From letters to words, from digits to numbers:

Similarities, differences and modal influences on writing processes from different perspectives

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Log. Stefanie Jung, geb. Keim
aus Berlin

Tübingen
2019

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

Tag der mündlichen Qualifikation:	04.03.2020
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Jürgen Heller
2. Berichterstatter:	Prof. Dr. Korbinian Möller
3. Berichterstatter	Prof. Dr. Margarete Delazer

Summary

The evolution of writing is closely related to the evolution of the (Arabic) number system and the notation of quantity. Yet, distinct scripts have developed for writing letters and digits. Domain-specific cognitive processes involved in writing of those scripts have merely been addressed by linguistic and numerical research within their respective domain. Domain-general cognitive processes involved in the graphomotor execution of writing reflect the common origins of both scripts. However, they have not yet been investigated in an integrated theoretic model on writing. Nowadays, digital media is increasingly replacing handwriting. Consequently, both domain-specific and domain-general cognitive processes of writing are subject to changes in the execution of writing. Understanding these changes when switching from handwriting to typing in both the language and the number domain seems essential to explain the different influences of writing media (i.e., handwriting or typing) on orthography and arithmetic skills.

This issue is addressed in the present thesis by means of six studies: beginning with the cognitive study of writing alphabetic and numerical scripts in general and number transcoding in particular, domain-specific and domain-general aspects of writing in both domains are investigated on the behavioral and the neuro-cognitive level. Typical, erroneous and impaired writing processes as well as arithmetic skills are examined in both typically and atypically developing children as well as in neurological patients. The influence of the actual writing mode (i.e., handwriting or typing) on orthography and arithmetic skills is further investigated to take into account the increasing change of writing from handwriting to typing caused by the use of digital media. Application of digital media in the diagnosis and intervention of orthography and arithmetic skills is finally outlined. Insights gained from these studies provide the foundation for the development of an integrated theoretical writing model which is extended to writing with digital media. This model is then evaluated with regard to its validity to comprehensively represent typical as well as impaired writing processes across domains and writing modalities.

By integrating different writing domains and modalities into a common theoretical writing model, the present thesis contributes to the investigation of cognitive processes involved in writing, keeping up with the latest methods and results in the study of writing as well as its evolution.

Zusammenfassung

Die Entwicklung der Schrift ist eng mit der Entwicklung des (arabischen) Zahlensystems und der Darstellung von Mengen verbunden. Dennoch haben sich unterschiedliche Schriftsysteme für Buchstaben und Zahlen entwickelt. Die domänenspezifischen kognitiven Prozesse, die den Schriftsystemen unterliegen, wurden von der Sprach- und Zahlenverarbeitungsforschung vorwiegend jeweils in ihrem Forschungsfeld betrachtet. Domänenübergreifende kognitive Prozesse, die an der graphomotorischen Ausführung des Schreibens beteiligt sind, spiegeln die gemeinsamen Wurzeln beider Schriftsysteme wider, wurden jedoch bisher nicht gemeinsam untersucht. Mit zunehmender Nutzung digitaler Medien ist davon auszugehen, dass sich sowohl die domänenspezifischen als auch die domänenübergreifenden kognitive Prozesse des Schreibens verändern. Diese Veränderungen zu verstehen ist unerlässlich, um die unterschiedlichen Einflüsse digitaler Medien auf orthographische und arithmetische Fähigkeiten erklären zu können.

Dieses Thema wird in der vorliegenden Arbeit anhand von sechs Studien behandelt: Ausgehend von den kognitiven Prozessen, die beim Schreiben des alphabetischen Schriftsystems, der Zahlen und des numerischen Transkodierens beteiligt sind, werden domänenspezifische und domänenübergreifende Aspekte des Schreibens auf Verhaltens- und neurokognitiver Ebene betrachtet. Typische, fehlerhafte und beeinträchtigte Prozesse des Schreibens sowie der arithmetischen Fähigkeiten werden bei Kindern mit und ohne besonderem Förderungsbedarf sowie bei neurologischen Patienten untersucht. Mit der Untersuchung verschiedener Modalitäten des Schreibens (d.h. handschriftliches Schreiben oder Tippen am Computer) wird der zunehmenden Veränderung des Schreibens durch den Einsatz digitaler Medien Rechnung getragen. Anschließend wird die Anwendung digitaler Medien für die Diagnostik orthographischer und arithmetischer Fähigkeiten und deren Förderung dargestellt. Die so gewonnenen Erkenntnisse bilden die Grundlage für die Entwicklung eines integrierten theoretischen Modells für das Schreiben von Buchstaben und Zahlen in unterschiedlichen Modalitäten. Dieses Modell wird abschließend dahingehend evaluiert, ob es typische sowie beeinträchtigte kognitive Prozesse des Schreibens über Domänen und Modalitäten hinweg umfassend abbilden kann.

Mit der Integration verschiedener Domänen und Modalitäten in ein gemeinsames Schreibmodell schlägt die vorliegende Arbeit einen theoretischen Rahmen für die Untersuchung der am Schreiben beteiligten kognitiver Prozesse vor, der sowohl aktuelle Forschungsergebnisse erklären als auch die Entwicklung des Schreibens abbilden kann.

Without writing,
nothing would ever be written.

Table of contents

PART III: GENERAL CONCLUSION

List of Figures

PART I: GENERAL INTRODUCTION

PART II: EMPIRICAL STUDIES

Section 1: Writing letters and digits: similar but not the same

Study 1: What letters can 'learn' from Arabic digits - fMRI-controlled single case therapy study of peripheral agraphia

Study 2: Differing connectivity of Exner's area for numbers and letters.

Section 2: Handwriting and typing: Modal influences on spelling and arithmetic

Study 3: Mode effect: an issue of perspective? The influence of writing mode in children with and without dyslexia

Study 4: Closing the gap. Bounded and unbounded number line estimation in secondary school children

Section 3: Writing and beyond: Assessment and training of orthography and numeracy

Study 5: Die TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)

Study 6: Behavioral and Neurocognitive Evaluation of a Web-Platform for Game-Based Learning of Orthography and Numeracy.

PART III: GENERAL DISCUSSION

List of Tables

PART I: GENERAL INTRODUCTION

PART II: EMPIRICAL STUDIES

Studies enclosed in the present thesis

Section 1: Writing letters and digits -similar but not the same

Study 1: **Jung, S**., Halm, K., Huber, W., Willmes, K., & Klein, E. (2015). What letters can "learn" from Arabic digits - fMRI-controlled single case therapy study of peripheral agraphia. *Brain and Language, 149*, 13-26.

Study 2: Klein, E., Willmes, K., **Jung, S**., Huber, S., Braga, L. W., & Moeller, K. (2016). Differing connectivity of Exner's area for numbers and letters. *Frontiers in Human Neuroscience, 10:281*.

Section 2: Handwriting and typing - modal influences on spelling and arithmetic

Study 3: **Jung, S**., Moeller, K., Klein, E., & Heller, J. (in revision). Mode effect: an issue of perspective? The influence of writing mode in children with and without dyslexia. *Assessing Writing.*

Study 4: **Jung, S**., Roesch, S., Klein, E., Dackermann, T., Heller, J. & Moeller, K., (in revision). Closing the gap. Bounded and unbounded number line estimation in secondary school children *Cognitive Development.*

Section 3: Writing and beyond - assessment and training of spelling and arithmetic

Study 5: **Jung, S**., Huber, S., Heller, J., Grust, T., Moeller, K., & Nuerk, H.-C. (2016). Die TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE): Individualisierte Förderung durch adaptive Lernspiele. *Lernen und Lernstörungen, 5(1)*, 7-15.

Study 6: Soltanlou, M., **Jung, S.,** Roesch, S., Ninaus, M., Brandelik, B., Heller, J., Grust, T., Nuerk, H.-C., & Moeller, K. (2017). Behavioral and Neurocognitive Evaluation of a Web-Platform for Game-Based Learning of Orthography and Numeracy. In J. Buder & F.W. Hesse (Eds.), *Informational Environments: Effects of Use, Effective Designs* (pp. 149-176). Cham, Switzerland: Springer International Publishing.

PART I: GENERAL INTRODUCTION

General introduction

Human cognition, its development and its impairments are studied in a wide range of scientific domains. In this vein, various domain-specific theoretical approaches have been proposed to explain typical and atypical cognitive processes as well as their neuronal correlates, each of them within a certain domain (e.g., language processing, memory, attention). Such a domain-specific distinction is, however, hardly justified by the architecture and function of the human brain – the joint subject of all these different research lines. Accordingly, such a distinction is neither reflected in data at the behavioral nor at the neural level.

On the contrary, at the behavioral level, cognitive abilities can dissociate. At the neuronal level, there is considerable consensus “that no brain region is by itself sufficient to perform a particular cognitive [...] process” (Eickhoff & Grefkes, 2011, p. 107). Each of the cognitive processes rely on the interplay of different brain areas and, thus, on the integrity of a widespread network integrating cortical areas across the entire brain via fiber pathways. From both levels it becomes obvious that human cognition in its entire complexity can only be understood from a cross-domain perspective that systematically identifies similarities and differences during cognitive processing in different domains. Nevertheless, each research domain still looks through its own scientific spectacles and may overlook a possible more general cross-domain notion of cognitive processing within the human cognitive system.

In the present thesis, I set aside these domain-specific spectacles. In particular, I aimed at integrating existing theoretical approaches of different domains into a joint cross-domain approach, using two very similar cognitive tasks, the writing of letters and numbers. Writing of letters and numbers has a joint origin. In the course of evolution, however, distinct notational systems have developed (Schmandt-Besserat & Erard, 2008), and have since been studied by different research perspectives (i.e., linguistics and numerical cognition). However, both letters and numbers are closely associated via their verbalized form, making it abundantly obvious to investigate them from a shared research perspective. By adopting this perspective, the superordinate goal of this thesis was to bring together the various theoretical approaches toward the cognitive study of writing both letters and numbers and to develop an integrated writing model for the cognitive processes required for writing in different notation systems.

