Measurement*, *57*(6), 995–1015. https://doi.org/10.1177/0013164497057006009

Vandewalle, D., Nerstad, C. G., & Dysvik, A. (2019). Goal Orientation: A Review of the Miles Traveled and the Miles to Go. *Annual Review of Organizational Psychology and Organizational Behavior*, *6*(1), 115–144. https://doi.org/10.1146/annurev-orgpsych-041015-062547

Vansteenkiste, M., Lens, W., Elliot, A. J., Soenens, B., & Mouratidis, A. (2014). Moving the Achievement Goal Approach One Step Forward: Toward a Systematic Examination of the Autonomous and Controlled Reasons Underlying Achievement Goals. *Educational Psychologist*, *49*(3), 153–174. https://doi.org/10.1080/00461520.2014.928598

Verghese, J., Lipton, R. B., Katz, M. J., Hall, C. B., Derby, C. A., Kuslansky, G., . . . Buschke, H. (2003). Leisure activities and the risk of dementia in the elderly. *The New England Journal of Medicine*, *348*(25), 2508–2516. https://doi.org/10.1056/NEJMoa022252

Watanabe, Y., & Koyama, D. (2008). A meta-analysis of second language cloze testing research. *Studies in Second Language Acquisition*, *26*, 103-133.

Webb, S., & Macalister, J. (2013). Is Text Written for Children Useful for L2 Extensive Reading? *TESOL Quarterly*, *47*(2), 300–322. https://doi.org/10.1002/tesq.70

Weir, C. J. (2005). *Language Testing and Validation*. London: Palgrave Macmillan UK.

Weir, C. J., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension. *Research Notes*. (31), 2–10.

Wigfield, & Eccles (2000). Expectancy-Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, *6*(1), 49–78. https://doi.org/10.1007/BF02209024

Williams, R. [Ray] (1986). 'Top ten' principles for teaching reading. *ELT Journal*, *40*(1), 42–45. https://doi.org/10.1093/elt/40.1.42

Williams, R. [Richard] (2019). Ordinal Independant Variables. Retrieved from https://www3.nd.edu/~rwilliam/xsoc73994/OrdinalIndependent.pdf

Yamashita, J. (2004). Reading attitudes in L1 and L2, and their influence on L2 extensive reading. *Reading in a Foreign Language*, *2004*(16), 1–19.

Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, *36*(4), 661–672. https://doi.org/10.1016/j.system.2008.04.003

9 References Library Books

9.1 British and American Childrens Books

9.1.1 Board Books

Pippi the panda (2001). London: Award Publications.

Allen, J. (2008). *I'm not cute!* London: Boxer.

Boynton, S. (1995). *Opposites*. New York: Little Simon.

Boynton, S. (2012). *Hippos go berserk!* New York: Little Simon.

Campbell, R. (2005). *Dear zoo*. London: Campbell.

Carle, E. (1994). *The very hungry caterpillar*. Harlow: Puffin Books.

Deegan, K. (2001). *My first book of opposites*. New York: Bloomsbury Children's Books.

Donaldson, J., & Scheffler, A. (2010). *The Gruffalo*. London: Macmillan.

Donaldson, J., & Scheffler, A. (2011). *Animal actions*. *My first Gruffalo*. London: Macmillan.

Donaldson, J., & Scheffler, A. (2011). *Opposites*. *My first Gruffalo*. London: Macmillan.

Eastman, P. D. (1998). *Are you my mother?* New York: Random House.

Ehlert, L. (2003). *Planting a rainbow*. Orlando, Florida: Red Wagon Books/Harcourt.

Gravett, E. (2014). *Matilda's cat*. New York: Macmillan.

Jones, R. L. (2013). *Look inside things that go*. London: Usborne.

Lionni, L. (2012). *A color of his own*. New York, Enfield: Knopf.

Litton, J. (2016). *Hello world: A celebration of languages and curiosities*. London: Caterpillar Books.

Sherry, K. (2011). *I'm the scariest thing in the castle*. New York: Dial Books.

Tafuri, N. (2013). *The big storm: A very soggy counting book*. New York: Little Simon.

9.1.2 Hard Cover or Paper Backs

Arlon, P., Gordon-Harris, T., & Hood, K. (2015). *The ultimate book of randomly awesome facts*. New York: Scholastic.

Bemelmans, L. (2000). *Madeline in London*. New York: Puffin Books.

Bishop, N. (2012). *Snakes*. New York: Scholastic.

Browne, A. (2003). *The shape game*. London: Doubleday.

Buckley, J., & Gigliotti, J. (2016). *Year in sports 2016*. New York: Scholastic.

Chichester Clark, E. (2015). *Love is my favourite thing: A Plumdog story*. London: Penguin Random House.

DiCamillo, K., McGhee, A., & Fucile, T. (2014). *Bink and Gollie, best friends for ever*. London: Walker Books.

Geisel, T. S. (2013). *The cat in the hat*. New York: Random House.

Lee, C. (2004). *Good dog, Paw*. Cambridge, Mass.: Candlewick Press.

Madonna (2003). *The English roses*. London: Puffin.

Mayer, M. (1993). *Little Critter's read-it-yourself storybook*. *Little Critter*. New York: Golden Books.

McCann, J. L. (2008). *Scooby-Doo! Super spooky double storybook*. New York: Scholastic.

Morse, J. C. (2015). *Book of world records 2015*. New York: Scholastic.

Morse, J. C. (2016). *Book of world records 2016*. New York: Scholastic.

Sewell, A. (1993). *Black Beauty*. London: Award Publications Limited.

Sharratt, N. (2006). *Ketchup on your cornflakes?* London, New York, Toronto: Scholastic.

Sís, P. (2014). *The pilot and the little prince: The life of Antoine de Saint-Exupéry*. London: Pushkin Children's Books.

Smith, W., & Nelson, K. (2005). *Just the two of us*. New York: Scholastic.

Trottier, M., & Reczuch, K. (2000). *Laura's journey: A childhood tale of Laura Secord*. Markham, Ontario: North Winds Press.

Wood, D. (1992). *Old turtle*. Duluth, MN: Pfeifer-Hamilton.

9.1.2.1 Breyer Stablemates Series

Earhart, K., & Andreasen, D. (2006). *Starlight*. *Breyer stablemates*. New York: Scholastic.

Mills, J. E., & Rogers, J. (2007). *Belle*. *Breyer stablemates*. New York: Scholastic.

Mills, J. E., & Rogers, J. (2011). *Stormy*. *Breyer stablemates*. New York: Scholastics.

Weyn, S., & Alba, E. (2009). *Diamond*. *Breyer stablemates*. New York: Scholastic.

Weyn, S., & Rogers, J. (2006). *Snowflake*. *Breyer stablemates*. New York: Scholastic.

9.1.2.2 Dirty Bertie Series

Li, A. (2007). Dirty Bertie: My joke book. Dirty Bertie. London: Stripes.

MacDonald, A. (2012). Dirty Bertie: Pirate! Dirty Bertie. London: Stripes.

MacDonald, A. (2014). Dirty Bertie: Rats! Dirty Bertie. London: Stripes.

MacDonald, A., & Roberts, D. (2006). Dirty Bertie: Worms! Dirty Bertie. London: Stripes.

MacDonald, A., & Roberts, D. (2010). Dirty Bertie: Fangs! Dirty Bertie: Vol. 12. London: Stripes.

MacDonald, A., & Roberts, D. (2011). Dirty Bertie: Kiss! Dirty Bertie. London: Stripes.

MacDonald, A., & Roberts, D. (2012). Dirty Bertie: Scream! Dirty Bertie. London: Stripes.

MacDonald, A., & Roberts, D. (2012). Dirty Berty: Ouch! Dirty Bertie. London: Stripes.

MacDonald, A., & Roberts, D. (2013). Dirty Berty: Dinosaur. Dirty Bertie. London: Stripes.

Roberts, D. (2003). Dirty Bertie. Dirty Bertie. London: Little Tiger Press.

Roberts, D., & MacDonald, A. (2008). Dirty Bertie: Yuck! Dirty Bertie. London: Stripes.

Roberts, D., & MacDonald, A. (2009). Dirty Bertie: Mud! Dirty Bertie. London: Stripes.

Roberts, D., & MacDonald, A. (2013). Dirty Bertie: Zombie! Dirty Bertie: Vol. 21. London: Stripes.

Roberts, D., & MacDonald, A. (2014). Dirty Bertie: Horror! Dirty Bertie: Vol. 24. London: Stripes.

9.1.3 Advanced and Adult Books

Collins, S., Ernst, D., & Kist, H. (2013). *The hunger games*. *Klett Readers*. Stuttgart: Klett.

Dahl, R. (2016). *The BFG*. London: Puffin.

Kinney, J. (2007). *Diary of a wimpy kid*. New York: Amulet Books.

Pilkey, D. (1997). *The adventures of Captain Underpants*. *The adventures of Captain Underpants*. New York: Scholastic.

Pilkey, D. (1999). *Captain Underpants and the invasion of the incredible naughty cafeteria ladies form outer space: and the subsequent assault of the equally evil lunchroom zombie nerds*. *The adventures of Captain Underpants*. New York: Scholastic.

Pilkey, D. (2002). *Captain Underpants and the attack of the talking toilets: Another epic novel*. *The adventures of Captain Underpants*. New York: Scholastic.

Stine, R. L. (2008). *Goosebumps: Night of the living dummy*. *Goosebumps: Vol. 7*. New York: Scholastic.

Stine, R. L. (2010). *Goosebumps: The scarecrow walks at midnight*. *Goosebumps*. New York: Scholastic.

9.1.3.1 Simpson Comics

Groening, M. (1992). *Simpsons comics: A go-go*. *Simpsons comics*. New York: HarperCollins.

Groening, M. (1994). *Simpsons comics: Extravaganza*. *Simpsons comics*. London: Titan.

Groening, M. (1995). *Simpsons comics: Spectacular*. *Simpsons comics*. New York: Titan.

Groening, M. (1996). *Simpsons comics: Simps-o-rama*. *Simpsons comics*. London: Titan Books.

Groening, M. (1996). *Simpsons comics: Strike back! Simpsons comics*. London: Titan Books.

Groening, M. (1997). *Simpson comics: Featuring Bartman: Best of the best*. *Simpsons comics*. London: Titan Books.

Groening, M. (1997). *Simpsons comics: Wingding*. *Simpsons comics*. London: Titan.

Groening, M. (1999). *Simpson comics: Big bonanza*. *Simpsons comics*. London: Titan Books.

Groening, M. (2001). *Simpson comics: Royale*. *Simpsons comics*. London: Titan Books.

Groening, M. (2002). *Simpsons comics: Unchained*. *Simpsons comics*. London: Titan Books.

Groening, M. (2005). *Simpson Comics: Jam-packed jamboree*. *Simpsons comics*. London: Titan Books.

Groening, M. (2005). *Simpsons comics: Barn burner*. *Simpsons comics*. London: Titan.

Groening, M. (2006). *Big bouncy book of Bart Simpson*. *Simpsons comics*. London: Titan.

Groening, M. (2007). *Simpson comics: Beach blanket bongo*. *Simpsons comics*. London: Titan Books.

Groening, M. (2008). *Big brilliant book of Bart Simpson*. *Simpsons comics*. London: Titan Books.

Groening, M. (2013). *Simpsons comics: Madness*. *Simpsons comics*. London: Titan Books.

Groening, M. (2014). *Simpson Comics: Shake-up*. *Simpsons comics*. London: Titan Books.

Richmond, R. (e.) (1997). *The Simpsons: A complete guide to your favorite family*. *Simpsons comics*. London: HarperCollins.

9.2 Easy Reader Series aimed at L1 Readers

9.2.1 Read with Biff, Chip and Kipper Series (Oxford Reading Tree)

9.2.1.1 Level 1

Hunt, R., & Brychta, A. (2014). *A good trick*. *Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). *Floppy did this! Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). *Get on*. *Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). *Six in a bed*. *Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). *The Pancake*. *Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). *Up you go*. *Read with Biff, Chip and Kipper. First Stories: Level 1*. Oxford: Oxford University Press.

9.2.1.2 Level 2

Hunt, R., & Brychta, A. (2014). Dad's Birthday. Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). Funny Fish. Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). Mum's new hat. Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). Picnic Time. Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). Silly Races! Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). The snowman. Read with Biff, Chip and Kipper. First Stories: Level 2. Oxford: Oxford University Press.

9.2.1.3 Level 3

Hunt, R., & Brychta, A. (2011). I can trick a tiger. Read with Biff, Chip and Kipper. First stories: Level 3. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Poor old rabbit. Read with Biff, Chip and Kipper. First stories: Level 3. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Super dad. Read with Biff, Chip and Kipper. First stories: Level 3. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2014). Floppy and the bone. Read with Biff, Chip and Kipper. First stories: Level 3. Oxford: Oxford University Press.

9.2.1.4 Level 4

Hunt, R., & Brychta, A. (2011). *Dragon danger*. *Read with Biff, Chip and Kipper. First stories: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *Missing! Read with Biff, Chip and Kipper. First stories: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *Quick! Quick! Read with Biff, Chip and Kipper. Phonics: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *The moon jet*. *Read with Biff, Chip and Kipper. Phonics: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *The raft race*. *Read with Biff, Chip and Kipper. First stories: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *The red coat*. *Read with Biff, Chip and Kipper. Phonics: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *The spaceship*. *Read with Biff, Chip and Kipper. First stories: Level 4*. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). *Wet feet*. *Read with Biff, Chip and Kipper. Phonics: Level 4*. Oxford: Oxford University Press.

9.2.1.5 Level 5

Hunt, R., & Alex, B. (2011). Craig saves the day. Read with Biff, Chip and Kipper. Phonics: Level 5. Oxford: Oxford University Press.

Hunt, R., & Alex, B. (2011). Husky adventure. Read with Biff, Chip and Kipper. First Stories: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Dolphin rescue. Read with Biff, Chip and Kipper. Phonics: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Egg fried rice. Read with Biff, Chip and Kipper. Phonics: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Hungry Floppy. Read with Biff, Chip and Kipper. First Stories: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Looking after Gran. Read with Biff, Chip and Kipper. First Stories: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Seasick. Read with Biff, Chip and Kipper. Phonics: Level 5. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Trapped! Read with Biff, Chip and Kipper. First Stories: Level 5. Oxford: Oxford University Press.

9.2.1.6 Level 6

Hunt, R., & Brychta, A. (2011). Gran's new blue shoes. Read with Biff, Chip and Kipper. Phonics: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Hairy-scary monster. Read with Biff, Chip and Kipper. First Stories: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Ice city. Read with Biff, Chip and Kipper. Phonics: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Mountain rescue. Read with Biff, Chip and Kipper. First Stories: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Save Pudding Wood. Read with Biff, Chip and Kipper. Phonics: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Secret of the sands. Read with Biff, Chip and Kipper. First Stories: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). The lost voice. Read with Biff, Chip and Kipper. First Stories: Level 6. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2011). Uncle Max. Read with Biff, Chip and Kipper. Phonics: Level 6. Oxford: Oxford University Press.

9.2.1.7 Level 7

Hunt, R., & Brychta, A. (2015). A tall tale. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). Detective adventure. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). Holiday in Japan. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). Magic tricks. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The portrait problem. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The time capsule. Read with Biff, Chip and Kipper stories. Decode and develop: Level 7. Oxford: Oxford University Press.

9.2.1.8 Level 8

Hunt, R., & Brychta, A. (2015). A good turn. Read with Biff, Chip and Kipper. Decode and Develop: Level 8. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). A lucky find. Read with Biff, Chip and Kipper. Decode and Develop: Level 8. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The beehive fence. Read with Biff, Chip and Kipper. Decode and Develop: Level 8. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The ogre's dinner. Read with Biff, Chip and Kipper. Decode and Develop: Level 8. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The secret pop star. Read with Biff, Chip and Kipper. Decode and Develop: Level 8. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The strange old house. Read with Biff, Chip and Kipper stories. Decode and develop: Level 8. Oxford: Oxford University Press.

9.2.1.9 Level 9

Hunt, R., & Brychta, A. (2015). A knight in town. Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). Fireball in the sky. Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). Princes in the tower. Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The fair-haired samurai. Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). The travelling players. Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

Hunt, R., & Brychta, A. (2015). What a Journey! Read with Biff, Chip and Kipper stories. Decode and develop: Level 9. Oxford: Oxford University Press.

9.2.2 Read it yourself with Ladybird Series

9.2.2.1 Level 1

Cinderella (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Fairy friends (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Goldilocks and the three bears (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Little creatures (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Little red hen (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Peppa Pig: Recycling fun! (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Rex the big dinosaur (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The emperor's new clothes (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The enormous turnip (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The magic porridge pot (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The princess and the pea (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The tale of Peter Rabbit (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The three billy goats Gruff (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

The ugly duckling (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Tinga Tanga Tales: Why giraffe has a long neck (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

Topsy and Tim: Go to the zoo (2015). Read it yourself with Ladybird: Level 1. London: Ladybird Books.

9.2.2.2 Level 2

Beauty and the beast (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Chicken licken (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Dom's dragon (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Little red riding hood (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Peppa Pig: Sports Day (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Peppa Pig: Nature trail (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Pirate school (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Rumpelstiltskin (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Sleeping beauty (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Sly fox and red hen (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

The gingerbread man (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

The three little pigs (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Tinga Tanga Tales: The tale of jemima puddle-duck (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Topsy and Tim: The big race (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Town mouse and country mouse (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Why lion roarrrs! (2015). Read it yourself with Ladybird: Level 2. London: Ladybird Books.

Horsley, L. (1997). My dad can't dance. Read with Ladybird: Level 2: Book 3. Leicestershire: Ladybird.

9.2.2.3 Level 3

Charlie and Lola: You won't like this present as much as I do! (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Hansel and Gretel (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Harry and the bucketful of dinosaurs (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Jack and the beanstalk (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Moshi Monsters: Poppet Stows Away (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Moshi Monsters: Furi and the music island (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

Rapunzel (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

The eleves and the shoemaker (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

The red knight (2015). Read it yourself with Ladybird: Level 3. London: Ladybird Books.

9.2.2.4 Level 4

Charlie and Lola: I am inventing an invention (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Harry and the dinosaurs united (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Heidi (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Moshi Monsters: Luvli and the glump-a-tron (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Moshi Monsters: Katsuma and the art thief (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Sam and the robots (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

Snow white and the seven dwarfs (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

The pied piper of hamelin (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

The wizard of Oz (2015). Read it yourself with Ladybird: Level 4. London: Ladybird Books.

9.2.3 Scholastic Readers

9.2.3.1 Level 1

Burgess, M., & Hall, L. (2005). *Billy Elliot*. *Scholastic readers. Level 1*. London: Scholastic.

9.2.3.2 Level 2

Arnold, T. (2013). *Fly Guy presents: Space*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2014). *Fly Guy presents: Firefighters*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2014). *Fly Guy presents: Dinosaurs*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2015). *Fly Guy presents: Insects*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2015). *Fly Guy presents: Bats*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2016). *Fly Guy presents: The White House*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2016). *Fly Guy presents: Snakes*. *Fly Guy presents*. New York: Scholastic.

Arnold, T. (2016). *Fly Guy presents: Weather*. *Fly Guy presents*. New York: Scholastic.

Berger, M., & Berger, G. (2014). *Ugly cute animals*. *Scholastic reader. Level 2*. New York: Scholastic.

Carson, M. K., & Bracken, C. (2014). *The magic school bus presents ocean adventure*. *The magic school bus*. New York: Scholastic.

Cole, J., & Degen, B. (1995). *The magic school bus inside the earth*. *The magic school bus*. New York: Scholastic.

Cole, J., & Degen, B. (2014). *The magic school bus presents planet earth*. *The magic school bus*. New York: Scholastic.

Emerson, J. (2013). *Stinky bugs*. *Scholastic reader. Level 2*. New York: Scholastic.

Emerson, J. (2015). *Snow babies*. *Scholastic reader. Level 2*. New York: Scholastic.

9.2.3.3 Level 4

Bourke, A., Rendall, J., & Revell, J. (2016). *A lion called Christian*. *Scholastic readers*. London: Scholastic.

9.2.4 Scholastic Discover More Series

9.2.4.1 For brand-new readers

Arlon, P., & Gordon-Harris, T. (2014). *Dolphins*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2015). *Animal faces*. *Scholastic discover more*. New York: Scholastic.

Buckley, J. (2013). *Things that go! Scholastic discover more readers*. New York: Scholastic.

9.2.4.2 For emergent readers

Arlon, P., & Gordon-Harris, T. (2012). *See me grow*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2013). *Puppies and kittens*. *Scholastic discover more*. New York: Scholastic.

Pinnington, A., & Gordon-Harris, T. (2012). *Animal babies*. *Scholastic discover more*. New York, NY: Scholastic.

9.2.4.3 For confident readers

Arlon, P. (2014). *Ancient Egypt*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2013). *Rainforest*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2013). *Reptiles*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2013). *Weather*. *Scholastic discover more*. New York: Scholastic.

Arlon, P., & Gordon-Harris, T. (2015). *Explorers*. *Scholastic discover more*. New York: Scholastic.

Burnie, D. (2013). *Sharks*. *Scholastic discover more*. New York: Scholastic.

Hayes, S., & Gordon-Harris, T. (2014). *Polar Animals*. *Scholastic discover more*. New York: Scholastic.

9.2.5 Scholastic "I am" Series

Norwich, G., & Alba, E. (2013). *I am George Lucas*. *Scholastic I am*. New York: Scholastic.

9.2.6 Scholastics Solo Reading Series

Carroll, J., & Barrett, V. (1997). *Lots and lots of chicken pox*. *Scholastic solo reading*. Markham, Ontario: Scholastic.

Cummings, P., & Cox, D. (1997). *The great Jimbo James*. *Scholastic solo reading*. Markham, Ont.: Scholastic.

Kropp, P., & Lesynski, L. (2002). *What a story! Scholastic solo reading*. New York, Toronto: Scholastic.

Rodda, E., & Jellett, T. (1999). *Fuzz, the famous fly*. *Scholastic solo reading*. New York, Toronto: Scholastic.

Walker, K., & Smith, C. (2000). *Sticky stuff*. *Scholastic solo reading*. Markham, Ont.: Scholastic.

9.2.7 Scholastics Questions and Answers Series

Berger, M., Berger, G., & Bond, H. (2000). Do tornadoes really twist? Questions and answers about tornadoes and hurricanes. Scholastic question and answer. New York: Scholastic.

Berger, M., Berger, G., & Di Vincent, F. (2000). Can you hear a shout in space? Questions and answers about space exploration. Scholastic questions and answer. New York: Scholastic.

Berger, M., Berger, G., & Effler, J. M. (2001). Where have all the pandas gone? Questions and answers about endangered species. Scholastic question and answer. New York: Scholastic.

Berger, M., Berger, G., & Rice, J. (2000). What makes an ocean wave? Questions and answers about oceans and ocean life. Scholastic question and answer. New York: Scholastic.

9.2.8 Hello Reader! Science

Hopping, L. J., & Wheeler, J. (1993). Wild weather.: Tornadoes! Hello reader! New York: Cartwheel Books.

9.2.9 News Non-fiction Readers

Kennedy, M. M. (2009). The story of the White House. Scholastic news non-fiction readers. New York: Children's Press.

Miller, A. (2009). What is Air Force One? Scholastic news non-fiction readers. New York: Children's Press.

9.2.10 National Geographic Kids First Big Book of Series

Carney, E. (2015). Little kids first big book of the world. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Esbaum, J. (2016). Little kids first big book of how. National Geografic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Hughes, C. (2010). Little kids first big book of animals. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Hughes, C. (2011). Little kids first big book of dinosaurs. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Hughes, C. (2012). Little kids first big book of space. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Hughes, C. (2014). Little kids first big book of bugs. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Hughes, C. (2017). First big book of the ocean. National Geographic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

Schields, A. (2011). Little kids first big book of why. National Geografic Kids First Big Book of Series. Washington D.C.: National Geographic Kids.

9.2.11 Usborne Very First Reading Series

Mackinnon, M., & Simpson, S. (2010). A bus for Miss Moss. Usborne very first reading: Vol. 3. London: Usborne.

9.2.12 O'Brien Flyers Series

McGann, O. (2006). Mad Grandad and the mutant river. Flyer: Vol. 14. Dublin: O'Brien.

9.2.13 And I Can Read Books

Dakos, K. (2000). *The bug in teacher's coffee and other school poems*. New York: HarperCollins World.

9.2.14 Graphic Library

Lassieur, A., Frenz, R., & Barnett, C. (2007). *Lord of the sea: The vikings explore the north atlantic*. Mankato, Minnesota: Capstone Press.

9.3 Easy Reader Series aimed at L2 Readers

9.3.1 Diesterweg First Readers

9.3.1.1 1. Lernjahr

Koerner, A. F., & Marckwort, U. (2015). *Ruby's story*. *First readers*. Braunschweig: Diesterweg.

Koerner, A. F., & Naumann, A. (2015). *The giant called Ed*. *First readers*. Braunschweig: Diesterweg.

Schade, D. (2002). *A knife in the back*. *First readers*. Frankfurt am Main: Diesterweg.

Taylor, J. (1999). *Perfect holiday*. *First readers*. Frankfurt am Main: Diesterweg.

9.3.1.2 2. Lernjahr

Fine, A. (2005). *Loudmouth Louis*. *First readers*. Braunschweig: Diesterweg.

Taylor, J. (1998). *The secret of the lost dogs*. *First readers*. Frankfurt am Main: Diesterweg.

Taylor, J. (1999). *Perfect present*. *First readers*. Frankfurt am Main: Diesterweg.

9.3.1.3 3. Lernjahr

Taylor, J. (1998). *The secret of the cemetery*. *First readers*. Frankfurt am Main: Diesterweg.

Taylor, J. (1999). *Perfect love*. *First readers*. Frankfurt am Main: Diesterweg.

9.3.2 Diesterweg Readers

9.3.2.1 Starter

Wilde, O. (2006). *The Canterville Ghost*. *Diesterweg Readers. Starter Level*. Slough, Braunschweig: Richmond Publ; Diesterweg.

9.3.2.2 Level 1

Chatwin, B. (2006). *Robin Hood*. *Diesterweg Readers. Level 1*. Slough, Braunschweig: Richmond Publ; Diesterweg.

Chatwin, B. (2006). *Titanic*. *Diesterweg Readers. Level 1*. Slough, Braunschweig: Richmond Publ; Diesterweg.

Dickens, C. (2006). *A Christmas Carol*. *Diesterweg Readers. Level 1*. Slough, Braunschweig: Richmond Publ; Diesterweg.

Jerome, J. K. (2006). *Ghost stories*. *Diesterweg Readers. Level 1*. Slough, Braunschweig: Richmond Publ; Diesterweg.

Robinson, I. (2006). *The haunted castle*. *Diesterweg Readers. Level 1*. Slough, Braunschweig: Richmond Publ; Diesterweg.

9.3.3 Hueber Lektüren

9.3.3.1 Level 1 Grade 5

Bean, J. (2006). The young riders. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

Bowring, J. (2007). Cousins and crocodiles. Hueber Lektüren Englisch. Level 1. Ismaning: Huber.

Bowring, J. (2008). Lost in the Rain. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

Craig, C., & Botté, E. (2006). The fast lane. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

Francis, P. (2006). Adventure in the Alps. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

Francis, P. (2007). The football match. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

McLean, A. C. (2009). The right team. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

Murray, S. (2007). Grizzly. Hueber Lektüren Englisch. Level 1. Ismaning: Hueber.

9.3.3.2 Level 2 Grade 6

Francis, P. (2008). Pirates! Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

Kirby, D. (2007). Between the flags. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

Murray, S. (2006). Blue Moon beach. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

Murray, S. (2010). The Show Must Go On. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

O'Carolan, P. (2006). Ride for your life. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

O'Carolan, P. (2007). The bride wore black. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

Smith, P. (2006). Shopping for trouble. Hueber Lektüren Englisch. Level 2. Ismaning: Hueber.

9.3.3.3 Level 3 Grade 7

Botté, E. (2006). In Sarah's dreams. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

Kirby, D., & Grantford, J. (2006). The key. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

O'Carolan, P. (2005). Stranger danger. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

O'Carolan, P. (2007). Double trouble. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

O'Carolan, P. (2007). The man next door. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

O'Carolan, P., & Botté, E. (2011). What Rose saw. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

Voysey, P. (2006). The rainbow girl. Hueber Lektüren Englisch. Level 3. Ismaning: Hueber.

9.3.3.4 Level 4 Grade 8

O'Carolan, P. (2007). The Alati collection. Hueber Lektüren Englisch. Level 4. Ismaning: Hueber.

O'Carolan, P., & Botté, E. (2008). Road to nowhere. Hueber Lektüren Englisch. Level 4. Ismaning: Hueber.

9.3.4 Cornelsen English Library

9.3.4.1 Schuljahr 5

Level 1

Hewitt, P. (2006). *Kim and the rainbow*. *Cornelsen English Library*. Berlin: Cornelsen Verlag.

Hoppenstedt, G. (2012). *Help me find Henry! Cornelsen English Library*. Berlin: Cornelsen.

Level 2

Gallagher, J. (2012). Adventure on the Hoe. Cornelsen English Library. Berlin: Cornelsen.

Taylor, C. (2001). Merlin's magnificent magic shop. Cornelsen English Library. Berlin: Cornelsen.

Woppert, A. F. (2006). The Case of the Corner Shop Robbers. Cornelsen English Library. Berlin: Cornelsen.

Level 3

Gallagher, J. (2014). *The golden ticket*. *Cornelsen English Library*. Berlin: Cornelsen Verlag.

9.3.4.2 Schuljahr 6

Level 1

Annis, L. (2008). Kisses and cokes. Cornelsen English Library. Berlin: Cornelsen.

Berning, J. (2003). Sharon's first case. Cornelsen English Library. Berlin: Cornelsen.

Inman, C. (2006). Harry's dog. Cornelsen English Library. Berlin: Cornelsen.

Level 2

Derkow Disselbeck, B. (2001). *Indian adventure*. *Cornelsen English Library*. Berlin: Cornelsen.

Harger, L. (2009). *A trip to New Zealand*. *Cornelsen English Library*. Berlin: Cornelsen.

Lamsdale, C. (2013). *Who is Henry Kazwell? Cornelsen English Library*. Berlin: Cornelsen.

Seidl, J. (2006). *The adventures of Mr D*. *Cornelsen English Library*. Berlin: Cornelsen.

Stewart, P. (2008). *An elephant never forgets*. *Cornelsen English Library*. Berlin: Cornelsen.

Level 3

DiCamillo, K. (2014). Because of Winn-Dixie. Cornelsen English Library. Berlin: Cornelsen.

Klingmann, A.-M., Erlemann, S., & Hinrichs, A. (2011). A viking musical. Cornelsen English Library. Berlin: Cornelsen.

9.3.4.3 Schuljahr 7

La Mare, C. de (2013). *Time flyer*. *Cornelsen English Library*. Berlin: Cornelsen.

9.3.5 Klett English Readers

Edwards, L. (2017). *The Royal Family* (1. Auflage). Stuttgart: Klett.

Shipton, V. (2012). *New Orleans fried chicken*. *English Readers + DVD*. Stuttgart: Klett Sprachen.

9.3.5.1 Lernjahr 1/2

Baer-Engel, J. (2007). *The four detectives*. *Klett English Reader*. Stuttgart, Leipzig: Klett.

Himmelstrup, K., & Stenzel, K. (2009). *Uncle Bill's will*. *Klett English Reader*. Stuttgart: Klett.

Posener, A. (2009). *It's a dog's life*. *Klett English Reader*. Stuttgart: Klett.

Posener, A. (2009). *Late again! Klett English Reader*. Stuttgart: Klett.

Sefton, C. (2011). *The ghost and Bertie Boggin*. *Klett English Reader*. Stuttgart: Klett.

Warner, K. L. (2008). *The secret of the island*. *Klett English Reader*. Stuttgart: Klett.

9.3.5.2 Lernjahr 2/3

Hellyer-Jones, R. (2013). The adventures of King Arthur and his knights of the round table: And his knights of the round table. Klett English Reader. Stuttgart, Leipzig: Klett.

Roth, R. W. (2012). Lost in the USA. Klett English Reader. Stuttgart, Leipzig: Klett.

Shipton, V. (2012). Hamburgers in Texas: [A2 ; Lehrprogramm gemäß §14 JuSchG]. Klett English Reader. Stuttgart: Klett.

Speight, S. (2009). The birdwatcher. Klett English Reader. Stuttgart: Klett.

9.3.5.3 Klasse 5

Level 1

Fuller, G. (2007). Tom's adventures in Catland and other stories. Klett English Reader. Stuttgart, Leipzig: Klett.

Hellyer-Jones, R. (2012). Barker's world. Klett English Reader. Stuttgart, Leipzig: Klett.

Level 2

Ashworth, P. (2017). Ships, pearls and a parrot. Klett English Reader. Stuttgart: Klett.

McBride, S. (2007). Pets 4 U. Klett English Reader. Stuttgart, Leipzig: Klett.

Level 3

Aziz, H. (2016). *The wildest party ever! Klett English Reader*. Stuttgart, Leipzig: Klett.

Baer-Engel, J. (2016). *Stories for a ghost*. *Klett English Reader*. Stuttgart, Leipzig: Klett.

Lambert, D. (2012). *Fenimore Castle: A mystery for the stage*. *Klett English Reader*. Stuttgart, Leipzig: Klett.

Weisshaar, H. (2007). *Sir Fox Bones: And the Buckingham Palace Mytery*. *Klett English Reader*. Stuttgart: Klett.

9.3.5.4 Klasse 6

Ashworth, P. (2010). Ok José or how José saved the day. Klett English Reader. Stuttgart: Klett.

McBride, S. (2010). The gift of the gab: A radio play from Ireland. Klett English Reader. Stuttgart: Klett.

9.3.5.5 English Graphic Novels (ab Klasse 7)

Fermer, D. (2012). *King Arthur*. *English graphic novels*. Stuttgart: Klett.

Fermer, D. (2013). *Sir Francis Drake*. *English graphic novels*. Stuttgart: Klett.

Fermer, D. (2016). *Guy Fawkes*. *English graphic novels*. Stuttgart: Klett.

Fermer, D., & Pflügner, M. (2013). *Robin Hood*. *English graphic novels*. Stuttgart: Klett.

9.3.6 Black Cat Cideb Reading and Training (grade 7 and higher)

Clemen, G. D. B. (2016). *Alien Alert in Seattle*. *Reading & training*. De Agostini Scuola S.p.A.: Black Cat CIDEB.

Clemen, G. D. B. (2016). *Miami Police File: The O'Nell Case*. *Reading & training*. De Agostini Scuola S.p.A.: Black Cat CIDEB.

Doyle, A. C. (2016). *Sherlock Holmes Stories*. *Reading & training*. De Agostini Scuola S.p.A.: Black Cat CIDEB.

Hutchinson, A. M. (2016). *Missing in Sydney*. *Reading & training*. De Agostini Scuola S.p.A.: Black Cat CIDEB.

James, M. R. (2015). *Stories for the Curious*. *Reading & training*. De Agostini Scuola S.p.A.: Black Cat CIDEB.

9.3.7 Easy Readers (ER/TR) Series

9.3.7.1 Easy Readers

Herriot, J. (2010). *If only they could talk*. *Easy Readers*. Copenhagen: Easy Readers.

Hinton, S. E., Keson, B.-K., & Illum, P. (2014). *The outsiders*. *Easy Readers*. Copenhagen: Easy Readers.

James, P. (2019). *The perfect murder*. *Easy Readers*. Copenhagen: Easy Readers.

Stanton, A. (2013). *Sterling and the Canary*. *Easy Readers*. Copenhagen: Easy Readers.

9.3.7.2 Teen Readers

Ferro, C. (2009). *The Cave Mystery*. *Teen readers*. Copenhagen: Easy Readers.

Heneghan, J. (2003). *Hit squad*. *Teen readers*. Copenhagen: Easy Readers.

Hewitt, P. (2010). *Secrets*. *Teen readers*. Copenhagen: Easy Readers.

Weatherly, L. (2008). *Them*. *Teen readers*. Copenhagen: Easy Readers.

9.3.8 Macmillan Readers

Cox, A. (2005). *Dangerous journey*. *Macmillan Readers*. Oxford: Macmillan.

Leroux, G. (2005). *The phantom of the opera*. *Macmillan Readers*. Oxford: Macmillan.

Milne, J. (2005). *The long tunnel*. *Macmillan Readers*. Oxford: Macmillan.

9.3.9 ELI Readers

9.3.9.1 Ready to Read

Staiano, E. (1997). The three little pigs. ELI readers. Ready to read. Recanati: ELI.

9.3.9.2 Young ELI Readers

Baum, L. F., Cadwallader, J., & Mazali, G. (2012). *The wonderful wizard of Oz*. *Young ELI Readers*. Recanati: ELI.

Cadwallader, J. (2014). *Uncle Jack in the Amazon Rainforest: Buch mit Audio-CD. Englische Lektüre für das 1. und 2. Lernjahr*. *Young ELI Readers*. Recanati: ELI.

Carroll, L., & Brown, R. (2009). *Alice in Wonderland*. *Young ELI Readers*. Recanati: ELI.

Staiano, E. (1997). *The three little pigs*. *Young ELI Readers*. Recanati: ELI.

9.3.9.3 Teen ELI Readers

Chaucer, G. (2014). *The Canterbury Tales*. *Teen ELI readers*. Recanati: ELI.

Dickens, C., & Celija, M. (2010). *Oliver Twist*. *Teen ELI readers*. Recanati: ELI.

Ferretti, L. (2014). *Dear diary*. *Teen ELI readers*. Recanati: ELI.

Flagan, M. (2010). *The egyptian souvenir*. *Teen ELI readers*. Recanati: ELI.

Porter, E. H. (2016). *Pollyanna: Englische Lektüre für das 1. und 2. Lernjahr. Buch + Audio-CD*. *Teen ELI readers*. Recanati: ELI.

Tomkinson, A. (2014). *Loving London*. *Teen ELI readers*. Recanati: ELI.

Tomkinson, A. (2015). *Great friends! Teen ELI readers*. Recanati: ELI.

9.3.10 Ravensburg Blue Bird Stories

Brosche, H., Schulte, S., & Bonk, H. (2003). *Judy and the dog*. *Ravensburger Blauer Rabe*. Ravensburg: Ravensburger.

9.3.11 STARK Lektüre Englisch

Cartmell, A. (1999). *Harry Hedgehog*. Freising: Stark.

Cartmell, A. (1999). *The green feather*. Freising: Stark.

9.3.12 Cambridge Readers

9.3.12.1 Starter/Beginner

Johnson, M. (2010). *Big hair day*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2011). *Little trouble in california*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Prowse, P. (1999). *Arman's Journey*. *Cambridge English readers*. Cambridge: Cambridge University Press.

9.3.12.2 Level 1

Brennan, F. (2006). *Three tomorrows*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2010). *A Little Trouble in Dublin*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Prowse, P. (2005). *Don't stop now! Cambridge English readers*. Cambridge: Cambridge University Press.

9.3.12.3 Level 2

Alan Battersby (2012). *The Dark Side of the City*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Harmer, J. (2005). *Within high fences*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Johnson, M. (2011). *Ask Alice*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Johnson, M. (2011). *New Zealand*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Johnson, M., & Mansfield, K. (2010). *Parties and presents: Three short stories*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2009). *Little trouble in amsterdam*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2011). *Bad Company*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2016). *The New Zealand file*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Rollason, J. (2009). *Killer bees*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Tomlinson, B. (2018). *Superbird*. *Cambridge English readers*. Cambridge: Cambridge University Press.

9.3.12.4 Level 3

Loader, M. (2003). *Eye of the storm*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2009). *A little Trouble in the Yorkshire Dales*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2009). *Scotland*. *Cambridge English readers*. Cambridge: Cambridge University Press.

MacAndrew, R. (2010). *Not above the law*. *Cambridge English readers*. Cambridge: Cambridge University Press.

Rollason, J. (2009). *Tales of terror: Edgar Allan Poe and others*. *Cambridge English readers*. Cambridge: Cambridge University Press.

9.3.13 Lifelike Shortstories

Wunder, B. (2013). *Lifelike shortstories: Part 2*. Augsburg: Brigg Pädagogik.

Wunder, B. (2013). *Lifelike shortstories: Part 1*. Augsburg: Brigg Pädagogik.

Wunder, B. (2013). *Lifelike Shortstories: Part 3*. Augsburg: Brigg Pädagogik.

9.3.14 Penguin Readers

Falkner, J. M., & Strange, J. (2008). *Moonfleet*. *Penguin readers*. Harlow, Essex: Pearson.

9.3.15 Helbling Readers People

Beddall, F. (2013). Steve Jobs and the Story of Apple, Class Set. Helbling Readers Non-Fiction. Rum: Helbling.

9.3.16 PONS Lektüre

Butler, D. (2016). PONS Murder in the fog: Mörderische Kurzkrimis zum Englischlernen. PONS Lektüre. Stuttgart: PONS.

Butler, D. (2016). To die for: Mörderische Kurzkrimis zum Englischlernen. PONS Lektüre. Stuttgart: PONS.