To pursue this goal, I investigated domain-specific and domain-general cognitive mechanisms and neural correlates underlying writing of letters and numbers as well as the influence of modern writing technologies (i.e., computer keyboard typing) on writing performance. These investigations were carried out in different populations (i.e., neurological patients and typically and atypically developing children) by means of different writing technologies (i.e., handwriting and typing), in different clinical and learning situations (i.e., diagnosis and intervention of impaired writing and arithmetic). To this end, six studies were conducted, which addressed three sets of research questions, and hence were thematically grouped into three sections: in Section 1, neuropsychological case studies are presented, which show that the existing cognitive writing models in both domains are not sufficient to explain the observed writing performance. In Section 2, the influence of the actual writing mode (i.e., handwriting or typing) on spelling and arithmetic is investigated in both typically developing children and children with developmental dyslexia. In Section 3, the application of digital media in the diagnosis and intervention of orthography and arithmetic skills is examined.

Based on the insights gained from these studies, I propose a novel integrated writing model of alphabetic and numerical scripts in the discussion section, which covers technological advances in writing over the last decades by including different writing modalities. This integrated writing model is not only intended to provide a better understanding of the cognitive processes and neural correlates involved in writing letters and numbers, but it also serves as a diagnostic and interventional foundation for atypically developed writing skills or acquired writing disorders. The model thus has clearly relevance for writing in different contexts and due to its cross-domain approach also beyond.

In the following, I will introduce the different theories on writing letters and words as well as on writing digits and numbers in the general introduction section to acquaint the reader with the neuro-cognitive study of writing. Simultaneously, this introduction reveals the model's shortcomings in each domain, motivates the research objectives, and emphasizes the need for a more comprehensive writing model. In the general discussion, findings from the empirical studies will be evaluated for their integration into findings of the current literature and for their relevance to the development of a novel integrated writing model. Subsequently, the integrated writing model is outlined and evaluated in terms of its prediction for both typically and impaired writing processes across domains. Finally, future directions of research based on my integrated writing model conclude the present thesis.

1 Origins of alphabetic and numerical scripts

The evolution of writing has gone through ten thousand years of development, from rather artistic cave paintings more than at least 30,000 years ago to the use of alphabetic scripts written by hand, and, later, with analogue typewriters. In recent decades these types of writing have increasingly been replaced by new digital writing media (i.e., computers and tablets). The development of writing was closely related to the evolution of digits enabling (permanent) notation of quantities of, for instance, animals, goods or anything countable. Abstraction of the quantity from the actual designated object and its phonological equivalent has generated different sign systems for writing letters and digits (Schmandt-Besserat & Erard, 2008).

In the past century, at the latest, the origins as well as the evolution of writing letters and numbers have been the subject of growing scientific interest. Since then, both linguistic and numerical research have been investigating the cognitive mechanisms underlying writing in the respective domain. Investigations also included the issue how the human cognitive system has adapted to the evolution of writing. For this purpose, various theories have been proposed regarding the cognitive mechanisms of typical writing (for linguistic research: e.g., Ellis, 1982, 1988; Graham & Weintraub, 1996; Margolin, 1984; Van Galen, 1991; for numerical research: Amalric & Dehaene, 2018; Barrouillet, Camos, Perruchet, & Seron, 2004; Cipolotti & Butterworth, 1995; Dehaene, 1992; Dehaene & Cohen, 1995; Dehaene, Piazza, Pinel, & Cohen, 2003; Deloche & Seron, 1987; McCloskey, 1992; Verguts & Fias, 2006). Such theories specify typical cognitive representations necessary for a function (i.e., writing letters or digits). Hence, they are essential to understand and to investigate both developmental as well as acquired deficits specifically affecting this function (McCloskey & Rapp, 2017). Within this framework, they also provide the basis for the development of individually tailored diagnostics and interventions. This is of particular importance as mastering literacy (i.e., reading and writing capabilities) and numeracy in general as well as reading and writing of numbers in particular (Lopes-Silva et al., 2014) is an essential skill for everyday life in digital ages. Disorders affecting these skills are associated with severe disadvantages for the individual but also for society as a whole (e.g., Beddington et al., 2008; Daniel et al., 2006; Gross, 2006; Gross et al., 2009; Parsons & Brynner, 2005).

For several decades, digital media have been part of everyday life with considerable influence on leisure activities, communication, and writing. Digital media have not only changed writing habits (Mangen & Velay, 2010, for a review), but also offer new approaches to promote literacy and numeracy (i.e., web-based writing programs, game-based learning). Research on this issue is still in its infancy but studies investigating digital learning environments showed positive learning effects (Qian and Clark, 2016; for a review). However, these studies neither provided explanations for the positive learning effects on a theoretical basis nor directly evaluated the influence of the test mode (i.e., mode effect: differences in performance that only result from the applied medium). Against this background, theoretical considerations on writing acknowledge the development of writing toward digital media in order to reflect the changes to the underlying cognitive mechanisms (Pinet, Ziegler, & Alario, 2016).

Despite common origins in the evolution of writing (Schmandt-Besserat & Erard, 2008), linguistic and numerical research have addressed cognitive mechanisms of writing in theoretical approaches differently within their respective domain. Domain-specific approaches are crucial to identify cognitive mechanisms exclusively related to a specific function in a specific domain. Domain-general approaches, instead, aim at identifying common cognitive mechanisms related to a specific function in two or more domains. As regards research on writing, domain-specific linguistic and numerical processing certainly differ at the central processing level. The mere graphomotor act of writing at the peripheral processing level (i.e., execution of writing movements), can be assumed to be general for both domains (Lochy, Zoppoth, Domahs, & Delazer, 2003). Accordingly, it should be possible to account for the cognitive mechanisms of writing letters and numbers within one integrated theoretical approach. This approach should integrate domain-specific central as well as domain-general peripheral processing. However, such a theoretical model has not yet been proposed.

An integrated model for alphabetic and numerical scripts would provide insights into associated and dissociated processes in both domains: on the one hand, domain-specific differences (i.e., dissociations between language and number processing) may be addressed both in children (e.g., Landerl, Fussenegger, Moll, & Willburger, 2009) and in neuropsychological patients (e.g., Delazer & Lochy, 2002; Anderson, Damasio, & Damasio, 1990; Starrfelt, 2007). On the other hand, associated domain-general writing processes in language and number writing (e.g., Artemenko et al., 2018; for evidence of joint neural circuits involved in copying letters and numbers) as well as writing deficits affecting both scripts may be explored. On a more general level, such an integrated model may account for results in previous studies showing that poor literary skills accompany poor arithmetical performance in children (e.g., De Smedt, Taylor, Archibald, & Ansari, 2010, Moll, Fussenegger, Willburger, & Landerl, 2009, Peters & De Smedt, 2018, for a review including neuro-functional evidence). This model may also explain impaired letter and number writing in patients (Delazer & Denes, 1998; Granà, Girelli, Gattinoni, & Semenza, 2001).

The separate approaches proposed for writing letters and numbers in each domain are presented below as they serve as a theoretical framework for the development of my novel integrated writing model. To this end, the cognitive study of and relevant theoretical approaches for writing in both domains will be discussed briefly in the Chapters 3 and 4. Prior to this, terminology of writing as I use it in my thesis will be outlined.

2 Defining the scope of writing

Writing is a broad concept that comprises various processing levels and comprises several scripts. Hence, it is important to define the scope of writing as I use it in the present thesis.

In principle, it is a goal of theoretical approaches to be universally applicable for all conceivable contexts in which a function may occur. However, this goal can probably only be achieved applying a step-wise-procedure. In stepwise procedures cognitive functions or requirements for cognitive processing are initially defined for a circumscribed context or task (e.g., writing at the word level). This context can then be extended to higher cognitive tasks and contexts (e.g., writing at the text level). In this vein, I investigated handwriting and typing in my thesis by means of two writing tasks, namely in a) a writing-to-dictation task and b) a copying task[1]. These two types of tasks are frequently used in learning contexts in school (Berninger, Whitaker, Feng, Swanson, & Abbott, 1996; Deno, 2003), but also in the diagnosis and intervention of acquired writing difficulties (Kay, Lesser, & Coltheart, 1996).

[1] For convenience, the bullet character a) refers to writing to dictation and the bullet character b) refers to the copying task. The bullet characters introduced below i) and ii) differentiate between i) lexical and ii) sub-lexical processing within the tasks a) and b).

In writing, I focus on the production of letters and single digits respectively, which are assembled into words and multi-digit numbers at a central processing level. At this central processing level, orthographic knowledge (i.e., taking spelling rules into account) and number transcoding is investigated. As regards peripheral processing, different output modalities (i.e., different styles of handwriting as well as computer typing) are examined. Higher level writing skills such as the composition of sentences and coherent text generation are beyond the scope of this thesis.

3 Writing letters and words

To begin with the language domain and the theoretical models for writing letters and words as well as the model shortcomings which motivated the research objectives of the present thesis:

The cognitive study of written language in general and of (hand)writing in particular has long sought to identify the cognitive mechanisms of writing and to describe them in various models of written language production (e.g., Ellis, 1982, 1988; Graham & Weintraub, 1996; Margolin, 1984; Morton, 1980; Van Galen, 1991). These models have been predominantly elaborated on both typical and erroneous writing (i.e., slip-of-the-tongue or slip-of-the-pen phenomena[2] in reading or writing) as compared to impaired writing in brain-injured patients in order "to develop the theory of normal writing processes" (Ellis, 1988, p.100). Given such a framework, both developmental (i.e., dysgraphia) as well as acquired writing deficits (i.e., agraphia) can be related to inadequate functioning components or disrupted connections between certain components of the normal writing system (McCloskey & Rapp, 2017). Malfunctioning components in turn should be reflected by inadequate functioning components or disrupted connections on a neural level (e.g., Beeson & Rapcsak, 2015, Rapcsak, Beeson, & Hillis, 2002). Based on behavioral and neural evidence, theoretical writing models proved to be promising for the development of targeted and specific diagnostics and intervention.

So far, theoretical analysis of written language has been dominated by the dual-route-model approach (i.e., Ellis & Young, 2013, for a review; but see Dell, 1988; Dell, Burger, & Svec, 1997; Wilshire, 2008 for further approaches). The dual-route model approach postulates serial/sequential language processing via two independent routes connecting autonomous processing components (i.e., modules): a specific linguistic task (e.g., retrieval of a word meaning)

[2] Slip-of-the-pen phenomena, for instance, concern the writing of homophones (i.e., words that have the same pronunciation but different meanings; for English: *'their'* instead of *'there'*, in German: *'viel'* instead of '*fiel*', which then might be spelled incorrectly.

is executed autonomously by a specifically assigned processing module (e.g., semantic system) via the requisite route(s) (i.e., lexical-semantic processing). Such serial models are often graphically displayed by box-and-arrow graphs, in which a box represents a processing module, and an arrow represents the route in the direction of processing. Crucially, modules and routes are assumed to be affected and treated selectively (Ellis & Young, 2013 for a review). The explicit association of an impairment and its probable cause using serial writing models in language assessments provides precise implications for the intervention of this impairment. This is probably one reason why serial models have prevailed in linguistics in spite of all criticism (cf. Dell et al., 1997; Wilshire, 2008).

3.1 The cognitive processes of (hand)writing

The probably most influential serial model of written language originated from Ellis (1982, 1988). The model was proposed after Morton (1980) introduced his logogen model. It has been commonly used in the literature as a global model to delineate processes involved in handwriting on the basis of single words. In his model, Ellis described writing as a complex act involving linguistic (i.e., central) and motor (i.e., peripheral) components, acquired during childhood and automatically retrieved in literate adults. These central and peripheral components have been recently identified to rely on different neural substrates (Beeson & Rapcsak, 2015; Purcell, Turkeltaub, Eden, & Rapp, 2011; Rapcsak et al., 2002, for meta-anaylses) in a large cerebral network of handwriting (Planton, Jucla, Roux, & Démonet, 2013, for a meta-analysis). For the identification of neural substrates, various writing tasks have been examined, such as writing-to-dictation, copying, and written object naming (e.g., Bonin, Méot, Lagarrigue, & Roux, 2015; Planton et al., 2013; Purcell, Turkeltaub et al., 2011). These tasks allow to investigate different input specific perceptual and cognitive mechanisms: a) writing-to-dictation involves initial auditory-verbal (i.e., phonological) processing, while b) copying involves reading and thus initial visual-graphemic processing[3], and c) written object naming requires purely visual perception of an object.

Apart from such input differences, Ellis (1982) proposed analogue and joint writing components at central and peripheral processing for writing which are required to solve these three different tasks. In the following, only the first two tasks a) and b) will be examined. Figure 3.1 illustrates language processing in writing-to-dictation and copying of single words as proposed by Ellis (1982,1988). In the figure, model components are given for speech (articulatory output) and writing. The required processing stages for speech are grayed out as the focus is on written output.

[3] In language research, graphemes – as meta letters – unite the knowledge of all concrete letter forms (i.e., upper and lower case). For example, there is a grapheme [H] for the letters 'H' and 'h'. During copying, visual-graphemic processing requires the retrieval of graphemes to encode the visually presented word.

In the Ellis' model (1982), two different processing levels are distinguished: at a central processing level, cognitive processes for correct spelling (i.e., orthographic rules) are executed. At a peripheral processing level, graphomotor execution is carried out. In Figure 3.1, I indicated the different processing levels by adding curly braces to the Ellis' model (1982). Both processing levels are described briefly in the following.

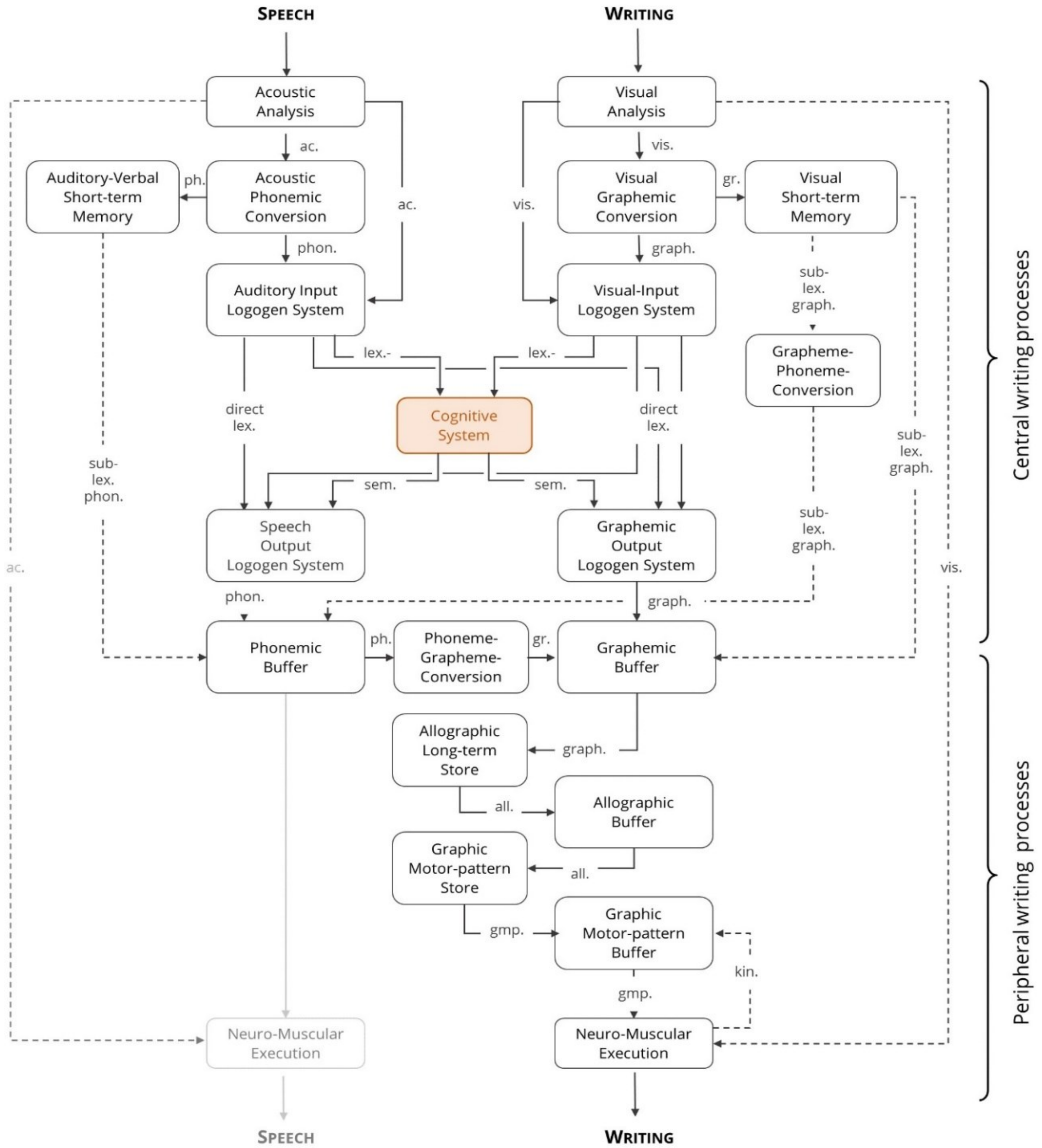


Fig. 3.1: Model of language functions adapted from Ellis (1982, p. 140): Model components are given for speech and writing according to Ellis (1982). Modules required for articulatory output are grayed out. The module necessary for semantic processing is indicated in orange. Several processing routes are distinguished: lexical processing (i.e., direct-lexical and lexical semantic) routes are given by solid arrows. Sub-lexical processing routes are given by dashed arrows. ac. = acoustic code, all. = allographic code, direct lex. = direct lexical code, graph./gr. = graphemic code, gmp. = graphic-motor pattern, kin. = kinesthetic, lex. - sem. = lexical-semantic code, sub.-lex.-graph. = sub-lexical graphemic code, vis. = visual code. Curly braces indicate serial processing stages.

3.1.1 Central components of writing according to Ellis (1982)

In the model by Ellis (1982), central components of writing (upper part of Fig. 3.1) are supposed to include independent i) lexical and ii) sub-lexical processing routes. i) In lexical processing, an input and an output logogen (i.e., lexicon) system are proposed. The input lexicon (i.e., auditory/visual-input logogen system) is particularly important for language comprehension and reading. The graphemic output lexicon (i.e., graphemic output logogen system), in turn, is specifically involved in writing processes. The graphemic output lexicon is proposed to embrace lexical information about all acquired words, for instance, information about the grammatical word class (i.e., a noun that has to be written in capital letters), grammatical gender (e.g., female), morphology, etc. The graphemic output lexicon also includes information about the orthographic irregularities of a word (e.g., in the phoneme-grapheme assignment when a phonemic *'[t]'* is realized as graphemic *'d'*, in the final obstruent in a syllable or a word, for instance, the German word *'Hand'* is pronounced as *'['hant]'*). Only by accessing these attributes stored in the graphemic output lexicon, an irregular word can be spelled correctly.

To stick with the example: a) in writing-to-dictation of the auditorily presented word [*'hant*], writing (i.e., starting with speech input, on the left side in Fig. 3.1) can be direct-lexically or indirectly processed via the lexical-semantic route (both indicated by solid arrows in Fig. 3.1). In direct-lexical processing (abbreviated as direct-lex. in Fig. 3.1), the word can be correctly written as *'Hand'* by direct processing routes between the orthographic input and the graphemic output lexicon. In contrast, lexical-semantic processing (abbreviated as lex-sem. in Fig. 3.1) involves a meaningful spelling of the word '*['hant]'* by resorting to the abstract semantic concept (i.e., body part, extremity, five fingers, etc.) Nevertheless, lexical access is necessary for the orthographically correct spelling as '*Hand'*. b) For copying of the visually presented word '*Hand'* (i.e., starting with written input, on the right side in Fig. 3.1), Ellis' (1982) model also suggests identical lexical and lexical-semantic processing routes.

ii) Sub-lexical processing (indicated by dashed arrows in Figure 3.1) proceeds at the surface of a verbal string neither activating a lexical entry nor semantic information. It is merely required in the (re-) production of unknown word material or writing non-words (Ellis, 1982; but see Bonin et al., 2015, for the role of sub-lexical processing in writing-to-dictation). Thereby the phonetic structure of a word is exactly assigned to the associated graphemes by phoneme-grapheme assignment. Typically, grapheme-phoneme assignment is highly flexible "This is easily demonstrated by devising a new word (i.e., a non-word) and asking someone to invent a plausible spelling for it. Alternatively yoo kann asque peepul tu kreeate olturnativ butt plorzibul spelingz forr fammiliyer wurdz." (Beattie & Ellis, 2017, p. 165).