10 Appendix

10.1 Ranking Books Borrowed

Books Borrowed	Titel	Series	Publisher
30	The Pancake	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
29	The raft race	Read with Biff, Chip and Kipper. First stories: Level 4	Oxford University Press
28	Picnic Time	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
28	Silly Races!	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
28	Six in a bed	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
27	Dad's Birthday	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
27	The snowman	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
26	Floppy and the bone	Read with Biff, Chip and Kipper. First stories: Level 3	Oxford University Press
25	Poor old rabbit	Read with Biff, Chip and Kipper. First stories: Level 3	Oxford University Press
24	Missing!	Read with Biff, Chip and Kipper. First stories: Level 4	Oxford University Press
24	Super dad	Read with Biff, Chip and Kipper. First stories: Level 3	Oxford University Press
24	The spaceship	Read with Biff, Chip and Kipper. First stories: Level 4	Oxford University Press
23	I can trick a tiger	Read with Biff, Chip and Kipper. First stories: Level 3	Oxford University Press
23	The red coat	Read with Biff, Chip and Kipper. Phonics: Level 4	Oxford University Press
21	Wet feet	Read with Biff, Chip and Kipper. Phonics: Level 4	Oxford University Press
20	Dragon danger	Read with Biff, Chip and Kipper. First stories: Level 4	Oxford University Press
20	Up you go	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
16	Dolphin rescue	Read with Biff, Chip and Kipper. Phonics: Level 5	Oxford University Press
15	Heidi	Read it yourself with Ladybird	Ladybird Books
15	Belle	Breyer stablemates	Scholastic
14	Snow white and the seven dwarfs	Read it yourself with Ladybird	Ladybird Books
14	Starlight	Breyer stablemates	Scholastic
14	Simpsons comics Strike back!	Simpsons comics	Titan Books
14	A good trick	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
14	Book of world records 2016		Scholastic
13	Little red riding hood	Read it yourself with Ladybird	Ladybird Books
13	Love is my favourite thing A Plumdog story		Penguin Random House
13	Looking after Gran	Read with Biff, Chip and Kipper. First Stories: Level 5	Oxford University Press
13	Funny Fish	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
12	Peppa Pig Recycling fun!	Read it yourself with Ladybird	Ladybird Books
12	Rapunzel	Read it yourself with Ladybird	Ladybird Books
12	Rex the big dinosaur	Read it yourself with Ladybird	Ladybird Books
12	Fly Guy presents Space	Fly Guy presents	Scholastic
12	Seasick	Read with Biff, Chip and Kipper. Phonics: Level 5	Oxford University Press
12	Trapped!	Read with Biff, Chip and Kipper. First Stories: Level 5	Oxford University Press
12	Stormy	Breyer stablemates	Scholastics
12	Animal babies	Scholastic discover more	Scholastic

12	Ketchup on your cornflakes?		Scholastic
12	Diamond	Breyer stablemates	Scholastic
11	Beauty and the beast	Read it yourself with Ladybird	Ladybird Books
11	Cinderella	Read it yourself with Ladybird	Ladybird Books
11	Peppa Pig Nature trail	Read it yourself with Ladybird	Ladybird Books
11	The ugly duckling	Read it yourself with Ladybird	Ladybird Books
11	The Gruffalo		Macmillan
11	Husky adventure	Read with Biff, Chip and Kipper. First Stories: Level 5	Oxford University Press
11	Hungry Floppy	Read with Biff, Chip and Kipper. First Stories: Level 5	Oxford University Press
11	Ice city	Read with Biff, Chip and Kipper. Phonics: Level 6	Oxford University Press
11	Quick! Quick!	Read with Biff, Chip and Kipper. Phonics: Level 4	Oxford University Press
11	Mum's new hat	Read with Biff, Chip and Kipper. First Stories: Level 2	Oxford University Press
10	Little creatures	Read it yourself with Ladybird	Ladybird Books
10	Peppa Pig Sports Day	Read it yourself with Ladybird	Ladybird Books
10	The princess and the pea	Read it yourself with Ladybird	Ladybird Books
10	Fly Guy presents Weather	Fly Guy presents	Scholastic
10	Simpsons comics Madness	Simpsons comics	Titan Books
10	Mountain rescue	Read with Biff, Chip and Kipper. First Stories: Level 6	Oxford University Press
10	Save Pudding Wood	Read with Biff, Chip and Kipper. Phonics: Level 6	Oxford University Press
10	Floppy did this!	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
10	Magic tricks	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
10	The portrait problem	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
10	The strange old house	Read with Biff, Chip and Kipper stories. Decode and develop: Level 8	Oxford University Press
10	What a Journey!	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
10	Snowflake	Breyer stablemates	Scholastic
9	Hansel and Gretel	Read it yourself with Ladybird	Ladybird Books
9	Jack and the beanstalk	Read it yourself with Ladybird	Ladybird Books
9	Moshi Monsters Luvli and the glump-a-tron	Read it yourself with Ladybird	Ladybird Books
9	Rumpelstiltskin	Read it yourself with Ladybird	Ladybird Books
9	Sleeping beauty	Read it yourself with Ladybird	Ladybird Books
9	The wizard of Oz	Read it yourself with Ladybird	Ladybird Books
9	Topsy and Tim: Go to the zoo	Read it yourself with Ladybird	Ladybird Books
9	Snakes		Scholastic
9	Sharks	Scholastic discover more	Scholastic
9	Animal actions	My first Gruffalo	Macmillan
9	Simpsons comics Unchained	Simpsons comics	Titan Books
9	Simpson comics Beach blanket bongo	Simpsons comics	Titan Books
9	Craig saves the day	Read with Biff, Chip and Kipper. Phonics: Level 5	Oxford University Press
9	Gran's new blue shoes	Read with Biff, Chip and Kipper. Phonics: Level 6	Oxford University Press

9	Hairy-scary monster	Read with Biff, Chip and Kipper. First Stories: Level 6	Oxford University Press
9	The moon jet	Read with Biff, Chip and Kipper. Phonics: Level 4	Oxford University Press
9	Blue Moon beach	Hueber Lektüren Englisch. Level 2	Hueber
9	Dirty Bertie	Dirty Bertie	Little Tiger Press
8	Little red hen	Read it yourself with Ladybird	Ladybird Books
8	The red knight	Read it yourself with Ladybird	Ladybird Books
8	The three little pigs	Read it yourself with Ladybird	Ladybird Books
8	Town mouse and country mouse	Read it yourself with Ladybird	Ladybird Books
8	Puppies and kittens	Scholastic discover more	Scholastic
8	Reptiles	Scholastic discover more	Scholastic
8	Hippos go berserk!		Little Simon
8	Secret of the sands	Read with Biff, Chip and Kipper. First Stories: Level 6	Oxford University Press
8	Uncle Max	Read with Biff, Chip and Kipper. Phonics: Level 6	Oxford University Press
8	Get on	Read with Biff, Chip and Kipper. First Stories: Level 1	Oxford University Press
8	The beehive fence	Read with Biff, Chip and Kipper. Decode and Develop: Level 8	Oxford University Press
8	Book of world records 2015		Scholastic
8	The adventures of Captain Underpants	The adventures of Captain Underpants	Scholastic
7	Moshi Monsters Furi and the music island	Read it yourself with Ladybird	Ladybird Books
7	Moshi Monsters Katsuma and the art thief	Read it yourself with Ladybird	Ladybird Books
7	The gingerbread man	Read it yourself with Ladybird	Ladybird Books
7	The pied piper of hamelin	Read it yourself with Ladybird	Ladybird Books
7	Dolphins	Scholastic discover more	Scholastic
7	The ultimate book of randomly awesome facts		Scholastic
7	The young riders	Hueber Lektüren Englisch. Level 1	Hueber
7	Judy and the dog	Ravensburger Blauer Rabe	Ravensburger
7	Year in sports 2016		Scholastic
7	Little kids first big book of the world	National Geographic Kids First Big Book of Series	National Geographic Kids
7	The fast lane	Hueber Lektüren Englisch. Level 1	Hueber
7	Are you my mother?		Random House
7	Big bouncy book of Bart Simpson	Simpsons comics	Titan
7	My dad can't dance	Read with Ladybird: Level 2	Ladybird
7	Egg fried rice	Read with Biff, Chip and Kipper. Phonics: Level 5	Oxford University Press
7	The lost voice	Read with Biff, Chip and Kipper. First Stories: Level 6	Oxford University Press
7	A lucky find	Read with Biff, Chip and Kipper. Decode and Develop: Level 8	Oxford University Press
7	Holiday in Japan	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
7	Diary of a wimpy kid		Amulet Books
7	Scooby-Doo! Super spooky double storybook		Scholastic

6	Dom's dragon	Read it yourself with Ladybird	Ladybird Books
6	Moshi Monsters Poppet Stows Away	Read it yourself with Ladybird	Ladybird Books
6	The eleves and the shoemaker	Read it yourself with Ladybird	Ladybird Books
6	The emperor's new clothes	Read it yourself with Ladybird	Ladybird Books
6	The tale of Peter Rabbit	Read it yourself with Ladybird	Ladybird Books
6	Fly Guy presents Firefighters	Fly Guy presents	Scholastic
6	Fly Guy presents Snakes	Fly Guy presents	Scholastic
6	Lost in the Rain	Hueber Lektüren Englisch. Level 1	Hueber
6	Dear zoo		Campbell
6	Bink and Gollie, best friends for ever		Walker Books
6	Opposites	My first Gruffalo	Macmillan
6	Robin Hood	English graphic novels	Klett
6	Adventure in the Alps	Hueber Lektüren Englisch. Level 1	Hueber
6	The cat in the hat		Random House
6	Simpsons comics Spectacular	Simpsons comics	Titan
6	Simpsons comics Simps-o-rama	Simpsons comics	Titan Books
6	Simpson comics Big bonanza	Simpsons comics	Titan Books
6	Simpson Comics Jam-packed jamboree	Simpsons comics	Titan Books
6	Simpsons comics Barn burner	Simpsons comics	Titan
6	Simpson Comics Shake-up	Simpsons comics	Titan Books
6	Little kids first big book of bugs	National Geographic Kids First Big Book of Series	National Geographic Kids
6	A tall tale	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
6	Detective adventure	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
6	Fireball in the sky	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
6	The secret pop star	Read with Biff, Chip and Kipper. Decode and Develop: Level 8	Oxford University Press
6	Look inside things that go		Usborne
6	Dirty Bertie Pirate!	Dirty Bertie	Stripes
6	Dirty Bertie Scream!	Dirty Bertie	Stripes
6	A bus for Miss Moss	Usborne very first reading	Usborne
6	The right team	Hueber Lektüren Englisch. Level 1	Hueber
6	A knife in the back	First readers	Diesterweg
6	I'm the scariest thing in the castle		Dial Books
5	Pippi the panda		Award Publications
5	Chicken licken	Read it yourself with Ladybird	Ladybird Books

5	Goldilocks and the three bears	Read it yourself with Ladybird	Ladybird Books
5	Sam and the robots	Read it yourself with Ladybird	Ladybird Books
5	Sly fox and red hen	Read it yourself with Ladybird	Ladybird Books
5	The magic porridge pot	Read it yourself with Ladybird	Ladybird Books
5	Ancient Egypt	Scholastic discover more	Scholastic
5	Animal faces	Scholastic discover more	Scholastic
5	Fly Guy presents Bats	Fly Guy presents	Scholastic
5	Fly Guy presents The White House	Fly Guy presents	Scholastic
5	The shape game		Doubleday
5	The football match	Hueber Lektüren Englisch. Level 1	Hueber
5	Simpsons comics Wingding	Simpsons comics	Titan
5	Polar Animals	Scholastic discover more	Scholastic
5	Little kids first big book of dinosaurs	National Geographic Kids First Big Book of Series	National Geographic Kids
5	Little kids first big book of space	National Geographic Kids First Big Book of Series	National Geographic Kids
5	Princes in the tower	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
5	Good dog, Paw		Candlewick Press
5	Hello world A celebration of languages and curiosities		Caterpillar Books
5	Dirty Bertie Kiss!	Dirty Bertie	Stripes
5	What is Air Force One?	Scholastic news non-fiction readers	Children's Press
5	The long tunnel	Macmillan Readers	Macmillan
4	Charlie and Lola You won't like this present as much as I do!	Read it yourself with Ladybird	Ladybird Books
4	The three billy goats Gruff	Read it yourself with Ladybird	Ladybird Books
4	Tinga Tanga Tales Why giraffe has a long neck	Read it yourself with Ladybird	Ladybird Books
4	Topsy and Tim: The big race	Read it yourself with Ladybird	Ladybird Books
4	I'm not cute!		Boxer
4	Rainforest	Scholastic discover more	Scholastic
4	The very hungry caterpillar		Puffin Books
4	Snow babies	Scholastic reader. Level 2	Scholastic
4	The Cave Mystery	Teen readers	Easy Readers
4	Pirates!	Hueber Lektüren Englisch. Level 2	Hueber
4	Big brilliant book of Bart Simpson	Simpsons comics	Titan Books
4	Little kids first big book of animals	National Geographic Kids First Big Book of Series	National Geographic Kids
4	First big book of the ocean	National Geographic Kids First Big Book of Series	National Geographic Kids

4	A good turn	Read with Biff, Chip and Kipper. Decode and Develop: Level 8	Oxford University Press
4	A knight in town	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
4	The fair-haired samurai	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
4	The ogre's dinner	Read with Biff, Chip and Kipper. Decode and Develop: Level 8	Oxford University Press
4	What a story!	Scholastic solo reading	Scholastic
4	The English roses		Puffin
4	Grizzly	Hueber Lektüren Englisch. Level 1	Hueber
4	Captain Underpants and the invasion of the incredible naughty cafeteria ladies form outer space and the subsequent assault of the equally evil lunchroom zombie nerds	The adventures of Captain Underpants	Scholastic
4	Dirty Bertie	Dirty Bertie	Stripes
4	Little kids first big book of why	National Geografic Kids First Big Book of Series	National Geographic Kids
4	The big storm A very soggy counting book		Little Simon
3	Harry and the bucketful of dinosaurs	Read it yourself with Ladybird	Ladybird Books
3	Pirate school	Read it yourself with Ladybird	Ladybird Books
3	The enormous turnip	Read it yourself with Ladybird	Ladybird Books
3	Tinga Tanga Tales The tale of jemima puddle-duck	Read it yourself with Ladybird	Ladybird Books
3	Tinga Tanga Tales Why lion roarrrs!	Read it yourself with Ladybird	Ladybird Books
3	See me grow	Scholastic discover more	Scholastic
3	Explorers	Scholastic discover more	Scholastic
3	The wildest party ever!	Klett English Reader	Klett
3	The four detectives	Klett English Reader	Klett
3	Steve Jobs and the Story of Apple, Class Set	Helbling Readers Non-Fiction	Helbling
3	Ugly cute animals	Scholastic reader. Level 2	Scholastic
3	Cousins and crocodiles	Hueber Lektüren Englisch. Level 1	Huber
3	Things that go!	Scholastic discover more readers	Scholastic
3	To die for Mörderische Kurzkrimis zum Englischlernen	PONS Lektüre	PONS
3	The magic school bus presents ocean adventure	The magic school bus	Scholastic
3	Alien Alert in Seattle	Reading & training	Black Cat CIDEB

3	The magic school bus inside the earth	The magic school bus	Scholastic
3	The bug in teacher's coffee and other school poems		HarperCollins World
3	My first book of opposites		Bloomsbury Children's Books
3	Sherlock Holmes Stories	Reading & training	Black Cat CIDEB
3	Planting a rainbow		Red Wagon Books/Harcourt
3	Sir Francis Drake	English graphic novels	Klett
3	Matilda's cat		Macmillan
3	Simpsons comics A go-go	Simpsons comics	HarperCollins
3	Simpson comics Featuring Bartman: Best of the best	Simpsons comics	Titan Books
3	Help me find Henry!	Cornelsen English Library	Cornelsen
3	Wild weather. Tornadoes!	Hello reader!	Cartwheel Books
3	The time capsule	Read with Biff, Chip and Kipper stories. Decode and develop: Level 7	Oxford University Press
3	The giant called Ed	First readers	Diesterweg
3	The phantom of the opera	Macmillan Readers	Macmillan
3	Dirty Bertie Worms!	Dirty Bertie	Stripes
3	Dirty Bertie Fangs!	Dirty Bertie	Stripes
3	The Show Must Go On	Hueber Lektüren Englisch. Level 2	Hueber
3	Captain Underpants and the attack of the talking toilets Another epic novel	The adventures of Captain Underpants	Scholastic
3	It's a dog's life	Klett English Reader	Klett
3	The Simpsons A complete guide to your favorite family	Simpsons comics	HarperCollins
3	Dirty Bertie Horror!	Dirty Bertie	Stripes
3	Lost in the USA	Klett English Reader	Klett
3	Black Beauty		Award Publications Limited
3	Goosebumps Night of the living dummy	Goosebumps	Scholastic
3	Loving London	Teen ELI readers	ELI
3	Laura's journey A childhood tale of Laura Secord		North Winds Press
3	Sir Fox Bones And the Buckingham Palace Mytery	Klett English Reader	Klett
2	Charlie and Lola I am inventing an invention	Read it yourself with Ladybird	Ladybird Books
2	Fairy friends	Read it yourself with Ladybird	Ladybird Books

2	Harry and the dinosaurs united	Read it yourself with Ladybird	Ladybird Books
2	Fly Guy presents Dinosaurs	Fly Guy presents	Scholastic
2	Fly Guy presents Insects	Fly Guy presents	Scholastic
2	The wonderful wizard of Oz	Young ELI Readers	ELI
2	Opposites		Little Simon
2	PONS Murder in the fog Mörderische Kurzkrimis zum Englischlernen	PONS Lektüre	PONS
2	Uncle Jack in the Amazon Rainforest	Young ELI Readers	ELI
2	Little kids first big book of how	National Geografic Kids First Big Book of Series	National Geographic Kids
2	King Arthur	English graphic novels	Klett
2	Simpson comics Royale	Simpsons comics	Titan Books
2	Barker's world	Klett English Reader	Klett
2	Uncle Bill's will	Klett English Reader	Klett
2	The travelling players	Read with Biff, Chip and Kipper stories. Decode and develop: Level 9	Oxford University Press
2	Stories for the Curious	Reading & training	Black Cat CIDEB
2	The key	Hueber Lektüren Englisch. Level 3	Hueber
2	Dirty Bertie My joke book	Dirty Bertie	Stripes
2	A color of his own		Knopf
2	A Little Trouble in Dublin	Cambridge English readers	Cambridge University Press
2	Dirty Bertie Rats!	Dirty Bertie	Stripes
2	Little Critter's read-it-yourself storybook	Little Critter	Golden Books
2	I am George Lucas	Scholastic I am	Scholastic
2	Ride for your life	Hueber Lektüren Englisch. Level 2	Hueber
2	The Alati collection	Hueber Lektüren Englisch. Level 4	Hueber
2	Pollyanna	Teen ELI readers	ELI
2	Late again!	Klett English Reader	Klett
2	Tales of terror Edgar Allan Poe and others	Cambridge English readers	Cambridge University Press
2	Hamburgers in Texas	Klett English Reader	Klett
2	The three little pigs	ELI readers. Ready to read	ELI
2	An elephant never forgets	Cornelsen English Library	Cornelsen
1	The Dark Side of the City	Cambridge English readers	Cambridge University Press
1	Weather	Scholastic discover more	Scholastic
1	Ok José or how José saved the day	Klett English Reader	Klett
1	Madeline in London		Puffin Books
1	Do tornadoes really twist?	Scholastic question and answer	Scholastic

1	Where have all the pandas gone?	Scholastic question and answer	Scholastic
1	A lion called Christian	Scholastic readers	Scholastic
1	Lots and lots of chicken pox	Scholastic solo reading	Scholastic
1	The green feather		Stark
1	Titanic	Diesterweg Readers. Level 1	Richmond Publ; Diesterweg
1	The Canterbury Tales	Teen ELI readers	ELI
1	Miami Police File: The O'Nell Case	Reading & training	Black Cat CIDEB
1	The magic school bus presents planet earth	The magic school bus	Scholastic
1	The hunger games	Klett Readers	Klett
1	The great Jimbo James	Scholastic solo reading	Scholastic
1	The BFG		Puffin
1	The Royal Family		Klett
1	Loudmouth Louis	First readers	Diesterweg
1	Tom's adventures in Catland and other stories	Klett English Reader	Klett
1	Adventure on the Hoe	Cornelsen English Library	Cornelsen
1	Simpsons comics Extravaganza	Simpsons comics	Titan
1	Hit squad	Teen readers	Easy Readers
1	Secrets	Teen readers	Easy Readers
1	Missing in Sydney	Reading & training	Black Cat CIDEB
1	Harry's dog	Cornelsen English Library	Cornelsen
1	The perfect murder	Easy Readers	Easy Readers
1	Big hair day	Cambridge English readers	Cambridge University Press
1	The story of the White House	Scholastic news non-fiction readers	Children's Press
1	Between the flags	Hueber Lektüren Englisch. Level 2	Hueber
1	Ruby's story	First readers	Diesterweg
1	Fenimore Castle	Klett English Reader	Klett
1	Who is Henry Kazwell?	Cornelsen English Library	Cornelsen
1	Lord of the sea The vikings explore the north atlantic		Capstone Press
1	Eye of the storm	Cambridge English readers	Cambridge University Press
1	Little trouble in amsterdam	Cambridge English readers	Cambridge University Press
1	Not above the law	Cambridge English readers	Cambridge University Press
1	Little trouble in california	Cambridge English readers	Cambridge University Press
1	Dirty Berty Ouch!	Dirty Bertie	Stripes
1	Mad Grandad and the mutant river	Flyer	O'Brien
1	Stranger danger	Hueber Lektüren Englisch. Level 3	Hueber