Referring to that example, a) writing-to-dictation of the auditorily presented German word '['*hant*]" via the sub-lexical processing route (sub-lexical graphemic code [=sub.-lex. graph.] in Fig. 3.1) would result in a plausible but incorrect spelling (i.e., '*hant';* Tainturier & Rapp, 2001, for spelling consistent and inconsistent words). b) For copying the visual string '*Hand'* the Ellis' (1982) model suggests two sub-lexical processing routes: first, a visual string can be copied by pure visual depiction of the letter forms (i.e., visual [=vis.] code in Fig. 3.1). Second, graphemes are again sub-lexically transferred by grapheme-phoneme conversion. This processing route is only required when the visually encoded letter string is internally verbalized (i.e., read) and then noted.

The information processed by the lexical and sub-lexical spelling route of the respective task (i.e., writing-to-dictation or copying) finally converges at the graphemic buffer (Ellis, 1982). The graphemic buffer has been labeled more recently as 'orthographic working memory' (McCloskey & Rapp, 2017; Purcell, Turkeltaub et al., 2011) to emphasize its functional role in writing. The graphemic buffer is considered to be the final central processing module that serves as the interface between central and peripheral writing components (Beeson & Rapcsak, 2015; Rapcsak et al., 2002).

3.1.2 Peripheral components of writing according to Ellis (1988)

In the model by Ellis (1988), peripheral components of writing (depicted in the lower part of Fig.3.1) represent – in analogy to articulation in speech – a complex process of the production of letters. This process involves multiple stages of cognitive processing which are regarded as essential steps toward handwriting. First, abstract representations of letter identities also referred to as graphemes have to be activated in the graphemic buffer. Graphemes are abstract meta letters without any individual specification (Teulings, 1996). In the example above, the graphemes {[H] [A] [N] [D]} need to be kept in the graphemic buffer (i.e., working memory) in this particular order to write the auditorily presented word *'[hant]'* correctly. Second, graphemes are converted into specific letter shapes or allographs (but see Menichelli, Rapp, & Semenza, 2008; Rapp & Caramazza, 1997, Van Galen, 1991, for another approach in which graphemes are assumed to be directly converted into graphic motor plans[4]). Allographs depict various concrete representations

[4] The term 'graphic motor plans' was introduced by Ellis (1982,1988) and is used exchangeably with the term 'graphomotor patterns' in this thesis.

of the letter's shape (e.g., case: UPPER and lower, style: print or *handwriting*). These allographs are stored in the allographic buffer in which transcoding processes[5] are performed between allographs. At the allographic level, only the initial allograph for graphemes {[H] [A] [N] [D]} in the example *'[hant]'* has to be activated in upper case {[H] [a] [n] [d]}. Third, graphic motor patterns for the selected allographs are specified (e.g., order, position, direction) and kept until execution. Graphic motor patterns contain information about the number of strokes within an allograph or the direction of writing (e.g., for the letter 'H' two parallel vertical strokes need to be combined by a short horizontal stroke in printed script). Finally, information encoded in graphic motor patterns is translated into motor commands for neuro-muscular execution of writing movements (i.e., *Hand*).

In contrast to central components, peripheral aspects of written language processing are those involved merely in writing execution. They differ depending on the output mode (e.g., handwriting, typing, touch-typing, etc.). However, in the initial version of his model, Ellis (1982) addressed only handwriting. Additional output domains (i.e., typing and spelling with letter blocks A H D N) which have to be assembled correctly, were introduced afterwards by Margolin (1984). Margolin suggested that the graphemes stored in the graphemic buffer[6] are transferred to a shared physical letter code (i.e., allographic level in Ellis' 1982 model) for handwriting and typing (but see Purcell, Turkeltaub et al., 2011, for another approach). "The physical letter code [specifies] which physical forms are acceptable versions of the [grapheme]" (Margolin, 1984, p. 469). For handwriting, and in accordance with Ellis (1982), the physical letter code is translated into motor commands for neuro-muscular execution of typing movements by activating graphic motor patterns. Analogous to graphic motor patterns, Margolin (1984) assumed the existence of typing motor patterns. Typing motor patterns entail a number of associations between letters and finger movements in order to type on a standard keyboard. By including typing in his model, as suggested by Margolin (1984), Ellis (1988) proposed a model framework in which cognitive mechanisms of writing can be explained using different writing modes (i.e., handwriting vs. typing), making the model still applicable in times of using recent digital writing technologies.

[5] Please note that, in the language domain, transcoding between different scripts proceeds merely at the peripheral processing level and is referred to the switching between different cases and styles in written output production. In the number domain, number transcoding proceeds at the central processing level at the transition to peripheral writing processes.

[6] Please note that Margolin (1984) used the term orthographic buffer in his model.

3.2 The influence of digital media on writing

The development of digital technology has been increasingly changing our everyday writing habits. The impact of these changes on the underlying cognitive mechanisms is still unknown (Pinet et al., 2016), as are the consequences for more general cognitive mechanisms (Longcamp et al., 2008; Longcamp, Zerbato-Poudou, & Velay, 2005) and writing processes (Goldberg, Russell & Cook, 2003; Mangen & Velay, 2010; Wollscheid, Sjaastad, & Tømte, 2016, for an overview). Evidence has been provided that handwriting facilitated letter recognition more than typing (Longcamp et al., 2008, 2005). These findings suggest cognitive processing differences due to the writing mode. Thus, the influence of the actual writing mode, the termed *mode effect,* has been the subject of many investigations (for typically developed written language: e.g., Connelly, Gee, & Walsh, 2007; Gaskill & Marshall, 2007; Goldberg, Russell, & Cook, 2003; Russell & Plati, 2002; Wollscheid, Sjaastad, & Tømte, 2016; for impaired written language: e.g., Berninger, Abbott, Augsburger, & Garcia, 2009). The vast majority of these studies investigated dynamic writing measures such as writing latency on different language levels (i.e., grapheme, word and/or sentence level; e.g., Berninger et al., 2009). Writing latency is assumed to reflect not only peripheral processes but also central processes of writing (Delattre, Bonin, & Barry, 2006). Regarding central processing levels, writing latency seems to be sensitive to linguistic effects either on semantic, lexical or sub-lexical levels (Bonin & Meot, 2002; Donders, 1868). For a more detailed description on how linguistic effects disclose central cognitive writing processes please find the digression in the next paragraph.

However, only a few studies have reported on writing performance in terms of spelling accuracy, writing errors or revision behavior (e.g., Goldberg et al., 2003; Wollscheid et al., 2016). Results were found to differ considerably depending on the age group studied and the presence of written language deficits. Younger children wrote more slowly on the computer keyboard than by hand (i.e., Berninger, Abbott, Augsburger, & Garcia, 2009; Connelly, Gee, & Walsh, 2007; Wollscheid, Sjaastad, & Tømte, 2016). With growing age and probably increasing computer experience (Russell, 1999; Russell & Plati, 2002) this effect diminished. Older children and adults, instead, usually wrote faster but tended to revise more often during typing (Goldberg et al., 2003, for a meta-analysis). However, children with impaired written language processing (i.e., dyslexia or dysgraphia) were observed to type slower and less accurately as compared to their peers. Interestingly, they corrected their own writing as much as typically developing children (Morken & Helland, 2013). On a more general level, recent evidence has identified some applications of digital writing media (e.g., word processing, speech recognition) as beneficial for children with impaired written language production (MacArthur, 2009). In view of the (further) development of a theoretical model of handwriting and typing, it is imperative to understand how different writing media impact on cognitive writing processes.

So far, investigations of computer keyboard typing from a theoretical perspective similar to handwriting are scarce (Pinet et al., 2016). The few studies investigating this issue, mainly focused on peripheral components (i.e., the execution of a keystroke) rather than central components (i.e., selection of required graphemes) to illustrate typing processes in serial model approaches (Rumelhart & Norman, 1982; but see Logan & Crump, 2011, for a hierarchical dual-route model approach). During typing, very high typing rates (i.e., up to ten keystrokes per second) have been observed. These typing rates argue for a high-level of automation and parallel preparation of successive key presses (Teulings, 1996). To achieve such high automation, typing is assumed to develop with practice from a sequential into a parallel strategy by several fingers moving simultaneously (Gentner, 1983).

The theoretical discussion on typing is impeded not only by the open question about the nature of processing, but also by the fact that there is currently no consensus on the definition of the smallest unit (i.e., graphomotor plan) to be executed during typing. While some authors have assumed bigrams or trigrams or even words (i.e., Nottbusch, 2008, for further discussion on this issue), others define “a key stroke as [the smallest] unit in typewriting, because it is so discrete as compared to the fluent […] handwriting movement” (Teulings, 1996, p. 549). The latter definition is employed in the present thesis in order to provide comparable writing units (i.e., a letter in handwriting equals a keystroke in typing) in written language production. Thus, both healthy and impaired cognitive mechanisms of writing and typing can be illustrated within the same theoretical model and analyzed on identical linguistic levels.

In summary, from the investigation of the theoretical model of (written) language functions of Ellis (1982), two relevant points emerged: on the one hand, it has been shown that writing cannot be understood independently of its connection to linguistic (i.e., central) aspects. On the other hand, these aspects have rarely received attention in previous theoretical approaches to portray cognitive mechanisms in keyboard typing. As writing with digital media has been shown to differ significantly from handwriting (Mangen & Velay, 2010), it is vital to consider different output modalities (i.e., handwriting and typing) as well as central and peripheral aspects in a comprehensive theory of typical writing. Such a theory is indispensable for the development of precise diagnostic and individual interventions, whether for children or adults.

However, our cultural achievements do not only embrace mastery of alphabetic scripts, but also the notation of digits and multi-digit numbers. Acquisition of numerical scripts is regarded a milestone in mathematics teaching for children (Moura et al., 2013). Therefore, understanding the cognitive processes in writing digits and multi-digit numbers is of particular interest, as is writing letters and words.

Before the theoretical models for writing digits and multi-digit numbers are presented, I will briefly introduce how to 'read' serial writing models. This digression is intended to provide deeper insights into the theoretical writing model and to demonstrate how atypically writing processes can be assessed by means of psycholinguistic effects.

3.3 Digression: How to 'read' a serial writing model

In this digression, differences between i) lexical and ii) sub-lexical processes are outlined and linked to various psycholinguistic effects. Psycholinguistic effects are determined mainly on the basis of reaction time measurements which originated from the seminal approach by Donders (1868). Donders was the first to assume that cognitive processes can be measured over time. In psycholinguistics, such effects are frequently used to indicate cognitive processing differences, for instance, with regard to the lexicality of a word, word length, its frequency, or its imagery (these effects are discussed in depth below). Psycholinguistic effects are presumed to be specifically associated with specific processing routes and modules. Thereby, an explanation consistent with the model is provided (Teulings, 1996). Table 1 summarizes which psycholinguistic effects are associated with the respective processing routes and modules according to the Ellis model (1982), The table also gives an overview of the characteristics of the writing errors in the different processing stages.