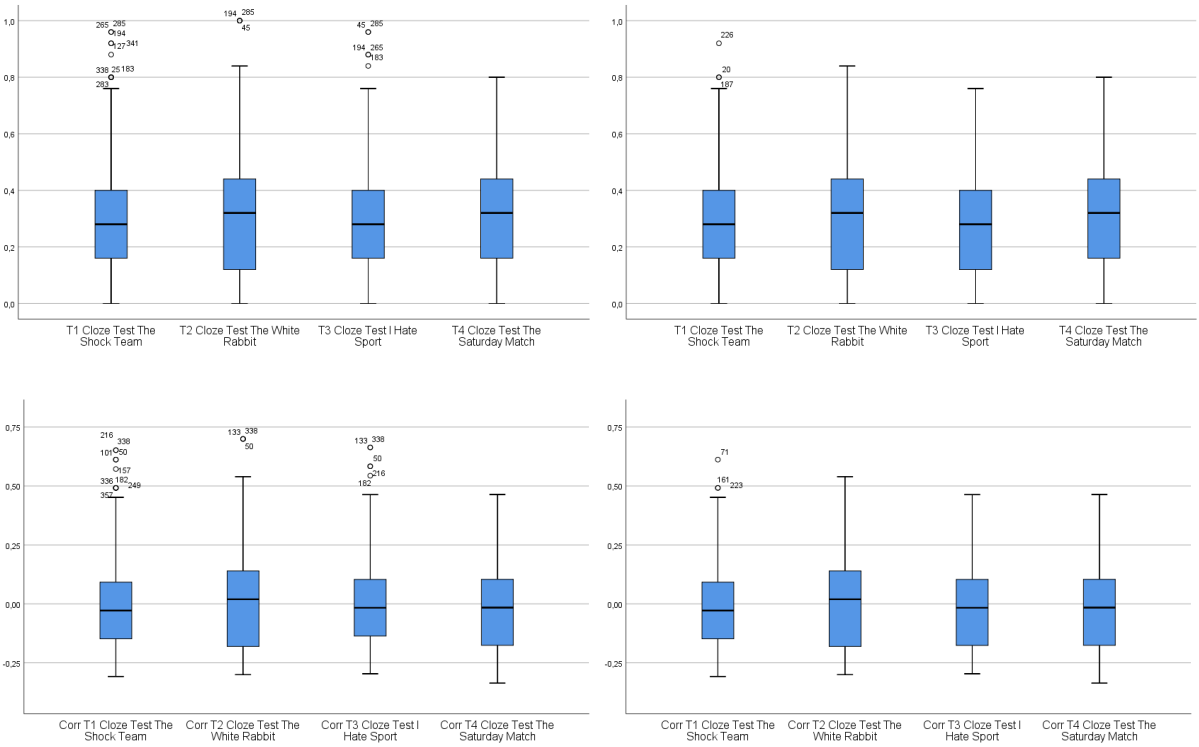
1	Double trouble	Hueber Lektüren Englisch. Level 3	Hueber
1	The bride wore black	Hueber Lektüren Englisch. Level 2	Hueber
1	The man next door	Hueber Lektüren Englisch. Level 3	Hueber
1	What Rose saw	Hueber Lektüren Englisch. Level 3	Hueber
1	Dirty Bertie Yuck!	Dirty Bertie	Stripes
1	Dirty Bertie Mud!	Dirty Bertie	Stripes
1	The ghost and Bertie Boggin	Klett English Reader	Klett
1	New Orleans fried chicken	English Readers + DVD	Klett Sprachen
1	Shopping for trouble	Hueber Lektüren Englisch. Level 2	Hueber
1	Just the two of us		Scholastic
1	The birdwatcher	Klett English Reader	Klett
1	Goosebumps The scarecrow walks at midnight	Goosebumps	Scholastic
1	Merlin's magnificent magic shop	Cornelsen English Library	Cornelsen
1	The secret of the cemetery	First readers	Diesterweg
1	Perfect holiday	First readers	Diesterweg
1	Great friends!	Teen ELI readers	ELI
1	The rainbow girl	Hueber Lektüren Englisch. Level 3	Hueber
1	The secret of the island	Klett English Reader	Klett
1	Old turtle		Pfeifer-Hamilton
0	Kisses and cokes	Cornelsen English Library	Cornelsen
0	Ships, pearls and a parrot	Klett English Reader	Klett
0	Stories for a ghost	Klett English Reader	Klett
0	Can you hear a shout in space?	Scholastic questions and answer	Scholastic
0	What makes an ocean wave?	Scholastic question and answer	Scholastic
0	Sharon's first case	Cornelsen English Library	Cornelsen
0	In Sarah's dreams	Hueber Lektüren Englisch. Level 3	Hueber
0	Three tomorrows	Cambridge English readers	Cambridge University Press
0	Billy Elliot	Scholastic readers. Level 1	Scholastic
0	Alice in Wonderland	Young ELI Readers	ELI
0	Harry Hedgehog		Stark
0	Robin Hood	Diesterweg Readers. Level 1	Richmond Publ; Diesterweg
0	Dangerous journey	Macmillan Readers	Macmillan
0	Indian adventure	Cornelsen English Library	Cornelsen
0	Because of Winn-Dixie	Cornelsen English Library	Cornelsen
0	A Christmas Carol	Diesterweg Readers. Level 1	Richmond Publ; Diesterweg
0	Oliver Twist	Teen ELI readers	ELI
0	Stinky bugs	Scholastic reader. Level 2	Scholastic
0	Moonfleet	Penguin readers	Pearson
0	Guy Fawkes	English graphic novels	Klett

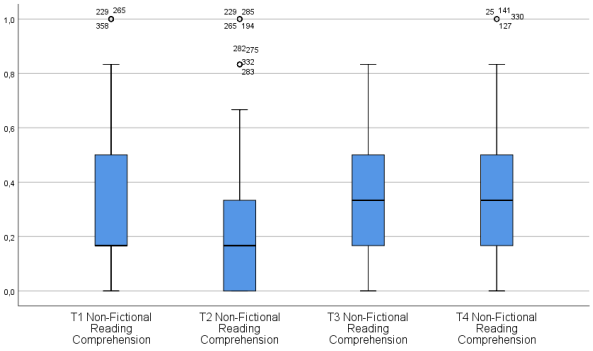
0	Dear diary	Teen ELI readers	ELI
0	The egyptian souvenir	Teen ELI readers	ELI
0	The golden ticket	Cornelsen English Library	Cornelsen Verlag
0	A trip to New Zealand	Cornelsen English Library	Cornelsen
0	Within high fences	Cambridge English readers	Cambridge University Press
0	The adventures of King Arthur and his knights of the round table And his knights of the round table	Klett English Reader	Klett
0	If only they could talk	Easy Readers	Easy Readers
0	Kim and the rainbow	Cornelsen English Library	Cornelsen Verlag
0	The outsiders	Easy Readers	Easy Readers
0	Ghost stories	Diesterweg Readers. Level 1	Richmond Publ; Diesterweg
0	Ask Alice	Cambridge English readers	Cambridge University Press
0	New Zealand	Cambridge English readers	Cambridge University Press
0	Parties and presents	Cambridge English readers	Cambridge University Press
0	A viking musical	Cornelsen English Library	Cornelsen
0	Time flyer	Cornelsen English Library	Cornelsen
0	A little Trouble in the Yorkshire Dales	Cambridge English readers	Cambridge University Press
0	Scotland	Cambridge English readers	Cambridge University Press
0	Bad Company	Cambridge English readers	Cambridge University Press
0	The New Zealand file	Cambridge English readers	Cambridge University Press
0	Dirty Berty Dinosaur	Dirty Bertie	Stripes
0	Pets 4 U	Klett English Reader	Klett
0	The gift of the gab	Klett English Reader	Klett
0	Road to nowhere	Hueber Lektüren Englisch. Level 4	Hueber
0	Arman's Journey	Cambridge English readers	Cambridge University Press
0	Don't stop now!	Cambridge English readers	Cambridge University Press
0	The haunted castle	Diesterweg Readers. Level 1	Richmond Publ; Diesterweg
0	Fuzz, the famous fly	Scholastic solo reading	Scholastic
0	Killer bees	Cambridge English readers	Cambridge University Press
0	The adventures of Mr D	Cornelsen English Library	Cornelsen
0	The pilot and the little prince		Pushkin Children's Books
0	Sterling and the Canary	Easy Readers	Easy Readers
0	The secret of the lost dogs	First readers	Diesterweg
0	Perfect love	First readers	Diesterweg
0	Perfect present	First readers	Diesterweg

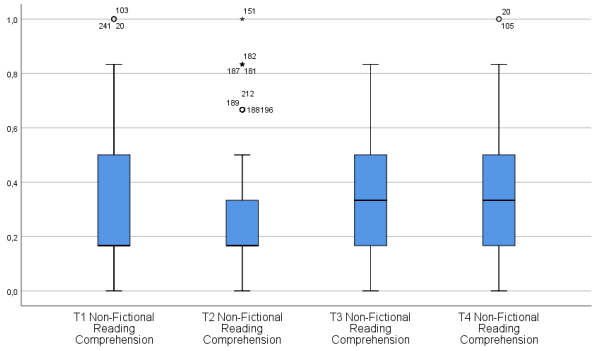
0	Superbird	Cambridge English readers	Cambridge University Press
0	Sticky stuff	Scholastic solo reading	Scholastic
0	Them	Teen readers	Easy Readers
0	The Canterville Ghost	Diesterweg Readers. Starter Level	Richmond Publ; Diesterweg
0	The Case of the Corner Shop Robbers	Cornelsen English Library	Cornelsen
0	Lifelike shortstories Part 2		Brigg Pädagogik
0	Lifelike shortstories Part 1		Brigg Pädagogik
0	Lifelike Shortstories Part 3		Brigg Pädagogik

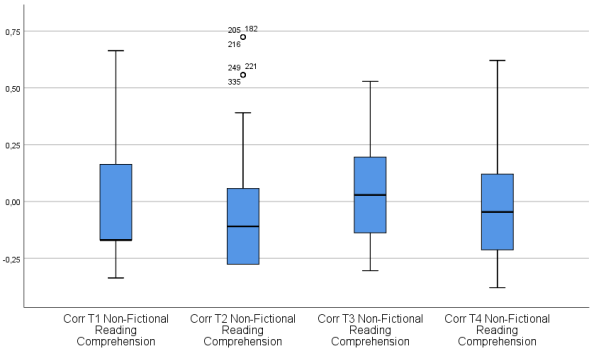
10.2 Meeting Criteria for Analysis

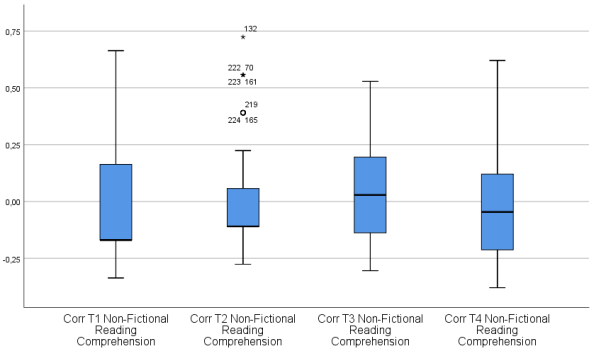
10.2.1 Outliers

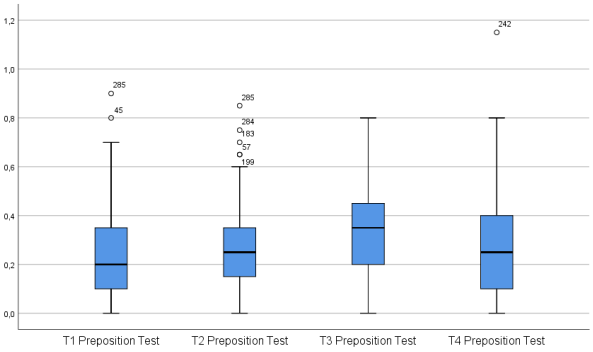


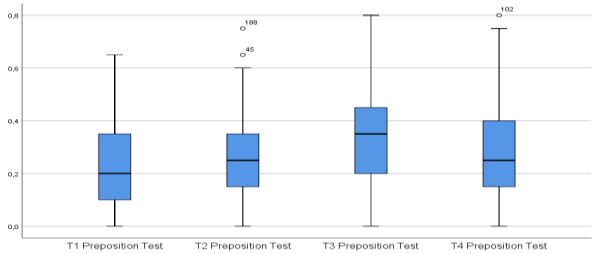

T1 Non-Fictional Reading Comprehension
T2 Non-Fictional Reading Comprehension
T3 Non-Fictional Reading Comprehension
T4 Non-Fictional Reading Comprehension

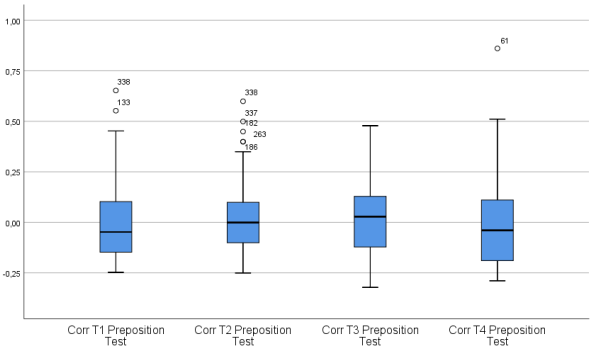

T1 Non-Fictional Reading Comprehension
T2 Non-Fictional Reading Comprehension
T3 Non-Fictional Reading Comprehension
T4 Non-Fictional Reading Comprehension

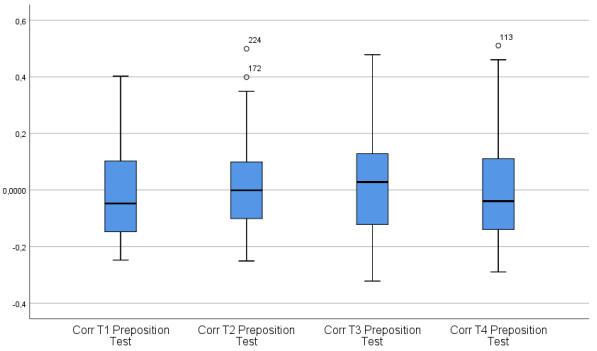

Corr T1 Non-Fictional Reading Comprehension
Corr T2 Non-Fictional Reading Comprehension
Corr T3 Non-Fictional Reading Comprehension
Corr T4 Non-Fictional Reading Comprehension

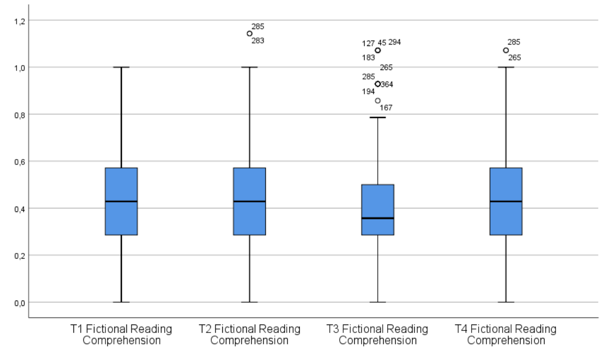

Corr T1 Non-Fictional Reading Comprehension
Corr T2 Non-Fictional Reading Comprehension
Corr T3 Non-Fictional Reading Comprehension
Corr T4 Non-Fictional Reading Comprehension

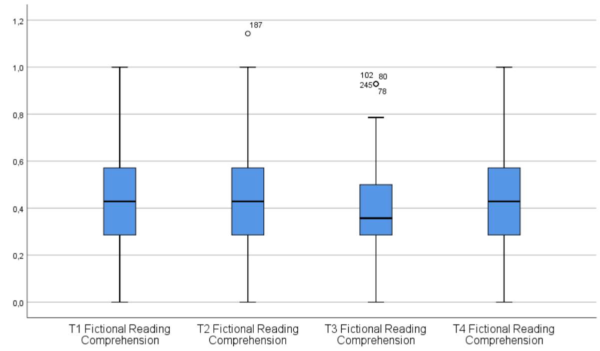

T1 Preposition Test
T2 Preposition Test
T3 Preposition Test
T4 Preposition Test

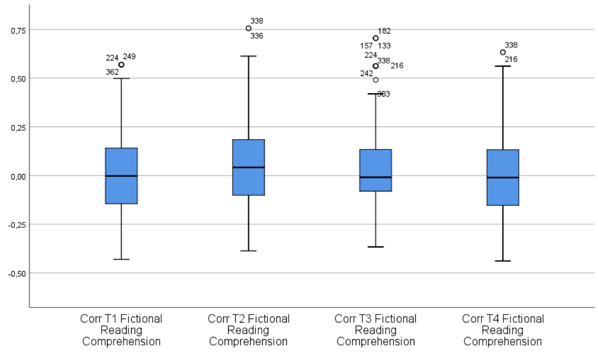

T1 Preposition Test
T2 Preposition Test
T3 Preposition Test
T4 Preposition Test

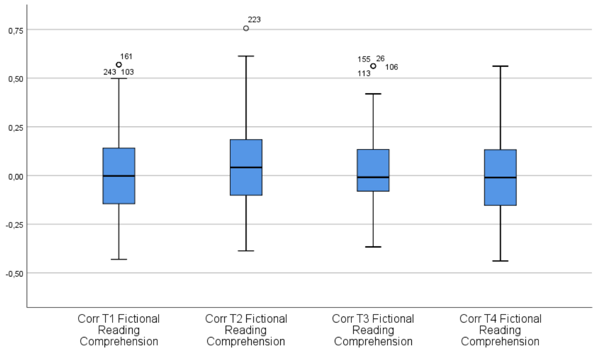

Corr T1 Preposition Test
Corr T2 Preposition Test
Corr T3 Preposition Test
Corr T4 Preposition Test

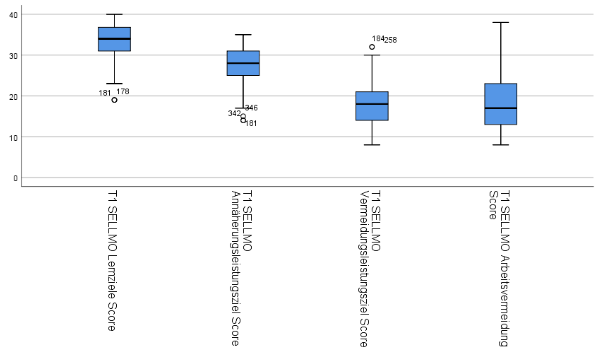

Corr T1 Preposition Test
Corr T2 Preposition Test
Corr T3 Preposition Test
Corr T4 Preposition Test

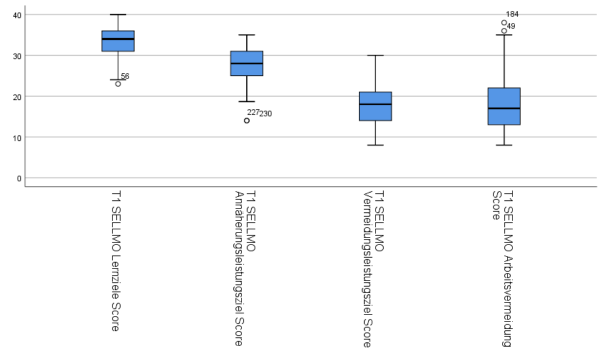

1,2
1,0
0,8
0,6
0,4
0,2
0,0
T1 Fictional Reading Comprehension
T2 Fictional Reading Comprehension
T3 Fictional Reading Comprehension
T4 Fictional Reading Comprehension

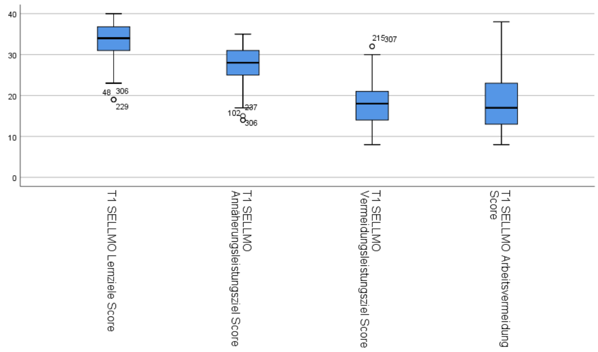

1,2
1,0
0,8
0,6
0,4
0,2
0,0
T1 Fictional Reading Comprehension
T2 Fictional Reading Comprehension
T3 Fictional Reading Comprehension
T4 Fictional Reading Comprehension

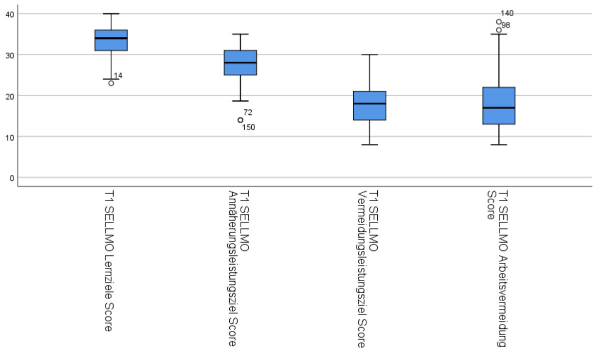

0,75
0,50
0,25
0,00
-0,25
-0,50
Corr T1 Fictional Reading Comprehension
Corr T2 Fictional Reading Comprehension
Corr T3 Fictional Reading Comprehension
Corr T4 Fictional Reading Comprehension

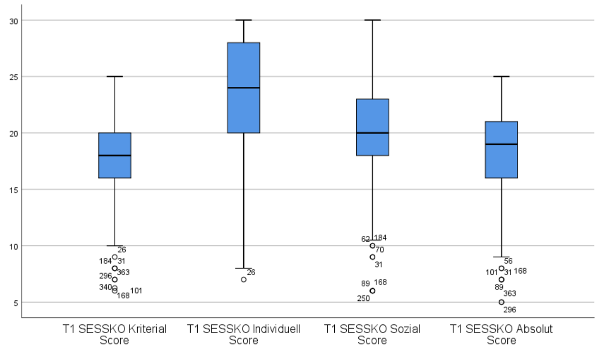

0,75
0,50
0,25
0,00
-0,25
-0,50
Corr T1 Fictional Reading Comprehension
Corr T2 Fictional Reading Comprehension
Corr T3 Fictional Reading Comprehension
Corr T4 Fictional Reading Comprehension

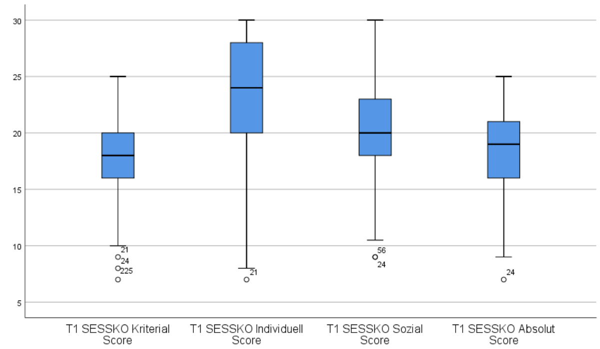

40
30
20
10
0
T1 SELLMO Lernziele Score
T1 SELLMO Annäherungsleistungsziel Score
T1 SELLMO Vermeidungsleistungsziel Score
T1 SELLMO Arbeitsvermeidung Score

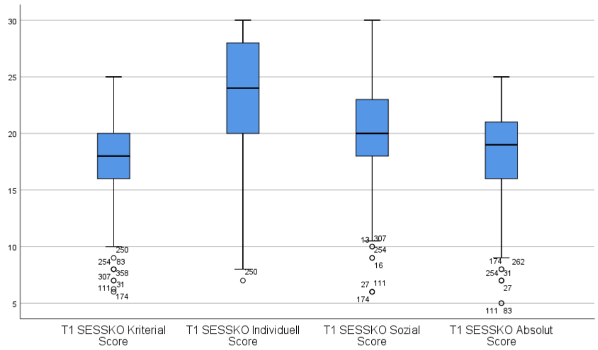

40
30
20
10
0
T1 SELLMO Lernziele Score
T1 SELLMO Annäherungsleistungsziel Score
T1 SELLMO Vermeidungsleistungsziel Score
T1 SELLMO Arbeitsvermeidung Score

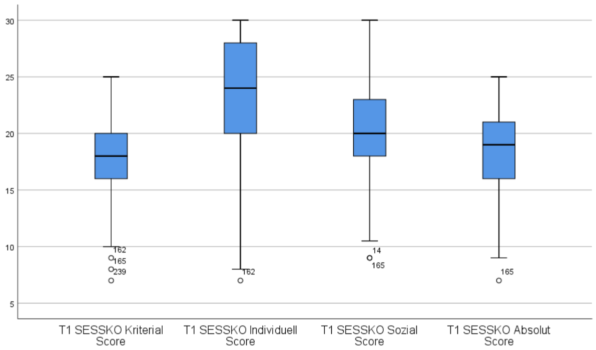

40
30
20
10
0
T1 SELLMO Lernziele Score
T1 SELLMO Annäherungsleistungsziel Score
T1 SELLMO Vermeidungsleistungsziel Score
T1 SELLMO Arbeitsvermeidung Score

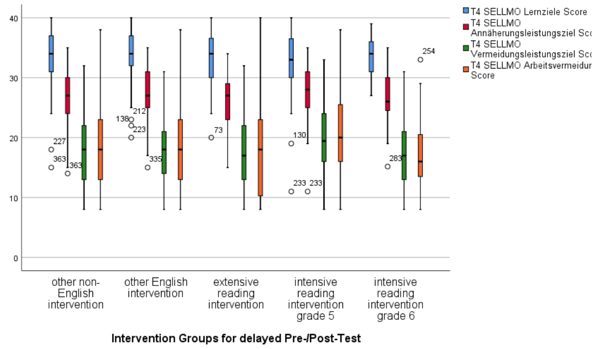

40
30
20
10
0
T1 SELLMO Lernziele Score
T1 SELLMO Annäherungsleistungsziel Score
T1 SELLMO Vermeidungsleistungsziel Score
T1 SELLMO Arbeitsvermeidung Score

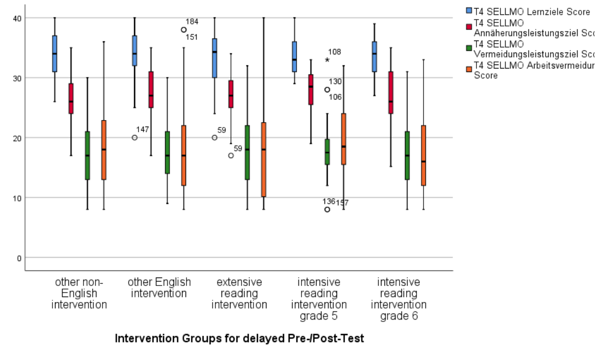

T1 SESSKO Kriterial Score
T1 SESSKO Individuell Score
T1 SESSKO Sozial Score
T1 SESSKO Absolut Score

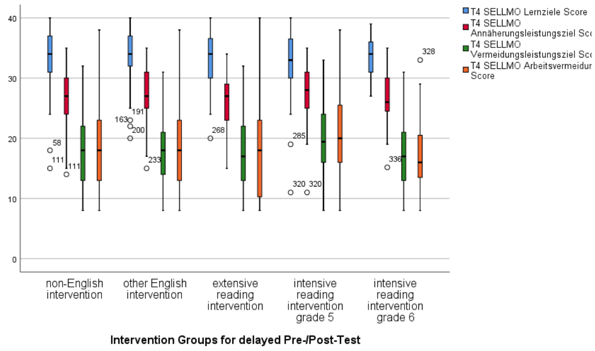

T1 SESSKO Kriterial Score
T1 SESSKO Individuell Score
T1 SESSKO Sozial Score
T1 SESSKO Absolut Score

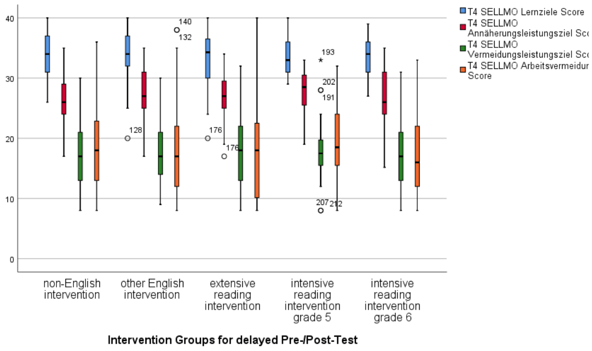

T1 SESSKO Kriterial Score
T1 SESSKO Individuell Score
T1 SESSKO Sozial Score
T1 SESSKO Absolut Score


T1 SESSKO Kriterial Score
T1 SESSKO Individuell Score
T1 SESSKO Sozial Score
T1 SESSKO Absolut Score


T4 SELLMO Lernziele Score
T4 SELLMO Annäherungsleistungsziel Sc
T4 SELLMO Vermeidungsleistungsziel Sc
T4 SELLMO Arbeitsvermeidur Score
other non-English intervention
other English intervention
extensive reading intervention
intensive reading intervention grade 5
intensive reading intervention grade 6
Intervention Groups for delayed Pre-/Post-Test

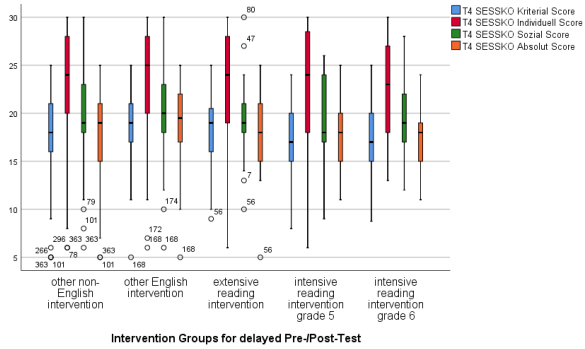

T4 SELLMO Lernziele Score
T4 SELLMO Annäherungsleistungsziel Sc
T4 SELLMO Vermeidungsleistungsziel Sc
T4 SELLMO Arbeitsvermeidur Score
other non-English intervention
other English intervention
extensive reading intervention
intensive reading intervention grade 5
intensive reading intervention grade 6
Intervention Groups for delayed Pre-/Post-Test

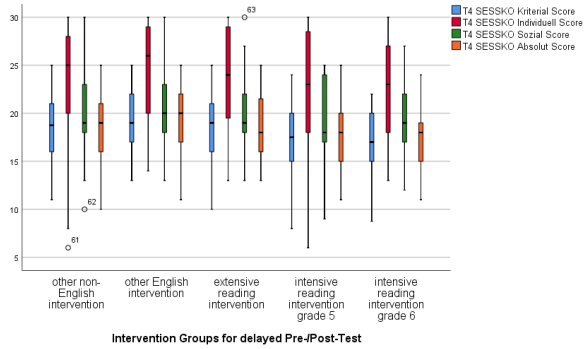

T4 SELLMO Lernziele Score
T4 SELLMO Annäherungsleistungsziel Sc
T4 SELLMO Vermeidungsleistungsziel Sc
T4 SELLMO Arbeitsvermeidur Score
non-English intervention
other English intervention
extensive reading intervention
intensive reading intervention grade 5
intensive reading intervention grade 6
Intervention Groups for delayed Pre-/Post-Test

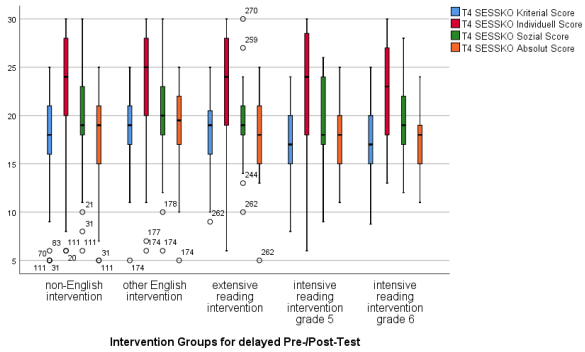

T4 SELLMO Lernziele Score
T4 SELLMO Annäherungsleistungsziel Sc
T4 SELLMO Vermeidungsleistungsziel Sc
T4 SELLMO Arbeitsvermeidur Score
non-English intervention
other English intervention
extensive reading intervention
intensive reading intervention grade 5
intensive reading intervention grade 6
Intervention Groups for delayed Pre-/Post-Test

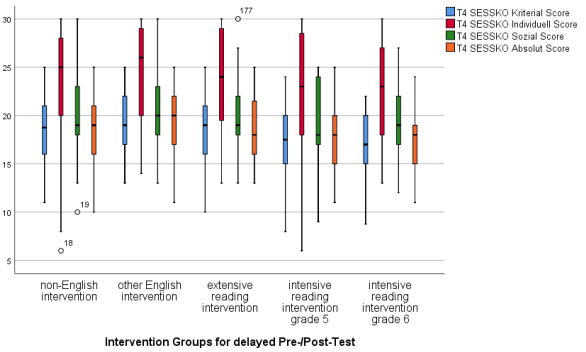
Analysis of Outliers - Z-Score SELLMO

	LZ T1	ALZ T1	VLZ T1	A T1	LZ T4	ALZ T4	VLZ T4	A T4
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
extreme |z|>3.29	4 (1.1)				3 (0.9)	1 (0.3)		
probable outlier 2.58 < |z| < 3.29	1 (0.3)	3 (0.9)	2 (0.6)	2 (0.6)	4 (1.1)	4 (1.1)	1 (0.3)	8 (2.3)
potential outliers 1.96 < |z| < 2.58	6 (1.7)	4 (1.2)	5 (1.5)	15 (4.4)	5 (1.4)	7 (2.0)	13 (3.7)	10 (2.9)
normal range |z| < 1.96	340 (96.9)	337 (97.9)	337 (98.0)	327 (95.1)	338 (96.6)	336 (96.6)	334 (96.0)	330 (94.8)







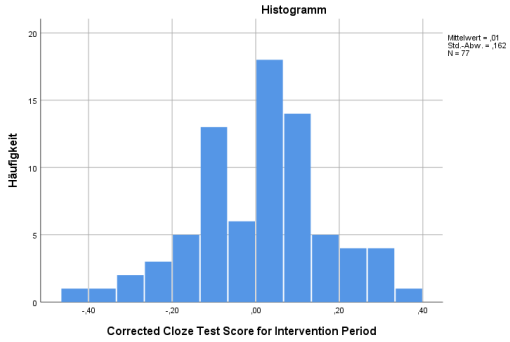


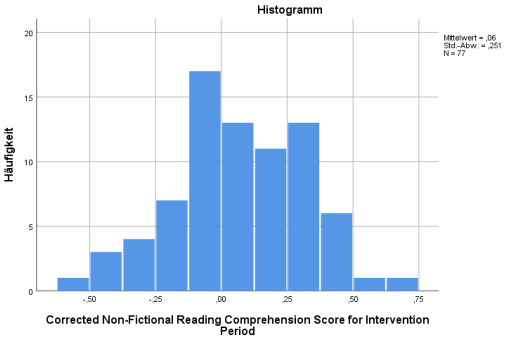
Analysis of Outliers - Z-Score SESSKO

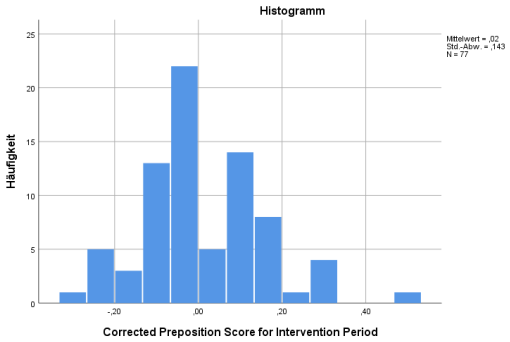
	krit. T1	indiv. T1	soz. T1	abs. T1	krit. T4	indiv. T4	soz. T4	abs. T4
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
extreme |z|>3.29			4 (1.2)	2 (0.6)	4 (1.2)		2 (0.6)	4 (1.2)
probable outlier 2.58 < |z| < 3.29	7 (2.1)	4 (1.2)	2 (0.6)	5 (1.5)	2 (0.6)	7 (2.1)	2 (0.6)	1 (0.3)
potential outliers 1.96 < |z| < 2.58	7 (2.1)	10 (2.9)	14 (4.1)	5 (1.5)	5 (1.5)	6 (1.8)	16 (4.7)	5 (1.5)
normal range |z| < 1.96	327 (95.9)	327 (95.9)	321 (94.1)	323 (96.4)	326 (96.7)	325 (96.2)	318 (94.1)	324 (97.9)