The theoretical analysis of writing and spelling errors in different output domains and at different levels of processing is one of the major accomplishments of the model by Ellis (1982) and its extensions (Ellis, 1988; Ellis & Young, 2013; Margolin, 1984). These errors can be used to identify writing strategies and are frequently analyzed in psycholinguistic assessments in order to evaluate cognitive mechanisms of writing (Kay et al., 1996). In general, the model assumes that the earlier a disorder occurs in the cognitive process, the more severe is its impact on subsequent processing steps (Ellis, 1982). For instance, sequence of processing steps a) in a writing-to-dictation task, writing typically begins by hearing the target word. b) In copying, the target word needs to be perceived visually first. These input processes at the entry of the central processing level (upper part of Fig. 3.1) serve as the basis for correct spelling (Purcell, Turkeltaub et al., 2011).

Table 1: Characteristics of writing errors and psycholinguistic effects in handwriting and typing associated with the different processing levels proposed by Ellis' (1982,1988)

Components	**Characteristics of writing errors**	**Examples**	**Psycholinguistic effects**
Central writing components			
Routes			
Sub-lexical	• preserved writing of regular words and non-words • regularization errors of irregular words	'hant' - hand 'serkit' - circuit	**Regularity effect** Beauvoir & Dérouesné (1981)
Lexical-semantic	• preserved writing of regular and irregular words • impaired writing of unknown words or non-words • frequent words better than less frequent words	'cough' - cough 'hand' - hand	**Lexicality effect** **Frequency effects** Shallice (1981)
Modules			
Semantic system	• semantic errors • writing of nouns > verbs > functional words	'mother' - father sea > see > she	**Concreteness effects** Bub & Kertesz (1982)
Phon. Graph. Conversion	• impaired writing of regular words and non-words • impaired assignment of sounds to letters (vice versa)	'bag' - back 'wan' - van	**Regularity effect** Beauvoir & Dérouesné (1981)
Graphemic Buffer	• longer words more error prone • difficulties with homophones • substitutions, omissions, additions, transitions • anticipations and perseverations of a grapheme which occurs later in the word order • confusion between identical graphemes but different allographs	 for - four - fore 'Gog…' - Cognitive 'english' -English	**Length effects** Sage & Ellis (2004) Hatfield & Patterson (1983) **Serial order effects** Caramazza et al. (1987) Teulings (1996)
	• serial order effects: substitutions, omissions, additions, transitions • spelling is captured by another but similar word, e.g., due to anticipations	'boold' - blood efficiency' - efficient	**Serial order effects** Grudin (1983) **Capture effects** Rumelhart & Norman (1982)
Peripheral writing components			
Allographic level	• omissions or haplographies of two identical allographs in a word • case and style substitutions of well-formed letters	'satified' - satisfied 'lamp' - lamb 'padoga' - pagoda	Teulings (1996) Ellis (1982) De Bastiani & Barry, (1989)
	• selection of the wrong, but related movement	'thses' - these	Margolin (1984)
Graphic motor plans	• distorted arrangements of strokes within a target letter		Margolin & Goodman-Schulman (1992)
	• transposition of keystrokes (both hands are in a different phase of typing) • doubling of key strokes	'wnet odnw' - went down 'scrren' – screen	Teulings (1996) Teulings (1996)

Note: Central and peripheral writing components are depicted separately for routes and modules. Characteristics of writing errors are given for each processing step. At the central processing level, psycholinguistic effects are assigned to the occurrence at the particular processing stage. From the transition of central to the peripheral processing level (from the graphemic buffer), modules are separated for writing and typing: handwriting is depicted above the dashed line and typing is depicted below.

A primary impairment to either of the input processes is excluded in the subsequent theoretical consideration. This means that it is assumed that the auditory-verbal analysis of the target word during writing-to-dictation and the visual-graphemic analysis of the target word during transcription are unaffected. Access to the respective input lexicons (i.e., auditory and visual input logogen system according to Ellis, 1982, see Fig. 3.1) is also considered to be unimpaired.

In the following, the impact of impairments on central processing routes will be described. At the central processing level, i) lexical (i.e., direct-lexical and lexical-semantic routes) and ii) sub-lexical processing routes are distinguished. These routes can be disrupted simultaneously or separately. The breakdown of both processing routes and modules respectively leads to a considerable impairment of written language (Ellis, 1982).

i) Lexical processing routes are predominantly used, even if not necessarily undisturbed, when the sub-lexical route is (selectively) interrupted (e.g., due to an impairment of phoneme-grapheme conversion, Shallice, 1981). In direct-lexical processing, information is processed directly from the input to the graphemic output lexicon (see Fig. 3.1). In this case, writing of regular and irregular words is preserved because access to the graphemic output lexicon is provided, but not writing of unknown words or non-words. The latter dissociation in performance is referred to as lexicality effect (Shallice, 1981). When the graphemic output lexicon can be accessed, but is impaired itself, specific frequency effects (i.e., better spelling performance for more frequently occurring words, Warrington & Shallice, 1979) may occur. The occurrence of frequency effects can be explained by the development of the graphemic output lexicon: both high frequent and early acquired words show processing advantages as compared to low frequent and later acquired words (Domahs, 2016). This effect applies in particular to the latter words, for which rather weak representations of the word exist due to the age of acquisition or to its lexical difficulty. Frequency effects can exclusively indicate a malfunctioning lexicon (Beauvois & Dérouesné, 1981; Ellis, 1982), but cannot explain deficits in semantic processing. Semantic processing deficits occur when access to the semantic system or the semantic system (depicted in orange in Fig. 3.1) itself is impaired in lexical-semantic processing. Deficits in lexical-semantic processing can be indicated, for instance, by the occurrence of concreteness effects (i.e., better spelling performance for concrete as compared to abstract and function words, Bub & Kertesz, 1982) or semantic writing errors (e.g., ‘father’ instead of ‘mother’; Alexander, Friedman, Loverso, & Fischer, 1992; Beeson & Rapcsak, 2015; Rapcsak et al., 2002).

ii) Sub-lexical writing is utilized when lexical processing routes are (selectively) impaired. Sub-lexical processing is characterized by the so called regularity effect (e.g., for German: Wasser - [ˈvasər], Fisch - [fɪʃ]) and irregular words (e.g., for German: 'Vase' - [ˈvaːzə], 'Vogel' - [ˈfoːgəl]) which is reflected in a regularity effect (i.e., better spelling performance for words whose phoneme-grapheme assignment is highly predictable from their sound (Beauvois & Dérouesné, 1981; McCloskey & Rapp, 2017). In this case, the phonological equivalent of the word is used to write irregular words for which lexical access is typically required (i.e., [ˈvaːzə], - ‘Wase’ or [ˈfoːgəl] –

‘Fogel’). Regularity effects were reported from conflicting (alternative) spellings generated by lexical and sub-lexical processes (Beauvois & Dérouesné, 1981; Tainturier & Rapp, 2001). Resolving this conflict was found to increase cognitive load and slow down writing times (Kandel & Perret, 2015).

At later central processing stages, spelling errors are also caused by deficits in the graphemic buffer. The graphemic buffer is supposed to maintain the respective order of graphemes generated by central processing components which are then converted into motor commands by the peripheral components of writing (Beeson & Rapcsak, 2015). The occurrence of length effects (i.e., decreased spelling performance with increasing word length, e.g., Sage & Ellis, 2004), serial order effects (i.e., letter substitutions, omissions, additions, transitions etc., for handwriting Caramazza, Miceli, Villa, & Romani, 1987; for typing: Grudin, 1983; Rumelhart & Norman, 1982) as well as capture effects (i.e., spelling is captured by another but similar word, Rumelhart & Norman, 1982) during writing is assigned to a malfunctioning graphemic buffer. These effects can be observed during handwriting and typing, as shown in Table 1 because the buffer applies to both writing modes (Margolin, 1984).

At the peripheral stages of writing motor execution, deficits at the allographic level, where information of shape, case, and font is assigned, are assumed to cause mere peripheral writing impairments. “The hallmark characteristic of damage to the allographic stage is the ability of an individual to spell in some modalities, fonts or case but not in others” (Menichelli et al., 2008, p. 861). In particular, case or style substitutions (De Bastiani & Barry, 1989) or selective case writing (Lochy et al., 2003, for an overview) have been described. In contrast to handwriting, no allographic level has been proposed for typing. Subsequent to the allographic level, graphomotor programs for handwriting, (i.e., number of strokes and their serial order, orientation, size) and typing (i.e., finger movements and their serial order) of the target letter have to be selected. In handwriting, a deficit at the selection stage of graphic motor programs is assumed to yield distorted arrangements of strokes within a target letter (Margolin & Goodman-Schulman, 1992). In typing, transposition errors (i.e., when keys are confused because both hands are in different positions in the word) or doubling of keystrokes may indicate a malfunctioning selection of graphomotor programs (Teulings, 1996, see also Table 1). Finally, specific neuro-muscular commands (i.e., force and size of production; Van Galen, 1991 for handwriting) are activated to terminate the (incorrect) typing process of the target word.

To sum up, this brief digression demonstrated how theoretical writing models may help to understand typically and atypically writing processes from a linguistic research perspective. How writing of digits and numbers is investigated from a theoretical perspective in numerical research is outlined below, as are the limitations of this perspective that underline the need for a more comprehensive writing model for writing in the number domain.

4 Writing digits and numbers - the nature of number transcoding

Writing digits and numbers differs significantly from writing letters and words in both script and central writing processes. Unlike letters (like A), digits can be written in two different formats in the Indo-Arabic decimal system: in an Arabic digit number format (like *2*) or a verbal number word format (like '*two*' or *[tu:])*. Switching between these number formats (i.e., mapping numbers from one format to another) is referred to as number transcoding (Deloche & Seron, 1987). Number transcoding is encountered in our daily activities, for instance, registering dates, telephone numbers or making mental calculations. Moreover, it is one of the most basic numerical abilities acquired in childhood as associations between the Arabic digit and the verbal number word format which establish during development (Zuber, Pixner, Moeller, & Nuerk, 2009).

The Arabic digit number format follows a base-10 place-value structure. Its lexicon contains a small set of ten digits (from 0–9). In the composition of multi-digit numbers, these ten digits are syntactically composed employing 10 as the base by a positional notation following the place-value principle (Zhang & Norman, 1995). According to this principle, the value of a digit is determined by its position in the sequence of digits with increasing values of powers of ten from right to the left. In this scheme, the number 2019 represents the sum of $(2 \times 10^3) + (0 \times 10^2) + (1 \times 10^1) + (9 \times 10^0)$. The verbal number word format was suggested to require similar processing pathways as words, for instance, for naming or accessing conceptual knowledge of numbers (Damian, 2004). It was further assumed to be linguistically structured comprising lexical and syntactical constraints (Fayol & Seron, 2005): single number words (like '*two', 'ten' and 'thousand')* are stored in an abstract number lexicon and arranged following language specific syntactic rules (like inversion of number words for teens in English ['*nine-teen'* instead of '*ten-nine*'] or for two-digit number words in German [like '*ein-und-zwanzig*' instead of '*zwanzig-und-eins*']). These rules are assumed to rely on two main principles, either an additive (e.g., *nine-teen* means *ten* plus *nine)* or multiplicative (e.g., '*two thousand'* means '*two times thousand*') composition of individual number words. Fayol and Seron (2005) pointed out that due to the lexical-syntactic peculiarities of each language, the base-10 structure of the verbal number word system is more or less transparent. Hence, various languages are differently prone to transcoding errors.

Transcoding errors were found to occur both during development in children (e.g., Nuerk, Moeller, Klein, Willmes, & Fischer, 2011, for a review; Zuber et al., 2009) and as a result of neuro-cognitive impairments mostly in adults (e.g., Cipolotti & Butterworth, 1995; Delazer & Bartha, 2001; Deloche & Seron, 1982; Nuerk et al., 2011, for a review). Due to the lexical-syntactic structure of the number formats, lexical (e.g., 'sixty-*three'* instead of 'sixty-*five'*) or syntactic transcoding errors (e.g., *'200019*' instead of '*2019'* or '*ten-nine'* instead of '*nine-teen'*) have been distinguished (Fayol & Seron, 2005). In their review on multi-digit number processing, Nuerk and colleagues (2011) subdivided syntactic transcoding errors into inversion errors ('*ten-nine'* instead of '*nine-teen'*), additive composition errors (e.g., *'200019'* instead of *'2019'*), and multiplicative composition errors (e.g., '*four-hundred'* is written as '*4100'*). Thereby, they made a distinction between the two basic principles of additive and multiplicative composition of multi-digit numbers. The authors further drew attention to the theoretical importance of investigating number transcoding (i.e., transferring numbers from one into another number format), for the ongoing controversy on semantic or asemantic processing routes of (multi-digit) number processing. This controversy is one of the most significant in numerical cognition. It is pertinent to the writing of numbers as there is still considerable ambiguity as to whether number transcoding always requires (semantic) processing or whether additional sub-semantic processing routes can be assumed in analogy to the dual-route model approach in linguistics. Crucially, the challenge for any transcoding model is to comprehensively explain how the decimal structure of the multi-digit number and the verbal number word format are integrated (Dotan & Friedmann, 2018b).

In recent years, various cognitive models of number transcoding, including semantic (McCloskey, 1992) and asemantic processing models (Barrouillet et al., 2004; Deloche & Seron, 1987) as well as models assuming multiple processing routes (Cipolotti & Butterworth, 1995) have been proposed. These models were mainly based on developmental data or findings in neuropsychological patients (Barrouillet et al., 2004). The most influential models will be briefly presented below. Analogous to the cognitive study of (hand)writing, first central processing levels will be outlined before peripheral writing components in (written) number transcoding are examined.

4.1 Central components of number transcoding

McCloskey (1992) was among the first to propose a semantic model for number transcoding based on several single-case studies (the Abstract Code Model, McCloskey, 1992). This model was a one-route model assuming an obligatory abstract semantic representation about the numerical quantity. In his Abstract Code Model, this semantic quantity representation was assumed as the central system which is responsible for communication between number processing (i.e., number comprehension and production) and calculation subsystems. However, model assumptions have not escaped criticism. Several researchers expressed doubts about the necessity of an internal, abstract representation as a central system between number processing (i.e., input and output systems) and calculations (Dehaene, 1992). Critics advocated either for multi-route models of number processing (Cipolotti & Butterworth, 1995; Verguts & Fias, 2006, for a connectionist approach) or for mere asemantic transcoding (Barrouillet et al., 2004; Deloche & Seron, 1987) "that is, direct translation between [A]rabic and verbal notations without going through an intermediate semantic representation" (Dehaene, 1992, p. 29). These researchers, in turn, presented further models investigating number processing in general and writing of numbers in particular.

The currently most influential model of numerical cognition is the Triple Code Model (TCM, Dehaene, 1992) with its neural implementation (Dehaene & Cohen, 1995, 1997), and extensions (Dehaene et al., 2003; Nieder & Dehaene, 2009; Amalric & Dehaene, 2018; Dehaene, 2009). It was the first model to integrate both neuro-functional and behavioral aspects. In the TCM, an integrative network comprising three modules or *codes* connected via bi-directional transcoding routes is proposed. Figure 3.2. shows the bi-directional structure of the processing codes of the TCM in an early version of the model (Dehaene, 1992; Dehaene & Cohen, 1995).

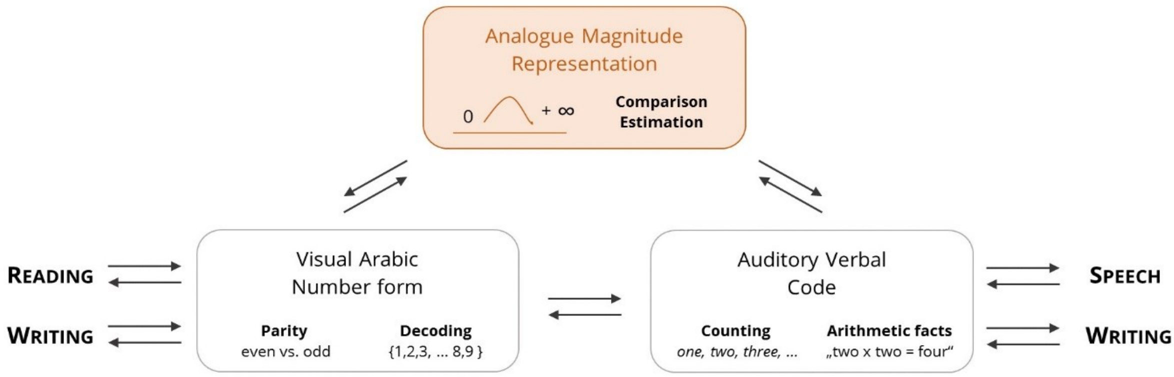


Fig. 3.2: Triple Code Model adapted from Dehaene (1992, p. 31, Dehaene and Cohen, 1995, p. 131): Semantic retrieval of analogue number magnitude information is depicted in orange.

First, an auditory verbal code represents numbers in their linguistic form. This representation is used for counting, naming, and phonologically coded operations such as the retrieval of arithmetic facts (e.g., multiplication tables). The auditory-verbal code is localized in left-hemispheric perisylvian language areas as well as basal ganglia and the thalamus (Dehaene &

Cohen, 1995, 1997). Retrieval of arithmetic facts was later specifically associated with the left angular gyrus (Dehaene et al., 2003, but see Bloechle et al., 2016; Menon, 2016).

Second, a visual Arabic number form is suggested to allow identification of Arabic digits in the ventral path of visual object processing. Its neural correlates were found to be located bilaterally in the occipito-temporal junction. Both, auditory and visual code are presumed to be asemantic.

Third, solely, the analogue magnitude code is proposed to be semantic. It is the only code proposed of the TCM that encodes semantic information about the magnitude information of a number and is often described in Western culture by the metaphor of a mental number line oriented from left to right. Neuroimaging studies have provided evidence that number magnitude information is activated even if it is irrelevant for solving a particular task at hand (e.g., Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003; Klein, Moeller, Nuerk, & Willmes, 2010). Its neural correlate was found to be located bi-hemispherically in the parietal cortex around the horizontal segment of the intraparietal sulcus (hIPS, Deheane et al., 2003). In addition to these codes, contributions of (pre)frontal brain areas are assumed to support numerical processing by more general cognitive processes such as attention, working memory and cognitive control. These processes are required in order to solve more complex mathematical problems (Arsalidou & Taylor, 2011; Imbo, Vandierendonck, & De Rammelaere, 2007).

In a nutshell, the previous models propose an abstract semantic code to be involved in every numerical task. However, a growing body of literature has provided evidence that a mere semantic approach is not sufficient to reflect typically and impaired number processing. In the latter context, clinical symptoms observed in brain-injured patients (e.g., preserved number comprehension and thus intact abstract semantic representations but impaired number transcoding; Cipolotti & Butterworth, 1995; Cohen & Dehaene, 1991) stressed the need for asemantic processing routes.

In their introduction of a connectionist model of number transcoding, Verguts and Fias (2006)[7] pointed to the difference between semantic and asemantic routes. In line with McCloskey (1992), they stated that "a semantic route is one in which a number is decomposed into its base-ten structure" (Verguts & Fias, 2006, p. 265). The asemantic route was suggested to proceed without "a correct (base-ten) interpretation of the input string" (Verguts & Fias, 2006, p. 265). Transcoding is assumed to act directly on the verbal input string. In this vein, Barrouillet and

[7] The present thesis focuses on serial processing models following the dual-route model approach. For this reason, the connectionist transcoding model by Verguts and Fias (2006) is not discussed in greater detail. However, the importance of connectionist models for a (computer-based) modelling of cognitive processes for numerical processing in analogy to linguistic research deserves to be mentioned.

colleagues (2004) proposed an asemantic procedural model for number transcoding from the verbal number word format to the Arabic digit number format (ADAPT: A Developmental, Asemantic, and Procedural Transcoding model, Barrouillet et al., 2004; but see Dotan & Friedmann, 2018a, for transcoding from the visual Arabic digit number format to the verbal number word format in reading). The ADAPT model was grounded on earlier work by Deloche and Seron (1987), which stated that transcoding does not require any (semantic) representation of numerical magnitude information. Figure 3.3 provides a depiction of the ADAPT model.