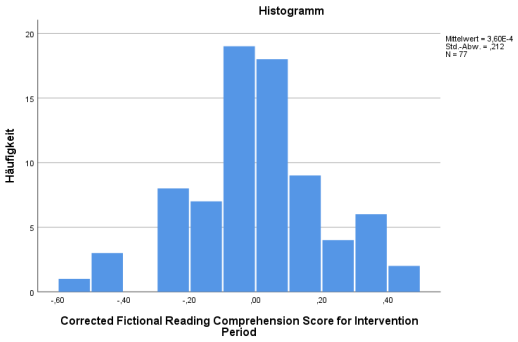
10.2.2 Normality – Histograms

Non-English Intervention

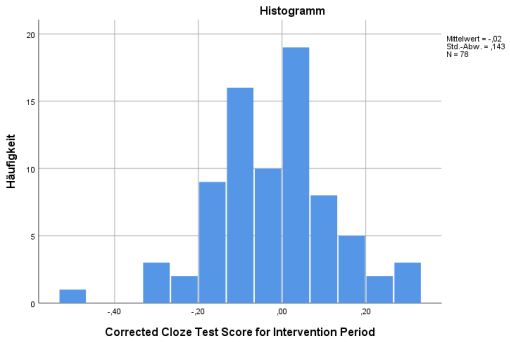


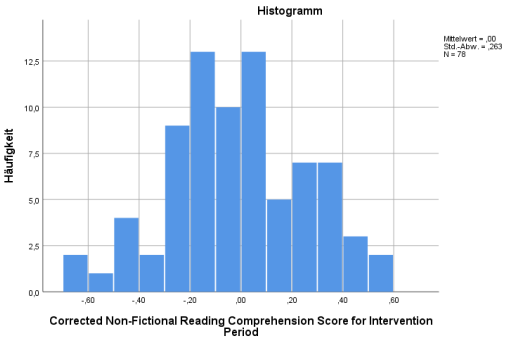


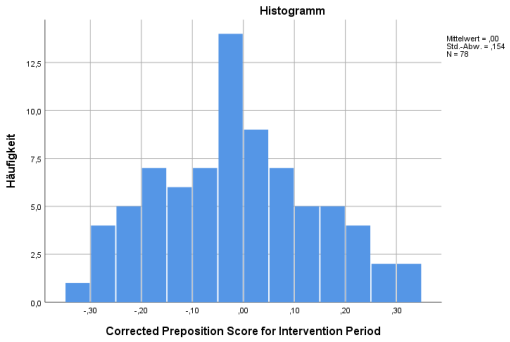


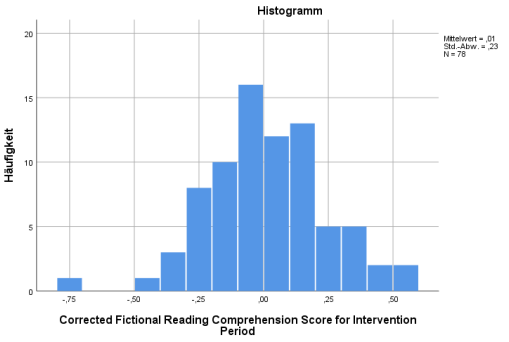


other English Intervention

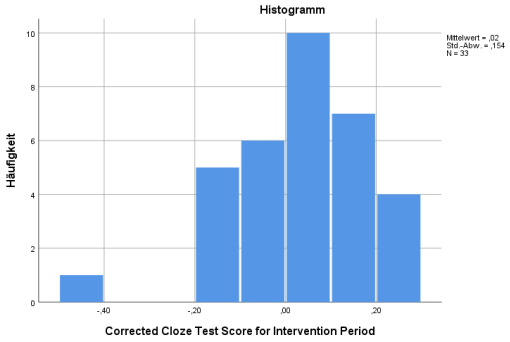


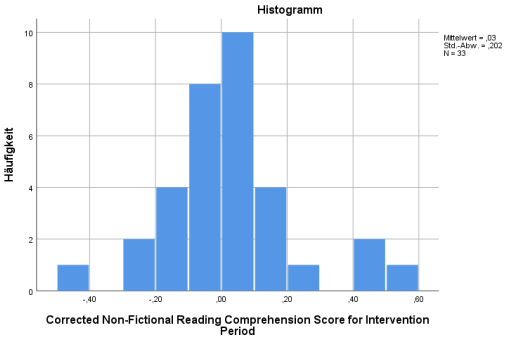


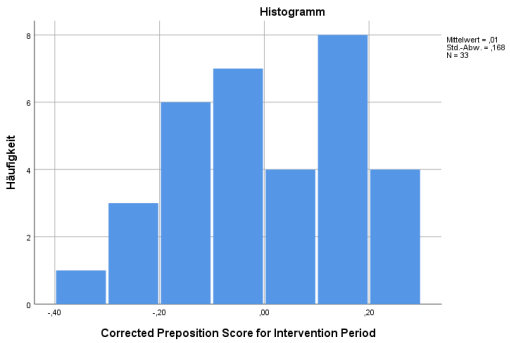


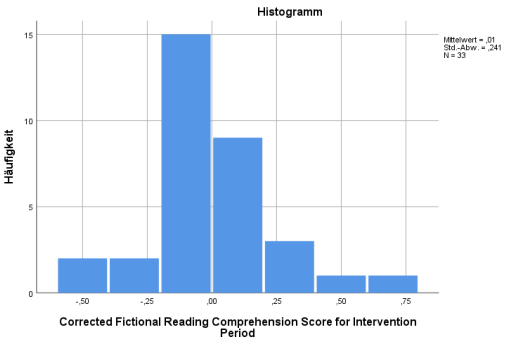


extensive reading intervention

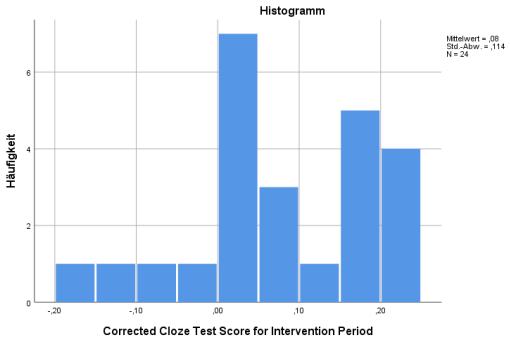


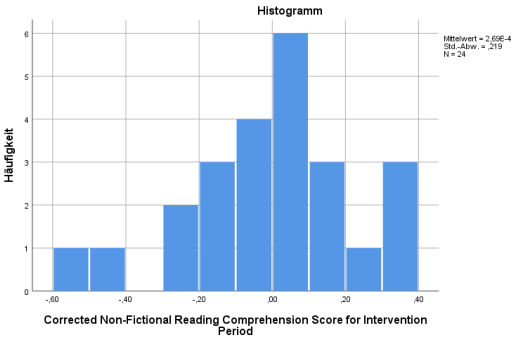


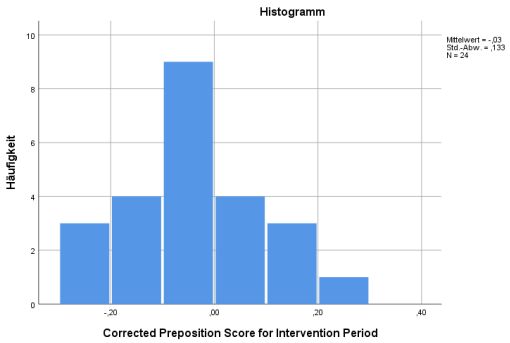


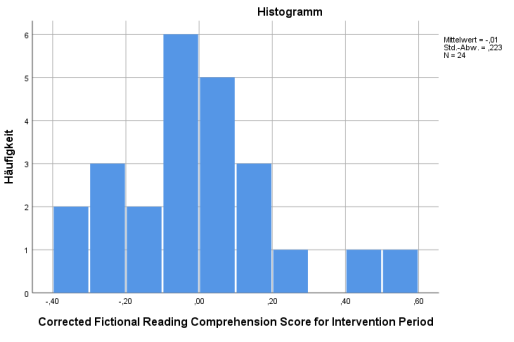


intensive reading intervention gerade 5

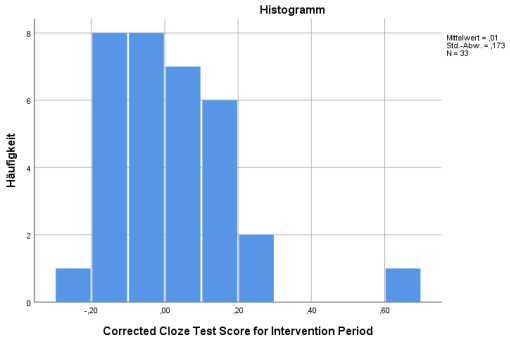


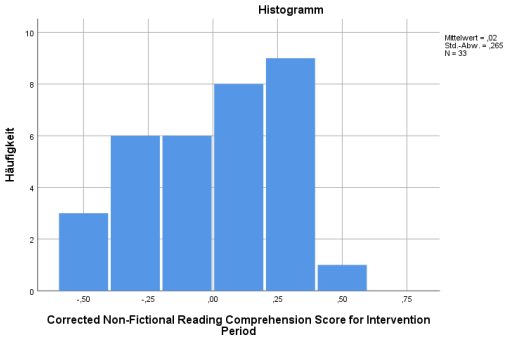


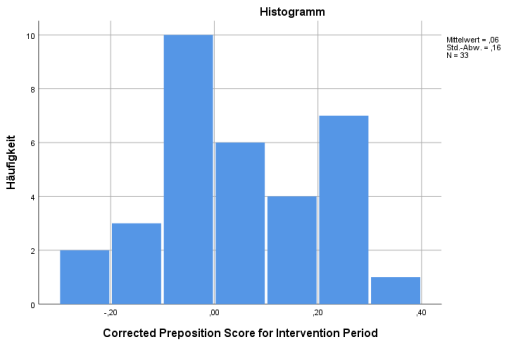


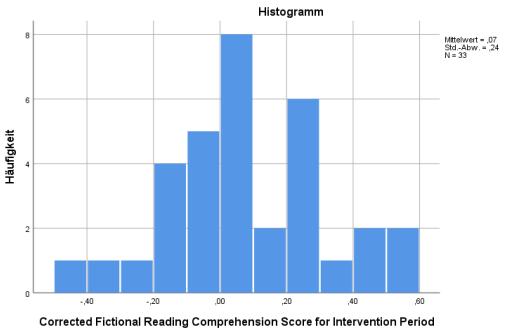


intensive reading intervention grade 6









10.3.1 Correlation Table

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
(1) T1 SELLMO Learning Goal Orientation	1																										
(2) T1 SELLMO Performance-Approach Goal Orientation	.403** .138 .431* .353**	1																									
(3) T1 SELLMO Performance-Avoid Goal Orientation	.050 -.135* -.308 .027	.422** .495** .034 .465**	1																								
(4) T1 SELLMO Work Avoidance	-.162 -.278 -.414* .009	.208 .352** -.026 .253	.507** .505** .740** .348**	1																							
(5) T1 SESSKO Criterial Self-Concept	.262* .087 -.227 .312*	.117 .369** -.069 .165	-.253* -.102 .216 -.162	-.136 -.025 .325 -.180	1																						
(6) T1 SESSKO Individual Self-Concept	.339** .172 .103 .228	.030 .334** -.014 .194	-.285* -.121 -027 -.188	-.146 -.083 .023 -.192	.624** .673** .467* .630**	1																					
(7) T1 SESSKO Social Self-Concept	.349** .122 .087 .266*	.026 .372** .105 .232	-.214 .009 .212 -.076	-.078 .039 .230 -.050	.771** .685** .760** .766**	.655** .586** .490** .672**	1																				
(8) T1 SESSKO Absolut Self-Concept	.276* .109 -.008 .251	.002 .412** -.024 .345**	-.291* -.051 .243 -.004	-.110 -.018 -198 .001	.815** .808** .837** .743**	.686** .685** .465* .637**	.779** .728** .808** .809**	1																			
(9) T4 SELLMO Learning Goal Orientation	.513** .354** .342 .345**	.288* -.058 .376 .179	-.052 -.105 -.435* -.229	-.164 -.291** -.566** -.284*	.112 .081 -.241 .247	.169 .004 -.161 .103	.142 -.101 -.206 .151	.084 .057 -.217 .165	1																		
(10) T4 SELLMO Performance-Approach Goal Orientation	.147 -.035 .304 .125	.640** .485** .426* .497**	.342** .241* -.111 .242	.068 .240* -.235 .031	.039 .211 -.058 .158	-.084 .194 -.154 .168	-.093 .176 .032 .289*	-.154 .298** .068 .348**	.343** .292** .760** .276*	1																	
(11) T4 SELLMO Performance-Avoid Goal Orientation	-.051 .009 -.276 .104	.281* .346** .284 .410**	.553** .425** .308 .622**	.274* .465** .202 .416**	-.166 -.035 .112 -.024	-.353** .009 -.187 -.017	-.137 -.016 .181 .072	-.247* .044 .230 .095	-.208 -.205 .051 -.094	.428** .427** .250 .389**	1																
(12) T4 SELLMO Work Avoidance	-.133 -.178 -.412* .008	.189 .314** .033 .093	.455** .315** .218 .351**	.529** .643** .383* .512**	-.025 .012 .203 .024	-.303** .079 -.116 .021	-.022 .150 .272 .083	-.110 .168 .134 .071	-.275* -.288** -.263 -.177	.167 .406** -.196 .256	.649** .624** .649** .500**	1														
(13) T4 SESSKO Criterial Self-Concept	.241* .045 -.043 .002	.049 .341** -.054 .103	-.282* .068 -.113 -.106	-.119 .001 -.218 -.121	.601** .713** .117 .468**	.444** .544** .156 .439**	.614** .562** -.086 .507**	.550** .662** .144 .540**	.365** .188 .218 .186	.019 .249* .281 .265*	-.308** .021 -.062 -.096	-.291* .035 -.385* .067	1													
(14) T4 SESSKO Individual Self-Concept	.217 .134 -.028 .106	.088 .094 -.010 .057	-.329** -.040 -.129 -.122	-.196 -.100 -.356 -.121	.405** .543** .145 .331*	.530** .563** .469* .538**	.424** .316** .063 .336*	.444** .501** .253 .451**	.392** .311 .262 .119	.056 .128 .359 .026	-.194 -.126 .229 -.171	-.351** -.106 -.232 -.162	.689** .621** .597** .728**	1												
(15) T4 SESSKO Social Self-Concept	.262* -.049 .180 .091	.046 .138 .295 .043	-.188 .075 -.005 -.040	-.124 .130 -.042 -.046	.525** .542** .290 .359**	.375** .397** .316 .304*	.699** .457** .391* .458**	.548** .475** .359 .473**	.203 .094 .098 .216	-.023 .265* .135 .198	-.152 .143 .230 .015	-.180 .238* .053 .004	.776** .644** .583** .801**	.544** .377** .467* .625**	1											
(16) T4 SESSKO Absolute Self-Concept	.137 .120 -.068 .004	.041 .190 .048 .092	-.280* -.022 -.023 .026	-.052 -.051 -.101 -.268*	.570** .650** .107 .424**	.430** .494** .328 .465**	.562** .475** .015 .440**	.596** .562** .134 .575**	.237* .306** -.037 .176	-.059 .250* .078 .310*	-.329** -.050 .084 -.029	-.312** -.014 -.212 .007	.832** .846** .805** .795**	.563** .657** .708** .692**	.729** .653** .687** .704**	1									
(17) Corrected Cloze Test Score for Intervention Period	.075 -.010 .231 -.368**	.159 -.066 -.147 -.164	.065 .029 -.102 -.115	.236* .072 -.326 -.279*	.091 -.157 -.217 .047	.072 -.155 -.059 .071	.080 .016 -.106 .095	.009 -.190 -.127 .073	-.125 .023 .252 -.186	.130 -.200 .372 -.008	.176 -.159 -.164 -.067	.199 -.096 -.438* -.156	-.012 -.041 .501** .054	-.024 -.061 .430* -.013	.028 .033 .318 -.091	.045 -.053 .354 .158	1									
(18) Corrected Non-Fictional Reading Comprehension Score for Intervention Period	-.110 .063 -.220 -.022	-.048 .100 -.101 .035	-.108 -.022 .132 -.045	-.072 -.012 .331 .164	-.021 .189 .117 -.110	-.042 .134 -.076 -.263*	-.062 .145 .021 -.166	-.041 .234* -.086 -.136	-.064 .005 -.081 -.026	-.089 .087 -.055 -.031	.032 .023 .016 .043	.034 .036 .034 .065	-.083 -.083 -.106 -.181	-.052 .214 .111 -.337*	-.050 .098 -.069 -.222	-.090 .052 -.104 -.346**	.120 -.027 .165 .065	1								
(19) Corrected Preposition Score for Intervention Period	.091 .002 .016 .031	-.214 -.074 -.100 -.015	-.309** -.203 -.150 -.086	-.204 -.094 -.274 .157	.011 -.144 .000 -.044	.148 .057 .341 -.044	-.005 .080 -.230 -.079	.097 -.006 -.168 -.058	.047 -.072 .135 .061	-.125 .061 -.212 -.010	-.193 .026 -.399* -.164	-.140 .060 -.421* .012	-.038 -.050 .343 -.095	-.030 -.018 .264 -.082	.009 .045 .167 -.109	-.014 -.102 .308 -.192	.154 .067 .177 -.061	.043 .217* .027 .234	1							
(20) Corrected Fictional Reading Comprehension Score for Intervention Period	-.219 -.038 -.113 -.143	-.082 -.038 -.235 -.065	.048 .012 -.056 -.037	.122 .007 .037 .098	-.111 .026 .302 -.149	-.051 -.048 .045 .135	-.077 -.028 .264 -.134	-.160 -.020 .294 -.069	-.123 -.120 .089 -.150	.021 -.019 .284 -.236	.018 -.045 .288 -.008	.014 -.052 .295 -.021	-.176 -.008 .133 -.010	-.108 -.024 .064 .123	-.077 -.038 .273 .015	-.082 .039 .119 -.038	.115 -.096 .104 .176	.037 .154 .100 .218	.079 -.011 -.247 .276*	1						
(21) Corrected Total Points Difference for Intervention Period	-.112 .015 -.090 -.185	-.078 -.009 -.001 -.070	-.116 -.070 -.060 -.103	.039 -.014 -.031 .091	-.029 .021 .161 -.125	.027 .023 .095 -.059	-.045 .104 .044 -.143	-.066 .062 .023 -.097	-.127 -.077 .154 -.124	-.034 -.003 .198 -.135	.029 -.053 -.046 -.055	.054 -.016 -.124 -.122	-.150 .013 .336 -.108	-.101 .086 .237 -.139	-.053 .065 .306 -.167	-.080 -.004 .262 -.202	.545** .287 .575** .410**	.650** .734 .637** .699**	.446** .505 .306 .544**	.575** .570 .593** .731**	1						
(22) Corrected Cloze Test Score for Semester Period	.005 .161 .339 -.060	.152 .121 .175 .042	.095 .086 -.168 -.049	.079 .025 -.388* -.121	-.364** .011 -.157 -.219	-.242* .043 -.048 -.085	-.082 .011 -.143 -.076	-.296** -.038 -.116 .009	-.054 .006 .446* .158	.063 -.036 .225 .112	.177 .035 -.233 .200	.127 .024 -.295 -.264*	-.184 -.013 .323 .050	-.023 .022 .030 .044	-.082 .057 .183 .023	-.191 .020 .091 .093	.290* .314** .325 .188	.038 .220* -.236 -.006	.009 .185 .100 .027	-.021 .089 .139 -.114	.122 .344** .119 .009	1					
(23) Corrected Non-Fictional Reading Comprehension Score for Semester Period	-.222 .033 -.163 .052	-.005 .005 -.204 .144	.002 .155 -.412* -.036	.066 -.100 -.281 -.041	-.165 -.169 -.195 -.111	-.076 -.172 -.361 -.154	-.189 -.169 -.299 -.107	-.002 -.191 -.391* .010	-.124 -.033 .109 .077	.037 -.004 -.102 -.005	.083 .009 -.268 .027	.062 -.166 -.139 -.188	-.194 -.147 -.215 -.297*	-.163 -.094 -.244 -.216	-.132 -.099 -.449* -.272*	-.181 -.136 -.309 -.285*	-.015 .104 .157 .072	.506** .339** .563** .378**	.014 .190 .103 .095	-.016 .013 .054 .217	.282* .303** .414* .346**	.100 .312** -.122 .031	1					
(24) Corrected Preposition Score for Semester Period	.203 -.057 .044 -.010	-.044 .028 .159 -.172	-.164 .073 .239 -.143	-.082 .087 .029 -.082	.055 .083 .140 -.029	.169 .087 .044 .087	.156 .007 -.016 -.040	.178 -.012 .185 -.068	.296** .060 .303 -.061	-.030 .081 .346 -.039	-.108 -.060 .078 -.042	-.095 -.139 -.183 -.014	.097 .104 .382* .046	.034 .096 .306 -.038	.125 .061 .133 -.102	.118 .095 .125 -.043	.047 .066 .218 .151	-.023 .097 .025 .236	.357** .265* .351 .384**	.006 .304** .009 .389**	.124 .331** .229 .473**	-.008 .311** .236 .066	-.092 .193 -.249 .134	1				
(25) Corrected Fictional Reading Comprehension Score for Semester Period	-.138 .077 -.132 -.056	-.142 -.125 .018 .028	-.050 -.013 -.218 -.111	.014 -.140 -.215 -.053	-.247* -.086 .075 -.145	-.022 -.115 .024 -.055	-.146 -.033 .059 -.109	-.122 -.103 .064 -.165	-.024 -.001 .153 -.052	-.028 -.155 .072 -.097	-.034 -.087 -.137 -.116	-.132 -.156 -.067 -.225	-.140 -.095 .040 -.110	.053 .023 .005 .060	-.160 .014 -.067 -.069	-.123 .037 -.093 -.139	-.003 -.030 -.035 .114	.181 .071 .033 .084	.086 .152 .183 .262*	.378** .265* .215 .470**	.315** .223* .201 .385**	.045 .224* .099 .088	.071 .131 .161 .239	-.014 .228* .233 .373**	1			
(26) Corrected Total Points Difference for Semester Period	-.133 .081 -.007 -.013	-.050 .002 .013 .057	-.058 .069 -.348 -.125	.043 -.070 -.428* -.104	-.341** -.090 -.097 -.240	-.075 -.091 -.204 -.112	-.165 -.093 -.228 -.143	-.101 -.153 -.186 -.083	.006 .003 .424* .047	.016 -.048 .179 -.026	.048 -.036 -.300 -.102	-.031 -.181 -.296 -.296*	-.222 -.083 .150 -.194	-.063 -.002 -.301 -.100	-.151 -.009 -.178 -.213	-.198 -.019 -.151 -.214	.104 .164 .278 .187	.413** .302** .253 .327*	.197 .297** .305 .291*	.200 .232* .203 .415**	.437** .458** .465* .513**	.416** .666** .472* .349**	.644** .724** .536** .736**	.314** .595** .385* .556**	.620** .593** .721** .706**	1		
(27) Difference in grades before and after the intervention	-.093 -.055 .030 .005	.025 .105 -.089 .024	.017 .159 .416* -.002	.119 .054 .201 .128	-.004 -.079 -335 .303*	.006 -.051 -.004 .238	-.050 .089 .222 .222	.017 .012 .520** .374**	-.063 -.115 -.234 -.055	.019 .138 -.010 .049	.091 .042 .006 -.055	.080 .230* -.110 .079	-.115 .031 -.073 .192	-.017 -.094 -.102 .219	-.007 .016 -.032 .131	-.058 .092 -.234 .282*	.111 -.114 -.159 .225	-.199 -.202 -.016 -.047	.007 -.150 -.034 -.023	.148 -.144 -.053 .067	.001 -.281* -.110 .071	.018 -.222* -.073 .239	-.047 -.143 -.201 .106	.044 -.199 .378 -.202	-.047 -.109 .046 -.085	-.032 -.251* -.006 .030	1	

**. Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.
*. Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.
first row: non-English intervention group; second row: English intervention; third row: extensive reading intervention; fourth row: intensive reading intervention
color coding: blue – significant positive correlations; yellow – significant negative correlations; green – significant positive and negative correlations

10.3.2 Moderation Analyses

10.3.2.1 Moderation Analyses for Total Points

	direct pre-/posttest				delayed pre-posttest			
Variables	***b***	***SE B***	***t***	***p***	***b***	***SE B***	***t***	***p***
Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups (Step 2)								
Constant	0.02 [0.00, 0.04]	0.01	2.50	.01	0.00 [-0.02, 0.02]	0.01	0.16	.88
Group (non-English vs. reading English)	0.00 [-0.03, 0.04]	0.02	0.07	.94	0.00 [-0.04, 0.04]	0.02	0.13	.90
Self-Concept (criterial)	0.00 [-0.01, 0.00]	0.00	-0.33	.74	-0.01 [-0.02, 0.00]	0.00	-3.32	<.01
Group x Self-Concept (criterial	0.00 [-0.01, 0.01]	0.01	-0.04	.97	-0.01 [-0.02, 0.01]	0.01	-0.94	.35
	$F(3; 157) = 0.04$; $R^2 = .00$; $p = .99$				$F(3; 157) = 3.80$; $R^2 = .07$; $p = .01$			
Constant	0.02 [0.00, 0.04]	0.01	2.50	.02	0.00 [-0.02, 0.02]	0.01	0.08	.94
Group (non-English vs. reading English)	0.00 [-0.03, 0.04]	0.02	0.14	.89	0.00 [-0.04, 0.04]	0.02	-0.10	.93
Self-Concept (social)	0.00 [0.00, 0.00]	0.00	-0.80	.43	-0.01 [-0.01, 0.00]	0.00	-2.05	.04
Group x Self-Concept (social)	0.00 [-0.01, 0.01]	0.00	0.24	.81	0.00 [-0.01. 0.01]	0.01	0.13	.90
	$F(3; 157) = 0.23$; $R^2 = .00$; $p = .87$				$F(3; 157) = 1.45$; $R^2 = .03$; $p = .23$			
Constant	0.09 [0.00, 0.18]	0.04	2.08	.04	0.12 [0.03, 0.21]	0.05	2.51	.01
Group (non-English vs. reading English)	-0.02 [-0.08, 0.03]	0.03	-0.83	.41	-0.03 [-0.09, 0.03]	0.03	-0.93	.36
Non-fictional reading comprehension points T1	-0.26 [-0.49, -0.03]	0.12	-2.20	.03	-0.41 [-0.66, -0.16]	0.13	-3.23	<.001
Group x non-fictional RC points T1	0.09 [-0.05, 0.24]	0.07	1.33	.18	-0.11 [-0.05, 0.26]	0.08	1.38	.17
	$F(3; 157) = 3.63$; $R^2 = .06$; $p = .01$				$F(3; 157) = 13.70$; $R^2 = .21$; $p = .00$			
Constant	0.10 [0.00, 0.04]	0.05	2.11	.04	0.06 [-0.02, 0.02]	0.06	1.15	.25
Group (non-English vs. reading English)	-0.04 [-0.03, 0.04]	0.03	-1.31	.19	-0.02 [-0.04, 0.04]	0.04	-0.53	.60
Preposition test points T1	-0.36 [0.00, 0.00]	0.17	-2.07	.04	-0.24 [-0.01, 0.00]	0.20	-1.21	.23
Group x Preposition test points T1	0.19 [-0.01, 0.01]	0.12	1.66	.10	0.06 [-0.01. 0.01]	0.13	0.48	.64
	$F(3; 157) = 1.70$; $R^2 = .03$; $p = .17$				$F(3; 157) = 1.89$; $R^2 = .03$; $p = .13$			
Reading Intervention Group vs. Other English Intervention Groups (Step 3)								
Constant	0.01 [-0.01, 0.03]	0.01	0.74	.46	0.02 [-0.01, 0.04]	0.01	1.32	.19
Group (reading English vs. other English)	-0.02 [-0.07, 0.00]	0.02	-1.74	.08	0.02 [-0.03, 0.07]	0.02	0.87	.38
Work Avoidance	0.00 [0.00, 0.00]	0.00	0.24	.81	0.00 [-0.01, 0.00]	0.00	-2.15	.03
Group x Work Avoidance	0.00 [-0.01, 0.00]	0.00	-0.58	.58	0.00 [-0.01, 0.01]	0.00	0.29	.77
	$F(3; 153) = 1.14$; $R^2 = .02$; $p = .33$				$F(3; 153) = 1.88$; $R^2 = .04$; $p = .14$			
Constant	-0.01 [-0.12, 0.09]	0.05	-0.27	.78	-0.10 [-0.23, 0.03]	0.07	-1.53	.13
Group (reading English vs. other English)	0.01 [-0.06, 0.08]	0.04	0.38	.71	0.10 [0.01, 0.19]	0.04	2.30	.02
Cloze Test Points T1	0.01 [0.00, 0.02]	0.01	1.55	.12	0.01 [-0.01, 0.03]	0.01	1.37	.17
Group x Cloze Test Points T1	-0.01 [-0.01, 0.00]	0.00	-1.53	.13	-0.01 [-0.02, 0.00]	0.01	-1.99	.05
	$F(3; 153) = 1.85$; $R^2 = .04$; $p = .14$				$F(3; 153) = 2.48$; $R^2 = .05$; $p = .06$			
Constant	0.12 [0.03, 0.20]	0.04	2.73	.01	-0.10 [0.00, 0.21]	0.05	2.01	.05
Group (reading English vs. other English)	-0.05	0.03	-1.83	.06	-0.01	0.03	-0.43	.67