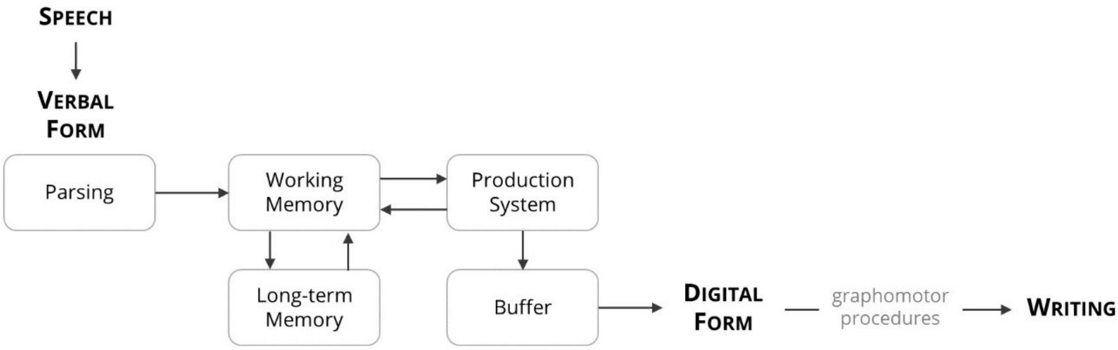


Fig. 3.3: ADAPT Model postulated by Barroulliet et al. (2004). Illustration adapted from Camos et al. (2008, p. 41)

Number transcoding according to the ADAPT model is proposed to proceed as follows, as shown in Figure 3.3: First, the verbal string (e.g., *'zweitausendneunzehn'*) is phonologically encoded (e.g., *'['tsvai]' [ˈtauznt] '[ˈnɔɪ̯ntseːn]'*). Second, a parsing system sequentially segments the verbal chain into lexical units (i.e., *'['tsvai]' – '2', '[ˈtauznt] – 'tausend = _ _ _ _ ', '[ˈnɔɪ̯ntseːn]' – '19'*). The parsed units are then sequentially sent to a working memory system in which they are stored temporarily for further processing. Third, the stored units are subsequently compared with internal lexical representations of numbers stored in long-term memory. In line with the lexicon proposed in the language domain, these internal lexical representations of numbers are assumed to develop during number acquisition in childhood (Moura et al., 2013). They comprise information about the serial order of lexical units. In the ADAPT model lexical entries are proposed to be arranged following different transcoding rules (P1 – P4, Barroulliet et al., 2004). Therefore, a production system is assumed to execute these transcoding rules: P1 rule embraces lexical retrieval of information from long-term memory, P2 and P3 rules are devoted to the number quantity and the arrangement of the syntactic frame (i.e., identification of how many slots are required to arrange the parsed lexical units, e.g., *'2_19'*) For syntactic arrangement, the P2 rule processes the separator *'hundred'*, the P3 rule processes the separator *'thousand'.* These separators were also isolated within the parsing process. P4 rules finally fill the empty slots (if any) with intermediary 0s (e.g., '*2019'*), and trigger the transformation of the abstract Arabic multi-digit number into output processes through graphomotor procedures (Camos, 2008). However, the way in which these processes operate was not specified. The ADAPT model, thus, omits peripheral aspects of writing digits and multi-digit numbers.

Regarding the asemantic nature of their model, Barroulliet and colleagues (2014) mentioned in a footnote that they do not exclude abstract semantic representations that are processed in parallel to asemantic processing. However, semantic processing was not considered to be necessarily involved in number transcoding. Accordingly, an interplay of potential semantic and asemantic processes during number transcoding was not provided. This issue might limit generalizability of the ADAPT model.

Let us consider a simple example: Imagine being in a (neuro-) psychological assessment. The German speaking patient or child to be examined listens to two multi-digit numbers which are auditorily presented (e.g., *'32'* and *'23'*). The task is to transcode the verbal number word format into the Arabic digit number format and to write down the numbers in increasing order (i.e., by numerical magnitude) which requires processing of the numerical quantity. The patient or the child does not succeed and writes again (e.g., *'32'* and *'23'*). What causes the incorrect response? The answer could be threefold: First, it could be an inversion-related syntactic transcoding error probably related to incorrect realization of the inversion of tens and digits in German language. Simultaneously, the error could be related to an incorrect syntactic assembly and storage of the required digits or to limited working memory capacity. Second, it could be a lexical transcoding error due to the incorrect retrieval of the visual format of the required digits. Third, it could be a lexical-semantic transcoding error because the abstract number magnitude representation cannot be activated correctly or is not yet established, for instance, for larger numbers in children. Crucially, the last possibility cannot be covered by the ADAPT model. Further investigations on central processes of number transcoding are required, for instance, a number comparison task to fully understand the origin of the transcoding error. This short example should illustrate that for a precise analysis of number transcoding (errors), in particular, and writing of digits and multi-digit numbers, in general, both semantic and asemantic processes appear to be relevant at central processing levels (see Verguts & Fias, 2006, for a similar approach).

4.2 Peripheral components in number transcoding

Writing of Arabic digits, numbers, and specifically number words do also comprise peripheral writing processes which are relevant for graphomotor planning and writing execution. However, the models presented above did not specify the aspect of written output in number transcoding. So far, only a few studies have addressed peripheral aspects of writing Arabic digits (Delazer, Lochy, Jenner, Domahs, & Benke, 2002, for an overview) and compared it to writing of letters. These studies were mainly inspired by insights from neurological patients. In some patients, commonalities have been described with regard to impaired language and number processing (Delazer & Denes, 1998; Gerstmann, 1940; Granà et al., 2001; Schonauer & Denes, 1994). In other patients, dissociations have been reported between impaired writing of alphabetic scripts and preserved writing of (single) Arabic digits (Anderson et al., 1990) and multi-digit numbers in peripheral writing processes (Delazer et al., 2002; Starrfelt, 2007; Tohgi et al., 1995). Table 2 compiles the opposing findings reported in the literature. In the discussion part of this thesis, I will evaluate my proposed integrated writing model for alphabetic and numerical scripts against these findings as the causes of impairments cannot be fully explained by the aforementioned writing models for letters and numbers.

Despite the findings in favor of domain-specific processing (Tohgi et al., 1995), there is no reason to assume that writing of letters is more difficult than writing of Arabic digits at the peripheral level of writing processes (Lochy et al., 2003). However, empirical evidence on this issue is missing. So far, only one study has been published which evaluates differences between writing of letters and Arabic digits (Lochy et al., 2003). For this purpose, the authors applied the model of language functions by Ellis (1988) to the number domain (see lower part of the model in Fig. 3.2). In their model, Lochy et al. (2003) proposed peripheral writing processes to start at the buffer level. The buffer was assumed to keep the abstract number identities in the correct order following the positional Arabic number system. Deficits to the buffer level were suggested by the authors to cause syntactic errors (e.g., *'1503' – '15003'*), lexical errors (e.g., *'2813' – '2613')* and mixed errors in number transcoding *('32104' – '31024',* Deloche & Seron, 1982). Lochy and colleagues (2003) argued that number transcoding occurs at the transition from central to peripheral components of writing at the stage of working memory processing. This is in line with findings from other studies showing that working memory capacity influences number transcoding performance (Barrouillet et al., 2004; Camos, 2008; Moura et al., 2013, for a similar role of working memory capacity in number transcoding).

Table 2: Studies on collectively and separately impaired linguistic and numerical processing

Studies	**Symptoms**
Collectively impaired linguistic and numerical processing	
Patient studies	
Gerstmann (1927, 1940) Gerstmann syndrome	• agraphia (impaired writing of alphabetic script) • acalculia (difficulties in arithmetic with normal intelligence) • finger agnosia (impaired identification of one's extremities) • left-right disorientation
Delazer & Denes (1998), Schonauer & Denes (1994) CK, female, 64 years old, right-handed left hemispheric ischemic stroke, mild right hemiparesis	• initial aphasia but well recovered spontaneous speech, auditory comprehension, and articulation • preserved number comprehension and magnitude comparison • severe agraphia and impaired writing of Arabic multi-digit numbers, i.e., syntax errors, lexical errors with length effects
Granà, Girelli, Gattinoni, & Semenza (2001) LD, male, 65 years old, right-handed left hemispheric stroke (CVA) right hemiparesis and hemianopsia	• anomic aphasia • normal phonological working memory (verbal and digit material) • severe alexia and agraphia for alphabetic and numerical scripts • faster recovery of Arabic digits and multi-digit-numbers during
Developmental studies	
Kinsbourne & Warrington (1963) Developmental Gerstmann syndrome	• symptoms are given above, but referred to as dysgraphia and dyscalculia
De Smedt et al. (2010); Peters & De Smedt (2018) Landerl et al. (2009)	• overlap between written language processing and arithmetic skills with phonological abilities as one key competence • development of language and arithmetic skills can be collectively or selectively affected
Dissociation between writing of letters and digits	
Patient studies	
Patterson & Wing (1989) DK, male, 60 years old, right-handed CVA in the left parietal region hemianopia and hemiplegia	• good comprehension, naming, and intact repetition of single words but impaired arithmetic skills • severely impaired writing of both scripts • case preferences in writing letters, mixed errors
Anderson, Damasio & Damasio (1990) female, 58 years old, right-handed circumscribed lesion in the left premotor area (BA 6) due to a metastasis from an adenocarcinoma	• no evidence of impairment of intellect or memory • severely affected reading, writing (but not spelling) and copying • severely distorted graphemes • preserved writing of single Arabic digits and multi-digit numbers
Toghi et al. 1995 male, 59 years old, right-handed left pre-frontal haemorrhagic infarction into the middle frontal gyrus and pre-frontal gyrus	• preserved reading comprehension, orally object naming, reading • impaired writing of Kana syllabograms but not Kanji morphograms • preserved writing of digits and multi-digit numbers • preserved magnitude comparison, simple addition and subtraction
Delazer et al. (2002) JS, male, 69 years old, right-handed lesion in the superior parietal (BA 7) and adjacent occipital (BA 19, precuneus) cortex post stroke	• initial but recovered ideomotor apraxia and remaining agraphia • well preserved number processing and calculation • preserved language comprehension, reading • impaired writing but not copying of alphabetic scripts
Starrfeld (2006) MT, male, 18 years old, right-handed head trauma following a car accident	• impaired reading and writing of alphabetic script but not numbers • preserved number writing and written arithmetic
Keller & Meister (2014) female, 61 years old, right-handed left hemispheric stroke to Exner's area	• acute global aphasia and right hemiparesis • impaired reading and writing of alphabetic script but not numbers

Note: The table comprises a selection of studies providing empirical evidence showing that language and number processing was found to be collectively impaired suggesting potential overlap between language and number processing. These studies were contrasted with findings in which writing of alphabetic and numerical scripts were presented to be selectively impaired.