	[-0.10, 0.00]				[-0.08, 0.05]			
Non-fictional reading comprehension points T1	-0.21 [0.45, 0.03]	0.12	-1.71	.09	-0.40 [-0.68, -0.11]	0.14	-2.75	.01
Group x non-fictional RC points T1	0.05 [-0.11, 0.20]	0.08	0.58	.57	0.10 [-0.09, 0.028]	0.09	1.05	.29
	$F(3; 153) = 5.55$; $R^2 = .10$; $p = .001$				$F(3; 153) = 10.62$; $R^2 = .17$; $p = .00$			

10.3.2.2 Moderation Analyses for Cloze Test Result

	direct pre-/posttest				delayed pre-posttest			
Variables	***b***	***SE B***	***t***	***p***	***b***	***SE B***	***t***	***p***
Step 2: Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups								
Constant	0.03 [0.00; 0.05]	0.01	2.10	.04	0.00 [-0.02; 0.03]	0.02	0.24	.81
Group (non-English vs. reading English)	-0.02 [-0.07; 0.03]	0.02	-0.94	.35	0.00 [-0.05; 0.05]	0.03	0.00	.99
Performance Approach Orientation	0.00 [-0.01; 0.01]	0.00	-0.08	.94	0.00 [0.00; 0.01]	0.00	1.47	.14
Group x Performance Approach Orientation	0.01 [0.00; 0.02]	0.01	1.98	.05	0.00 [-0.01; 0.01]	0.01	0.25	.80
	$F(3; 157) = 1.61$; $R^2 = .03$; $p = .19$				$F(3; 157) = 0.77$; $R^2 = .01$; $p = .51$			
Constant	0.03 [0.00; 0.05]	0.01	2.07	.04	0.00 [-0.02; 0.03]	0.01	0.27	.79
Group (non-English vs. reading English)	-0.02 [-0.07; 0.02]	0.02	-0.97	.33	0.00 [-0.06; 0.05]	0.03	-0.14	.89
Work Avoidance	0.00 [0.00; 0.00]	0.00	-0.20	.84	0.00 [-0.01; 0.00]	0.00	-0.98	.33
Group x Work Avoidance	0.01 [0.00; 0.02]	0.00	3.37	.00	0.01 [0.00; 0.02]	0.00	1.89	.06
	$F(3; 157) = 4.22$; $R^2 = .07$; $p = .01$				$F(3; 157) = 1.68$; $R^2 = .03$; $p = .17$			
Constant	0.02 [0.00; 0.05]	0.01	1.86	.06	0.00 [-0.02; 0.03]	0.01	0.30	.77
Group (non-English vs. reading English)	-0.02 [-0.07; 0.03]	0.03	-0.96	.34	0.00 [-0.05; 0.05]	0.03	0.11	.91
Self-Concept (criterial)	0.00 [-0.01; 0.01]	0.00	0.31	.76	-0.01 [0.02; -0.01]	0.00	-3.55	.00
Group x Self-Concept (criterial)	0.01 [0.01; 0.02]	0.01	0.90	.37	-0.01 [-0.02; 0.01]	0.01	-1.17	.25
	$F(3; 157) = 0.58$; $R^2 = .01$; $p = .63$				$F(3; 157) = 4.38$; $R^2 = .07$; $p = .01$			
Constant	0.02 [0.00; 0.05]	0.01	1.89	.06	0.00 [-0.02; 0.03]	0.01	0.29	.77
Group (non-English vs. reading English)	-0.02 [-0.07; 0.03]	0.03	-0.98	.33	0.00 [-0.05; 0.05]	0.03	0.01	.99
Self-Concept (individual)	0.00 [0.00; 0.01]	0.00	0.61	.54	-0.01 [-0.01; 0.00]	0.00	-1.93	.06
Group x Self-Concept (individual)	0.00 [-0.01; 0.01]	0.01	0.36	.72	-0.01 [0.02; 0.00]	0.01	-1.08	.28
	$F(3; 157) = 0.45$; $R^2 = .01$; $p = .72$				$F(3; 157) = 1.50$; $R^2 = .03$; $p = .21$			
Constant	0.02 [0.00; 0.05]	0.01	1.90	.06	0.01 [-0.02; 0.04]	0.01	0.41	.68
Group (non-English vs. reading English)	-0.02 [-0.07; 0.03]	0.03	-0.94	.35	0.00 [-0.06; 0.05]	0.03	0.14	.89
Self-Concept (absolute)	0.00 [0.01; 0.01]	0.00	0.76	.94	-0.01 [0.00; 0.00]	0.00	-2.14	.03
Group x Self-Concept (absolute)	0.00 [-0.01; 0.01]	0.01	0.05	.96	-0.01 [-0.01; 0.01]	0.01	-1.70	.09
	$F(3; 157) = 0.29$; $R^2 = .01$; $p = .83$				$F(3; 139) = 01.09$; $R^2 = .04$; $p = .10$			
Constant	-0.04 [-0.17; 0.10]	0.07	-0.51	.61	0.11 [-0.03; 0.24]	0.07	1.57	.12
Group (non-English vs. reading English)	0.06 [-0.03; 0.15]	0.05	1.27	.21	0.02 [-0.07; 0.10]	0.04	0.34	.73
Cloze Test Points T1	0.01 [-0.00; 0.03]	0.01	1.55	.12	-0.02 [-0.03; 0.00]	0.01	-1.98	.05
Group x Cloze Test Points T1	-0.01 [-0.01; 0.01]	0.01	-2.08	.04	-0.001 [-0.01; 0.01]	0.01	-0.20	.84
	$F(3; 157) = 2.33$; $R^2 = .04$; $p = .08$				$F(3; 157) = 15.72$; $R^2 = .23$; $p = .00$			

Reading Intervention Group vs. Other English Intervention Groups (Step 3)								
Constant	0.01 [-0.01; 0.03]	0.01	0.74	.46	0.00 [-0.03; 0.03]	0.01	0.07	.95
Group (reading English vs. other English)	-0.06 [-0.11; -0.01]	0.02	-2.44	.02	-0.01 [-0.06; 0.05]	0.03	-0.20	.84
Learning Goal Orientation	0.00 [-0.01; 0.00]	0.00	-1.26	.21	0.01 [0.00; 0.02]	0.00	1.81	.07
Group x Learning Goal Orientation	0.01 [-0.01; 0.02]	0.01	0.88	.38	0.00 [0.01; 0.02]	0.01	0.48	.63
	$F(3; 153) = 1.60$; $R^2 = .05$; $p = .05$				$F(3; 153) = 1.32$; $R^2 = .03$; $p = .27$			
Constant	0.01 [-0.02; 0.03]	0.01	0.67	.51	0.00 [-0.03; 0.03]	0.01	0.02	.98
Group (reading English vs. other English)	-0.06 [-0.10; -0.01]	0.02	-2.33	.02	-0.01 [-0.07; 0.05]	0.03	-0.36	.72
Performance Approach Orientation	0.00 [-0.01; 0.00]	0.00	-1.35	.18	0.01 [0.00; 0.01]	0.00	1.46	.15
Group x Performance Approach Orientation	0.00 [-0.01; 0.02]	0.01	0.71	.48	0.00 [-0.01; 0.02]	0.01	0.36	.72
	$F(3; 153) = 2.66$; $R^2 = .05$; $p = .05$				$F(3; 153) = 0.80$; $R^2 = .02$; $p = .49$			
Constant	0.01 [-0.01; 0.03]	0.01	0.74	.46	0.00 [-0.03; 0.03]	0.01	0.05	.96
Group (reading English vs. other English)	-0.06 [-0.11; -0.01]	0.02	-2.39	.02	-0.01 [-0.07; 0.05]	0.03	-0.31	.75
Performance Avoid Orientation	0.00 [-0.01; 0.00]	0.00	-0.31	.76	0.00 [-0.01; 0.00]	0.00	-1.09	.28
Group x Performance Avoid Orientation	0.01 [0.00, 0.02]	0.01	1.16	.25	0.00 [-0.01; 0.21	0.01	0.12	.90
	$F(3; 153) = 2.39$; $R^2 = .04$; $p = .07$				$F(3; 153) = 0.42$; $R^2 = .01$ $p = .74$			
Constant	0.01 [-0.01; 0.03]	0.01	0.84	.40	0.00 [-0.03; 0.03]	0.01	0.12	.90
Group (reading English vs. other English)	-0.06 [-0.11; -0.01]	0.02	-2.50	.01	-0.01 [-0.07; 0.05]	0.03	-0.35	.72
Work Avoidance	0.00 [-0.01; 0.00]	0.00	-1.63	.10	0.00 [-0.01; 0.00]	0.00	-1.41	.16
Group x Work Avoidance	0.01 [0.00; 0.02]	0.00	2.35	.02	0.01 [0.00; 0.01]	0.00	1.48	.14
	$F(3; 153) = 4.61$; $R^2 = .08$; $p = .004$				$F(3; 153) = 1.37$; $R^2 = .03$; $p = .26$			
Constant	0.01 [-0.02; 0.03]	0.01	0.74	.46	0.01 [-0.02; 0.04]	0.01	0.85	.40
Group (reading English vs. other English)	-0.07 [-0.11; -0.02]	0.02	-2.63	.01	0.01 [-0.05; 0.07]	0.03	0.23	.82
leisure reading behavior	0.00 [-0.01; 0.00]	0.00	-0.70	.49	0.00 [0.00; 0.01]	0.00	0.57	.57
Group x reading behavior	0.00 [-0.01; 0.01]	0.00	-0.04	.97	0.00 [-0.01; 0.01]	0.00	0.35	.73
	$F(3; 133) = 2.45$; $R^2 = .05$; $p = .07$				$F(3; 133) = 0.19$; $R^2 = .004$; $p = .90$			
Constant	0.01 [-0.01; 0.03]	0.01	0.83	.41	0.00 [-0.03; 0.03]	0.01	-0.16	.88
Group (reading English vs. other English)	-0.05 [-0.10; -0.00]	0.02	-2.14	.03	0.00 [-0.06; 0.05]	0.03	-0.16	.87
Self-Concept (criterial)	0.00 [-0.01; 0.00]	0.00	-1.36	.18	0.00 [0.01; 0.00]	0.00	-0.94	.35
Group x Self-Concept (criterial)	-0.01 [-0.02; 0.01]	0.01	-0.83	.41	0.01 [-0.01; 0.03]	0.01	1.27	.21
	$F(3; 153) = 2.66$; $R^2 = .05$; $p = .05$				$F(3; 153) = 1.02$; $R^2 = .02$; $p = .38$			
Constant	0.01 [-0.01; 0.03]	0.01	0.80	.43	0.00 [-0.03; 0.03]	0.01	0.01	.99
Group (reading English vs. other English)	-0.06 [-0.10; -0.01]	0.02	-2.31	.02	-0.01 [-0.07; 0.05]	0.03	-0.26	.80
Self-Concept (individual)	0.00 [-0.01; 0.00]	0.00	-0.90	.37	0.00 [-0.01; 0.00]	0.00	-0.36	.72
Group x Self-Concept (individual)	0.01 [-0.01; 0.00]	0.00	-1.28	.20	0.00 [-0.01; 0.01]	0.01	0.50	.62
	$F(3; 153) = 2.69$; $R^2 = .05$; $p = .05$				$F(3; 153) = 0.16$; $R^2 = .00$; $p = .92$			

Constant	0.01 [-0.02; 0.03]	0.01	0.75	.45	0.00 [-0.03; 0.03]	0.01	-0.08	.94
Group (reading English vs. other English)	-0.06 [-0.11; -0.01]	0.02	-2.34	.02	-0.01 [-0.07; 0.05]	0.03	-0.23	.81
Self-Concept (social)	0.00 [0.00; 0.01]	0.00	-0.06	.95	0.00 [-0.01; 0.01]	0.00	-0.44	.66
Group x Self-Concept (social)	0.00 [-0.02; 0.01]	0.01	-0.41	.68	-0.01 [-0.01; 0.02]	0.01	0.85	.40
	$F(3; 153) = 1.95$; $R^2 = .04$; $p = .12$				$F(3; 153) = 0.41$; $R^2 = .01$; $p = .74$			
Constant	0.01 [-0.02; 0.03]	0.01	0.94	.35	0.00 [-0.02; 0.04]	0.01	0.03	.98
Group (reading English vs. other English)	-0.05 [-0.11; -0.02]	0.02	-2.11	.04	-0.01 [-0.05; 0.07]	0.03	-0.21	.83
Self-Concept (absolute)	0.00 [-0.01; 0.00]	0.00	-1.34	.18	0.00 [0.00; 0.01]	0.00	-0.39	.70
Group x Self-Concept (absolute)	-0.01 [-0.01; 0.01]	0.01	-1.45	.15	0.00 [-0.01; 0.01]	0.01	0.07	.94
	$F(3; 153) = 3.01$; $R^2 = .06$; $p = .03$				$F(3; 133) = 0.19$; $R^2 = .004$; $p = .97$			
Step 4: Intensive Reading Intervention Groups vs. Extensive Intervention Groups (Step 4)								
Constant	0.03 [0.00; 0.07]	0.02	2.17	.03	0.00 [-0.03; 0.04]	0.02	0.22	.82
Group (intensive vs. extensive reading)	0.01 [-0.06; 0.08]	0.03	0.38	.71	0.01 [-0.07; 0.09]	0.04	0.32	.75
Learning Goal Orientation	-0.01 [-0.02; 0.00]	0.01	-1.80	.08	0.00 [-0.01; 0.02]	0.01	0.70	.49
Group x Learning Goal Orientation	-0.03 [-0.05; -0.01]	0.01	-2.76	.01	-0.02 [-0.05; 0.00]	0.01	-1.87	.07
	$F(3; 80) = 3.27$; $R^2 = .11$; $p = .03$				$F(3; 80) = 1.51$; $R^2 = .05$; $p = .22$			
Constant	0.03 [0.00; 0.07]	0.02	2.13	.04	0.00 [-0.04; 0.04]	0.02	0.07	.94
Group (intensive vs. extensive reading)	0.02 [-0.05; 0.09]	0.04	0.70	.49	0.03 [-0.05; 0.11]	0.04	0.66	.51
Work Avoidance	0.01 [-0.01; 0.00]	0.00	-2.74	.01	-0.01 [-0.01; 0.00]	0.00	-1.98	.05
Group x Work Avoidance	0.00 [-0.01; 0.01]	0.01	0.10	.92	0.01 [0.00; 0.02]	0.01	1.39	.17
	$F(3; 80) = 2.60$; $R^2 = .09$; $p = .06$				$F(3; 80) = 1.63$; $R^2 = .06$; $p = .19$			
Constant	-0.01 [-0.23; 0.21]	0.11	-0.09	.93	0.17 [-0.07; 0.40]	0.12	1.43	.16
Group (intensive vs. extensive reading)	0.02 [-0.11; 0.15]	0.07	0.30	.77	-0.02 [-0.16; 0.11]	0.07	-0.36	.72
Cloze Test Points T1	0.00 [-0.03; 0.04]	0.02	0.28	.78	-0.04 [-0.07; 0.00]	0.12	-2.20	.03
Group x Cloze Test Points T1	0.00 [-0.02; 0.02]	0.01	-0.17	.86	0.01 [-0.01; 0.03]	0.01	1.17	.25
	$F(3; 80) = 2.60$; $R^2 = .09$; $p = .06$				$F(3; 80) = 7.29$; $R^2 = .21$; $p < .001$			

10.3.2.3 Moderation Analyses for Non-Fictional Reading Comprehension Test Results

	direct pre-/posttest				delayed pre-posttest			
Variables	***b***	***SE B***	***t***	***p***	***b***	***SE B***	***t***	***p***
Step 2: Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups (Step 2)								
Constant	0.04 [0.00; 0.08]	0.02	1.96	.05	0.01 [-0.04; 0.06]	0.02	0.32	.75
Group (non-English vs. reading English)	0.05 [-0.02; 0.13]	0.04	1.40	.16	-0.03 [-0.13; 0.07]	0.05	-0.61	.54
Self-Concept (social)	-0.01 [-0.02; 0.00]	0.00	-1.08	.28	-0.01 [-0.03; 0.00]	0.01	-2.22	.03
Group x Self-Concept (social)	0.00 [-0.02; 0.02]	0.01	0.27	.79	0.00 [-0.03; 0.02]	0.01	-0.02	.98
	$F(3; 157) = 0.96$; $R^2 = .02$; $p = .41$				$F(3; 157) = 1.91$; $R^2 = .04$; $p = .13$			
Constant	-0.13 [-0.35; 0.08]	0.11	-1.23	.22	-0.22 [-0.49; 0.05]	0.14	-1.60	.11
Group (non-English vs. reading English)	0.11 [-0.03; 0.25]	0.07	1.59	.11	0.15 [-0.03; 0.33]	0.09	1.67	.10

Cloze Test Points T1	0.01 [-0.01; 0.04]	0.01	1.10	.27	0.04 [0.01; 0.07]	0.02	2.52	.01
Group x Cloze Test Points T1	-0.01 [-0.03; 0.01]	0.01	-1.08	.28	-0.03 [-0.05; 0.01]	0.01	-2.59	.01
	$F(3; 157) = 0.95$; $R^2 = .02$; $p = .42$				$F(3; 157) = 2.51$; $R^2 = .05$; $p = .06$			
Step 3: Reading Intervention Group vs. Other English Intervention Groups (Step 3)								
Constant	0.00 [-0.04; 0.03]	0.02	-0.19	.85	0.06 [0.01; 0.11]	0.02	2.44	.02
Group (reading English vs. other English)	-0.04 [-0.11; 0.04]	0.04	-0.90	.37	0.06 [-0.03; 0.16]	0.05	1.30	.19
Work Avoidance	0.00 [0.00; 0.01]	0.00	1.17	.24	-0.01 [-0.01; 0.00]	0.00	-1.95	.05
Group x Work Avoidance	-0.01 [-0.02; 0.00]	0.01	-1.54	.13	0.00 [-0.02; 0.01]	0.01	-0.44	.66
	$F(3; 153) = 1.50$; $R^2 = .03$; $p = .22$				$F(3; 153) = 2.02$; $R^2 = .04$; $p = .11$			
Constant	-0.01 [-0.05; 0.03]	0.02	-0.32	.75	0.06 [0.01; 0.11]	0.03	2.50	.01
Group (reading English vs. other English)	-0.04 [-0.12; 0.03]	0.04	-1.11	.27	0.09 [-0.01; 0.19]	0.05	1.71	.09
Self-Concept (criterial)	0.01 [-0.01; 0.02]	0.01	0.91	.36	-0.02 [-0.03; 0.00]	0.01	-2.00	.05
Group x Self-Concept (criterial)	0.02 [-0.01; 0.04]	0.01	1.36	.18	-0.01 [-0.04; 0.02]	0.02	-0.35	.72
	$F(3; 153) = 1.08$; $R^2 = .02$; $p = .36$				$F(3; 153) = 2.00$; $R^2 = .04$; $p = .12$			
Constant	-0.01 [-0.04; 0.03]	0.02	-0.26	.79	0.06 [0.01; 0.11]	0.02	2.47	.01
Group (reading English vs. other English)	-0.04 [-0.11; 0.04]	0.04	-0.94	.35	0.06 [-0.02; 0.18]	0.05	1.62	.11
Self-Concept (individual)	0.00 [-0.01; 0.01]	0.00	-0.40	.69	-0.01 [-0.02; 0.00]	0.00	-2.68	.01
Group x Self-Concept (individual)	0.02 [0.00; 0.03]	0.01	2.08	.04	0.00 [-0.02; 0.02]	0.01	0.18	.86
	$F(3; 153) = 1.83$; $R^2 = .03$; $p = .14$				$F(3; 153) = 3.10$; $R^2 = .06$; $p = .03$			
Constant	-0.01 [-0.05; 0.03]	0.02	-0.32	.75	0.06 [0.01; 0.11]	0.03	2.46	.02
Group (reading English vs. other English)	-0.04 [-0.12; 0.04]	0.04	-1.01	.31	0.08 [-0.01; 0.18]	0.05	1.68	.09
Self-Concept (social)	0.00 [-0.01; 0.01]	0.01	0.30	.76	-0.01 [-0.03; 0.00]	0.01	-2.03	.04
Group x Self-Concept (social)	0.02 [0.00; 0.04]	0.01	1.55	.12	0.00 [-0.03; 0.03]	0.01	-0.11	.91
	$F(3; 153) = 1.10$; $R^2 = .02$; $p = .35$				$F(3; 153) = 2.09$; $R^2 = .04$; $p = .10$			
Constant	-0.01 [-0.04; 0.03]	0.02	-0.49	.62	0.07 [0.02; 0.12]	0.03	2.54	.01
Group (reading English vs. other English)	-0.04 [-0.12; 0.03]	0.04	-1.14	.26	0.09 [-0.02; 0.19]	0.05	1.73	.09
Self-Concept (absolute)	0.00 [-0.01; 0.02]	0.01	0.84	.40	-0.01 [0.00; 0.01]	0.01	-1.97	.05
Group x Self-Concept (absolute)	0.03 [0.00; 0.05]	0.01	2.27	.02	-0.01 [-0.02; 0.01]	0.01	-0.65	.52
	$F(3; 153) = 2.07$; $R^2 = .04$; $p = .11$				$F(3; 153) = 1.98$; $R^2 = .04$; $p = .12$			
Constant	-0.10 [-0.04; 0.03]	0.11	-0.86	.39	-0.35 [-0.63; -0.07]	0.14	-2.44	.02
Group (reading English vs. other English)	0.07 [-0.12; 0.03]	0.08	0.99	.32	0.27 [0.09; 0.46]	0.10	2.88	.005
Cloze Test Points T1	0.02 [-0.01; 0.02]	0.01	1.53	.13	0.04 [0.01; 0.08]	0.02	2.54	.01
Group x Cloze Test Points T1	-0.02 [0.00; 0.05]	0.01	-1.71	.09	-0.03 [-0.05; -0.01]	0.01	-2.53	.01
	$F(3; 153) = 2.07$; $R^2 = .04$; $p = .11$				$F(3; 153) = 2.87$; $R^2 = .05$; $p = .04$			
Step 4: Intensive Reading Intervention Groups vs. Extensive Intervention Groups (Step 4)								
Constant	0.03 [-0.03; 0.09]	0.03	1.01	.31	0.04 [-0.04; 0.11]	0.04	0.96	-34
Group (intensive vs. extensive reading)	0.02 [-0.11; 0.14]	0.06	0.29	.77	-0.06 [-0.22; 0.11]	0.08	-0.71	.48
leisure reading behavior	0.00 [-0.01; 0.01]	0.00	0.94	.35	0.01 [0.00; 0.02]	0.01	1.30	.19

Group x reading behavior	0.00 [-0.02; 0.02]	0.01	0.15	.89	0.03 [0.00; 0.06]	0.01	2.18	.03
	$F(3; 67) = 0.34$; $R^2 = .02$; $p = .80$				$F(3; 67) = 2.57$; $R^2 = .10$; $p = .06$			
Constant	0.02 [-0.04; 0.07]	0.03	0.59	.56	0.03 [-0.04; 0.10]	0.04	0.79	.43
Group (intensive vs. extensive reading)	-0.02 [-0.13; 0.09]	0.05	-0.30	.77	-0.01 [-0.17; 0.14]	0.08	-0.17	.86
Self-Concept (individual)	-0.01 [-0.02; 0.00]	0.00	1.96	.05	-0.01 [-0.03; 0.00]	0.01	-1.89	.06
Group x Self-Concept (individual)	-0.01 [-0.03; 0.01]	0.01	-1.02	.31	-0.01 [-0.02; 0.03]	0.01	0.49	.62
	$F(3; 80) = 1.48$; $R^2 = .05$; $p = .23$				$F(3; 80) = 1.41$; $R^2 = .05$; $p = .25$			
Constant	-0.01 [-0.04; 0.07]	0.14	-0.10	.92	-0.02 [-0.35; 0.31]	0.17	-0.13	.90
Group (intensive vs. extensive reading)	0.08 [-0.13; 0.09]	0.08	0.96	.34	0.18 [-0.02; 0.37]	0.10	1.80	.08
Non-fictional reading comprehension points T1	0.14 [-0.02; 0.00]	0.40	0.34	.73	0.08 [-0.87; 1.03]	0.48	0.17	.87
Group x non-fictional RC points T1	-0.28 [-0.03; 0.01]	0.23	-1.24	.22	-0.55 [-1.10; 0.00]	0.28	-1.97	.05
	$F(3; 80) = 3.68$; $R^2 = .12$; $p = .02$				$F(3; 80) = 14.44$; $R^2 = .35$; $p = .00$			

10.3.2.4 Moderation Analyses for Preposition Test Results

	direct pre-/posttest				delayed pre-posttest			
Variables	***b***	***SE B***	***t***	***p***	***b***	***SE B***	***t***	***p***
Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups (Step 2)								
Constant	0.01 [-0.01; 0.04]	0.01	1.20	.23	-0.01 [-0.03; 0.02]	0.01	-0.36	.72
Group (non-English vs. reading English)	0.00 [-0.05; 0.04]	0.02	-0.14	.89	0.00 [-0.05; 0.06]	0.03	0.03	.98
Performance Avoid Orientation	-0.01 [-0.01; 0.00]	0.00	-2.55	.01	0.00 [-0.01; 0.00]	0.00	-1.12	.27
Group x Performance Avoid Orientation	-0.01 [-0.02; 0.00]	0.00	-1.15	.25	-0.01 [-0.02; 0.01]	0.01	-0.97	.33
	$F(3; 157) = 2.63$; $R^2 = .05$; $p = .05$				$F(3; 157) = 0.74$; $R^2 = .01$; $p = .53$			
Constant	0.03 [-0.10; 0.15]	0.06	0.41	.68	0.10 [-0.04; 0.25]	0.07	1.47	.14
Group (non-English vs. reading English)	-0.02 [-0.10; 0.07]	0.04	-0.41	.68	-0.01 [-0.10; 0.09]	0.05	-0.16	.87
Preposition test points T1	-0.05 [-0.51; 0.41]	0.23	-0.23	.82	-0.53 [-1.04; -0.03]	0.26	-2.08	.04
Group x preposition test points T1	0.08 [-0.23; 0.38]	0.15	0.50	.62	0.08 [-0.26; 0.41]	0.17	0.44	.66
	$F(3; 157) = 0.27$; $R^2 = .01$; $p = .85$				$F(3; 157) = 8.68$; $R^2 = .14$; $p = .00$			
Reading Intervention Group vs. Other English Intervention Groups (step 3)								
Constant	-0.06 [-0.19; 0.08]	0.07	-0.85	.39	0.14 [-0.00; 0.28]	0.07	1.94	0.05
Group (reading English vs. other English)	0.07 [-0.03; 0.16]	0.05	1.43	.15	-0.04 [-0.14; 0.05]	0.05	-0.88	0.38
Preposition test points T1	0.34 [-0.14; 0.82]	0.24	1.38	.17	-0.79 [-1.30; -0.28]	0.26	-3.06	0.00
Group x Preposition test points T1	-0.31 [-0.64; 0.01]	0.17	-1.89	.06	0.33 [-0.02; 0.68]	0.18	1.88	0.06
	$F(3; 153) = 1.72.$; $R^2 = .03$; $p = .17$				$F(3; 153) = 6.54.$; $R^2 = .11$; $p = .0003$			
Intensive Reading Intervention Groups vs. Extensive Intervention Groups (step 4)								
Constant	-0.03 [-0.25; 0.18]	0.11	-0.31	.76	0.17 [-0.04; 0.37]	0.10	1.62	.11
Group (intensive vs. extensive reading)	0.03 [-0.10; 0.15]	0.06	0.42	.68	-0.04 [-0.16; 0.08]	0.06	-0.66	.51
Preposition test points T1	0.06 [-0.78; 0.90]	0.42	0.14	.89	-1.15 [1.96; -0.35]	0.41	-2.85	.01
Group x Preposition test points T1	-0.02 [-0.50; 0.45]	0.24	-0.10	.92	0.40 [-0.06; 0.85]	0.23	1.74	.09
	$F(3; 80) = 0.13$; $R^2 = .00$; $p = .94$				$F(3; 80) = 8.24$; $R^2 = .24$; $p = .0001$			

10.3.2.5 Moderation Analyses for Fictional Reading Comprehension Test Results

	direct pre-/posttest				delayed pre-posttest			
Variables	***b***	***SE B***	***t***	***p***	***b***	***SE B***	***t***	***p***
Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups (Step 2)								
Constant	0.01 [-0.02; 0.05]	0.02	0.70	.48	0.00 [-0.04; 0.04]	0.02	-0.14	.89
Group (non-English vs. reading English)	-0.02 [-0.09; 0.05]	0.04	-0.68	.50	0.01 [-0.07; 0.10]	0.04	0.33	.74
Learning Goal Orientation	-0.01 [-0.02; 0.00]	0.01	-2.13	.03	-0.01 [-0.02; 0.00]	0.01	-1.35	.18
Group x Learning Goal Orientation	0.00 [-0.02; 0.02]	0.01	-0.25	.80	0.00 [-0.03; 0.02]	0.01	-0.29	.77
	$F(3; 157) = 1.76$; $R^2 = .03$; $p = .16$				$F(3; 157) = 0.73$; $R^2 = .01$; $p = .54$			
Constant	0.01 [-0.02; 0.05]	0.02	0.75	.45	0.00 [-0.04; 0.04]	0.02	-0.07	.94
Group (non-English vs. reading English)	-0.02 [-0.09; 0.05]	0.04	-0.61	.54	0.02 [-0.06; 0.10]	0.04	0.51	.61
Self-Concept (criterial)	0.00 [-0.01; 0.01]	0.01	-0.46	.65	-0.01 [-0.03; 0.00]	0.01	-2.39	.02
Group x Self-Concept (criterial)	-0.01 [-0.03; 0.01]	0.01	-0.88	.38	-0.01 [-0.04; 0.01]	0.01	-1.03	.31
	$F(3; 157) = 0.44$; $R^2 = .01$; $p = .73$				$F(3; 157) = 2.12$; $R^2 = .04$; $p = .10$			
Constant	0.17 [-0.01; 0.35]	0.09	1.86	.06	0.14 [-0.07; 0.34]	0.10	1.29	.20
Group (non-English vs. reading English)	-0.07 [-0.18; 0.04]	0.06	-1.21	.23	-0.05 [-0.18; 0.08]	0.07	-0.72	.47
Non-fictional reading comprehension points T1	-0.43 [-0.91; 0.04]	0.24	-1.82	.07	-0.57 [-1.12; -0.02]	0.28	-2.06	.04
Group x Non-fictional RC points T1	0.18 [-0.11; 0.47]	0.15	1.20	.23	0.24 [-0.10; 0.58]	0.17	1.39	.17
	$F(3; 157) = 2.23$; $R^2 = .04$; $p = .09$				$F(3; 157) = 2.61$; $R^2 = .05$; $p = .05$			
Constant	0.22 [0.03; 0.42]	0.10	2.32	.02	0.09 [-0.13; 0.32]	0.11	0.80	.43
Group (non-English vs. reading English)	-0.12 [-0.25; 0.01]	0.06	-1.84	.07	-0.06 [-0.21; 0.09]	0.08	-0.74	.46
Preposition test points T1	-0.80 [-1.48; -0.10]	0.35	-2.27	.02	-0.51 [-1.33; 0.30]	0.41	-1.24	.22
Group x preposition test points T1	0.42 [-0.03; 0.88]	0.23	1.83	.07	0.31 [-0.23; 0.85]	0.27	1.14	.26
	$F(3; 157) = 2.16$; $R^2 = .04$; $p = .10$				$F(3; 157) = 0.55$; $R^2 = .01$; $p = .65$			
Constant	0.16 [-0.06; 0.38]	0.11	1.43	.15	0.24 [0.03; 0.44]	0.10	2.26	.03
Group (non-English vs. reading English)	-0.02 [-0.16; 0.12]	0.07	-0.27	.79	0.04 [-0.09; 0.18]	0.07	0.66	.51
Fictional reading comprehension points T1	-0.27 [-0.73; 0.19]	0.23	-1.17	.25	-0.63 [-1.07; -0.20]	0.22	-2.89	.004
Group x Fictional RC points T1	-0.01 [-0.30; 0.29]	0.15	-0.05	.96	-0.07 [-0.34; 0.21]	0.14	-0.46	.64
	$F(3; 157) = 4.86$; $R^2 = .09$; $p = .003$				$F(3; 157) = 35.79$; $R^2 = .41$; $p = .00$			
Reading Intervention Group vs. Other English Intervention Groups (step 3)								
Constant	0.21 [0.03; 0.39]	0.09	2.36	.02	0.16 [-0.02; 0.35]	0.09	1.74	0.08
Group (reading English vs. other English)	-0.12 [-0.23; 0.00]	0.06	-2.03	.04	-0.08 [-0.19; 0.04]	0.06	-1.29	0.20
Non-fictional reading comprehension points T1	-0.59 [-1.09; -0.09]	0.25	-2.32	.02	-0.59 [-1.10; -0.08]	0.26	-2.27	0.02
Group x Non-fictional RC points T1	0.33 [0.01; 0.65]	0.16	2.01	.05	0.26 [-0.08; 0.59]	0.17	1.52	0.13
	$F(3; 153) = 2.05.$; $R^2 = .04$; $p = .11$				$F(3; 153) = 2.96.$; $R^2 = .05$; $p = .03$			
Constant	0.27 [0.04; 0.51]	0.12	2.31	.02	0.30 [0.10; 0.49]	0.10	3.01	0.00
Group (reading English vs. other English)	-1.14 [-0.29; 0.02]	0.08	-1.70	.09	-0.02 [-0.15; 0.11]	0.07	-0.28	0.78
Fictional reading comprehension points T1	-0.55 [-1.04; -0.05]	0.25	-2.17	.03	-0.74 [-1.16; -0.33]	0.21	-3.54	0.00
Group x Fictional RC points T1	0.27	0.17	1.58	.12	0.04	0.14	0.31	0.76

	[-0.07; 0.60]				[-0.24; 0.32]			
	$F(3; 153) = 2.39.; R^2 = .04; p = .07$				$F(3; 153) = 32.33.; R^2 = .39; p = .00$			
Intensive Reading Intervention Groups vs. Extensive Intervention Groups (step 4)								
Constant	0.03 [-0.03; 0.08]	0.03	1.02	.31	-0.01 [-0.06; 0.05]	0.03	-0.27	.79
Group (intensive vs. extensive reading)	0.03 [-0.08; 0.14]	0.06	0.52	.60	0.05 [-0.07; 0.17]	0.06	0.89	.38
Self-Concept (criterial)	0.00 [-0.02; 0.01]	0.01	-0.13	.90	-0.01 [-0.03; 0.00]	0.01	-1.43	.16
Group x Self-Concept (criterial)	-0.03 [-0.06; 0.00]	0.01	-1.98	.05	-0.02 [-0.06; 0.01]	0.02	-1.49	.14
	$F(3; 80) = 1.44; R^2 = .05; p = .24$				$F(3; 80) = 1.45; R^2 = .05; p = .24$			
Constant	-0.94 [-0.60; -0.29]	0.33	-2.87	.01	-0.54 [-1.28; 0.21]	0.38	-1.43	.16
Group (intensive vs. extensive reading)	0.49 [0.15; 0.82]	0.17	2.89	.01	0.27 [-0.11; 0.66]	0.19	1.42	.16
Years of English in primary school	0.24 [0.07; 0.40]	0.08	2.83	.01	0.12 [-0.07; 0.31]	0.10	1.28	.20
Group x years of English in primary school	-0.12 [-0.20; -0.04]	0.04	-2.82	.01	-0.06 [-0.16; 0.03]	0.05	-1.30	.20
	$F(3; 76) = 2.84; R^2 = .10; p = .04$				$F(3; 76) = 0.68; R^2 = .03; p = .57$			
Constant	0.31 [-0.03; 0.66]	0.71	1.80	.08	0.02 [-0.36; 0.40]	0.19	0.12	.90
Group (intensive vs. extensive reading)	-0.17 [-0.36; 0.03]	0.10	-1.65	.10	-0.003 [-0.22; 0.22]	0.11	-0.03	.98
Cloze Test Points T1	-0.06 [-0.11; -0.01]	0.02	-2.38	.02	-0.02 [-0.07; 0.03]	0.03	-0.79	.43
Group x Cloze Test Points T1	-0.12 [0.01; 0.06]	0.01	2.42	.02	0.01 [-0.02; 0.04]	0.01	0.68	.50
	$F(3; 80) = 2.08; R^2 = .07; p = .11$				$F(3; 80) = 0.51; R^2 = .02; p = .68$			
Constant	0.19 [-0.19; 0.56]	0.19	0.99	.32	0.37 [0.06; 0.68]	0.16	2.39	.02
Group (intensive vs. extensive reading)	-0.03 [-0.25; 0.19]	0.11	-0.24	.81	-0.05 [-0.23; 0.13]	0.09	-0.56	.58
Fictional reading comprehension points T1	-0.62 [-1.47; -0.23]	0.43	-1.46	.15	-1.34 [-2.05; -0.64]	0.35	-3.81	.00
Group x Fictional RC points T1	-0.20 [-0.29; 0.68]	0.24	0.81	.42	0.37 [-0.03; 0.77]	0.20	1.83	.07
	$F(3; 80) = 2.64; R^2 = .09; p = .06$				$F(3; 80) = 22.18; R^2 = .45; p = .00$			

10.3.2.6 Moderation Analyses for School Grade Difference

	direct pre-/posttest			
Variables	***b***	***SE B***	***t***	***p***
Intensive or Extensive Reading Intervention Groups vs. non-English Intervention Groups (Step 2)				
Constant	0.36 [0.25; 0.47]	0.06	6.48	.00
Group (non-English vs. reading English)	0.03 [-0.19; 0.25]	0.11	0.24	81
Self-Concept (absolute)	0.04 [0.01; 0.07]	0.02	2.55	.01
Group x Self-Concept (absolute)	-0.07 [-0.13; -0.01]	0.03	-2.16	.03
	$F(3; 154) = 2.71; R^2 = .08\ p = .01$			
Step 3: Reading Intervention Group vs. Other English Intervention Groups				
Constant	0.32 [0.21; 0.43]	0.06	5.57	.00
Group (reading English vs. other English)	0.01 [-0.22; 0.24]	0.11	0.08	.93
Performance Avoid Orientation	0.03 [0.01; 0.05]	0.01	2.64	.01
Group x Performance Avoid Orientation	0.02 [-0.02; 0.07]	0.02	0.99	.32
	$F(3; 150) = 2.64.; R^2 = .05; p = .05$			
Constant	0.33 [0.22; 0.45]	0.06	5.77	.00

Group (reading English vs. other English)	-0.02 [-0.25; 0.21]	0.12	-0.19	.85
Self-Concept (criterial)	-0.02 [-0.01; 0.06]	0.02	1.37	.17
Group x Self-Concept (criterial)	-0.07 [-0.14; 0.00]	0.04	-2.05	.04
	$F(3; 150) = 2.41$; $R^2 = .05$; $p = .07$			
Constant	0.34 [0.22; 0.45]	0.06	5.90	.00
Group (reading English vs. other English)	-0.05 [-0.27; 0.18]	0.11	-0.40	.69
Self-Concept (absolute)	0.04 [0.01; 0.07]	0.02	2.61	.01
Group x Self-Concept (absolute)	-0.07 [-0.13; 0.00]	0.03	-2.04	.04
	$F(3; 150) = 4.57$; $R^2 = .08$; $p = .01$			
Step 4: Intensive Reading Intervention Groups vs. Extensive Intervention Groups				
Constant	0.23 [0.07; 0.39]	0.08	2.79	.01
Group (intensive vs. extensive reading)	-0.01 [-0.36; 0.35]	0.18	-0.04	.97
leisure reading behavior	-0.01 [-0.03; 0.02]	0.01	-0.38	.70
Group x reading behavior	-0.07 [.0.13; 0.00]	0.03	-2.14	.04
	$F(3; 66) = 1.66$; $R^2 = .07$; $p = .18$			
Constant	0.32 [0.17; 0.46]	0.07	4.43	.00
Group (intensive vs. extensive reading)	-0.16 [-0.47; 0.14]	0.15	-1.06	.29
Self-Concept (criterial)	0.06 [-0.02; 0.10]	0.02	2.91	.01
Group x Self-Concept (criterial)	0.01 [-0.07; 0.09]	0.04	0.15	.88
	$F(3; 79) = 3.13$; $R^2 = .11$; $p = .03$			
Constant	0.32 [0.17; 0.46]	0.07	4.31	.00
Group (intensive vs. extensive reading)	-0.16 [-0.48; 0.15]	0.16	-1.04	.30
Self-Concept (social)	0.04 [0.00; 0.08]	0.02	-2.02	.05
Group x Self-Concept (social)	0.00 [-0.08; 0.08]	0.04	0.07	.94
	$F(3; 79) = 1.61$; $R^2 = .06$; $p = .19$			
Constant	0.32 [0.18; 0.45]	0.07	4.68	.00
Group (intensive vs. extensive reading)	-0.15 [-0.44; 0.14]	0.15	-1.05	.30
Self-Concept (absolute)	0.07 [0.04; 0.11]	0.02	4.09	.00
Group x Self-Concept (absolute)	-0.01 [-0.09; 0.06]	0.04	-0.38	.70
	$F(3; 79) = 6.14$; $R^2 = .19$; $p = .001$			