An allographic level is then activated by the buffer level. At the allographic level, visuo-spatial aspects of the abstract number identities are held and retrieved. Allographic processing of digits was assumed to differ significantly from those of letters: numerical scripts contain a smaller set of digits (i.e., 10 Arabic digits) as compared to letters (i.e., 26 alphabetic letters). These digits do contain less invariants in handwriting like case and style. Accordingly, the set of allographs for abstract number identities is assumed to be smaller than the set for graphemes (Lochy et al., 2003). Subsequently, graphomotor patterns are assumed to be activated and translated into motor commands for neuro-muscular execution of writing. Evidence was provided that graphomotor patterns of digits are less complex as compared to the graphomotor patterns of letters because they contain less changes in direction (Lochy et al., 2003).

Consequently, letter and digit writing differ also at these processing stages. Further differences have been described regarding the contribution of semantics at peripheral writing processes, because semantic magnitude information is suggested to be evoked automatically in number processing (e.g., Dehaene & Cohen, 1997; Dehaene et al., 2003; Eger et al., 2003; Klein et al., 2010). Consequently, Arabic digits always comprise semantic information. Single letters or graphemes do not carry semantic information comparable to the numerical magnitude conveyed in numbers (Beeson et al., 2003). This difference may account for the observation of impaired writing of alphabetic but preserved numerical scripts at peripheral processing stages.

To sum up, in the language domain, research was significantly influenced by the serial dual-route model approach (but see Dell, 1988; Wilshire, 2008) adopted to the cognitive model of language functions by Ellis (1982,1988). This model suggests lexical (i.e., direct-lexical and lexical-semantic) and sub-lexical processing of (reading and) writing and later typing (Ellis & Young, 2013, Margolin, 1984) on two subsequent processing levels. On the central processing level, writing is referred to spelling performance; on the peripheral processing level, graphomotor execution of writing is carried out. More recent study results challenged the seriality of central and peripheral processing during writing (Kandel, Lassus-Sangosse, Grosjacques, & Perret, 2017; Kandel & Perret, 2015; McCloskey & Rapp, 2017) arguing for a continuous interaction between both levels. However, the existence of both levels has not been challenged.

In the number domain, semantic (i.e., semantic-lexical) and asemantic processing routes have been distinguished. The Abstract Code Model by McCloskey (1992) assumed number processing via one single, semantic route. Challenging this assumption, Cipolotti & Butterworth (1995) proposed an additional asemantic route involved in number transcoding. This multi-route-model approach was modified by the assumption of three abstract codes (i.e., auditory-verbal code, visual Arabic number format, analog magnitude representation) in the TCM by Dehaene (1992) and latest updates (e.g., Amalric & Dehaene, 2018; Dehaene, 2009; Dehaene & Cohen, 1995; Dehaene et al., 2003). However, cognitive mechanisms specifically related to serial writing processes were addressed only for asemantic number transcoding (Barrouillet et al., 2004). A serious limitation with all these models is that they consider neither peripheral aspects of writing nor number transcoding within different writing modes (i.e., handwriting or typing). So far, only Lochy and colleagues (2003) have acknowledged peripheral components of writing digits and numbers, but they did not acknowledge central processing levels. A model taking both processing levels into account has not been established in numerical research yet.

To address this issue, I propose a novel integrated model for writing alphabetic and numerical script in this thesis providing a theoretical framework for which individual tailored diagnostic procedures and interventions can be developed. For this purpose, writing of alphabetic and numerical scripts is considered at different processing levels in different (writing) contexts and from different empirical perspectives. Based on the above theoretical framework of central and peripheral writing components of both scripts, I pursued the following research objectives in the subsequent studies presented in the next paragraph for the development and evaluation of my integrated writing model:

(1) Starting from the differences in the cognitive processing of letters and digits at the behavioral and neuro-cognitive level in a patient study and in healthy controls, I investigated domain-specific and domain-general aspects of writing alphabetic and numerical scripts.

(2) In order to extend the integrated writing model to the changing writing habits due to digital writing media in the last decades, I examined the comparability of writing modes (i.e., handwriting and typing) in the diagnosis of writing and arithmetic skills in typically and atypically developing children.

(3) Finally, I evaluated the integrated writing model based on the findings derived from the assessment and training of orthography and arithmetic skills as well as based on the current literature in order to test its validity for written language and number processing.

5 Objectives addressed in the studies enclosed

The guiding question underlying the present thesis is whether the cognitive mechanisms of writing letters and numbers can be theoretically captured from a cross-domain perspective, and thus, can be integrated in one shared writing model for alphabetic and numerical scripts. However, both linguistics and numerical cognition have not yet provided sufficient evidence for such an integration. By means of the six studies presented in the following three sections, I aim at contributing to provide this missing evidence.

(1) The first section, inspired by an individually tailored writing regimen in a stroke patient, outlines two studies investigating neuronal correlates of writing single letters and digits. Neural correlates were examined in a patient study (Study 1); neural connectivity was studied in healthy adults (Study 2). Study results provide compelling insights concerning the necessity to extend current writing models.

(2) The second section investigates the influence of the actual writing mode (i.e., handwriting or typing) on spelling and arithmetic performance in different groups of children (i.e., typically developing children and children with dyslexia, Study 3), and different tasks, (i.e., writing-to-dictation and basic calculation, Study 4). Results are presented in view of the changing writing habits from handwriting to computer keyboard typing in recent decades. Furthermore, it will be demonstrated how predominantly use of digital technology may change the cognitive mechanism underlying writing and hence, may influence diagnostics in educational and therapeutic contexts.

(3) In the third section, two studies report on the application of a computerized learning platform for spelling and arithmetic in different learning contexts: from a diagnostic perspective, an individually tailored adaptive assessment for spelling and arithmetic is presented (Study 5). From an interventional perspective, learning games developed to promote literacy and numeracy are evaluated on a behavioral and a neuro-cognitive level in a training study (Study 6).

The individual objectives motivating each of the six studies will be outlined in the following as well as, what was investigated and why. Major results of the studies in the three sections are summarized in terms of their contribution to the development of the integrated writing model in the general discussion.

5.1. Section 1: Writing letters and digits – similar but not the same

The significance of writing – be it letters or Arabic digits – to cope with everyday life in our literate society cannot be emphasized enough. However, we often do not recognize this significance until writing is suddenly impaired, for instance, as a result of a stroke.

Writing can be partially or totally impaired at central and/or peripheral processing levels referred to as central or peripheral dysgraphia/agraphia. Central agraphia often occurs in association with other acquired language disorders such as aphasia (e.g., Delazer & Bartha, 2001) and may affect writing in all domains (i.e., letters and digits in the context of acalculia) and output modalities (i.e., oral spelling, handwriting and typing, Beeson & Rapcsak, 2015; Rapcsak et al., 2002). Peripheral agraphia, in contrast, is related to graphomotor planning and execution of writing and thus specific to the output domain and mode (Lochy et al., 2003; Purcell, Turkeltaub et al., 2011). However, writing of letters and digits is not always equally affected. Writing of numbers is often preserved as presented in the first part of this thesis. The exact nature of this dissociation has not yet been clarified.

5.1.1 Study 1: *What letters can 'learn' from Arabic digits – fMRI-controlled single case therapy study of peripheral agraphia*

Study 1 begins right at this point by reporting on an fMRI-controlled training study in a stroke patient. In this patient, handwriting (but not typing on the computer keyboard) of words and even single letters but not numbers was severely impaired. Starting from the question why numbers are often less impaired than letters, Study 1 presents an individually tailored writing training. In this training, mental images are assigned to all letters of the alphabet (*A* looks like a tepee, *B* looks like glasses, etc.). Mental images added additional semantic information to letters, similar to the number magnitude information (Dehaene & Cohen, 1995; Dehaene et al., 2003) automatically activated when processing Arabic digits (Eger et al., 2003; Klein et al., 2010). At the behavioral level, the overarching goal of the training was to enable the patient to write manually again. At the neuro-cognitive level, attempts were made to document the neuronal correlates of the writing training. The, the focus was on the different cortical regions specifically involved in writing letters and digits.

However, identification of cortical regions by fMRI data does not allow to comprehensively explain the functional mechanisms of the writing training presented in Study 1. Moreover, dissociations between writing letters and digits cannot be fully explained. The reason is that fMRI data do not necessarily indicate that activated areas are functionally critical for the task (Rorden & Karnath, 2004). Integration of behavioral data, functional data as well as structural data of the brain is a promising approach to overcome this issue. One approach for such an integration is provided by the observation of neuronal fiber pathway connections by tractography. Tractography allows to indicate by which fibers (white matter tracts) cortical regions (gray matter) are connected (e.g., Assaf & Pasternak, 2008, for a review; Basser, Pajevic, Pierpaoli, Duda, & Aldroubi, 2000; Poupon et al., 2000). Identifying such pathways is particularly important as brain lesions are often not circumscribed in the brain; damage to a specific region in the gray matter often impacts also on the associated fibers in the white matter. Thus, not only damage in a certain brain region can lead to a loss of function, but also damage to the fibers connecting regions involved in the same function (Catani et al., 2012, for a review).

5.1.2 Study 2: *Differing connectivity of Exner's area for numbers and letters*

Accordingly, in Study 2, atlas-based tractography was employed for the patient's anatomy (Study 1) to show which pathways were connected to lesioned or unaffected tissue in the brain. Subsequently, connectivity of the identified regions associated with writing letters or digits was investigated by means of deterministic fiber-tracking in healthy adults. The aim was to provide further evidence for the conclusions of Study 1 which say that dissociations between letter and digit writing are associated with differing connectivity of the brain regions specifically subserving language and number processing.

5.2 Section 2: Handwriting and typing - modal influences on spelling and arithmetic

How do writing processes change when we switch from pen(cil) to keyboard? Are these changes relevant to writing performance and how comparable are findings recorded in a paper-pencil or a computerized diagnostic procedure? These are the core questions of the subsequent paragraph. The answers to these questions were used to extend the writing model to different writing modes.

Three major aspects motivated these questions: first, digital technology has changed writing habits and processes in a way necessary to consider for the evaluation of writing in educational and therapeutic contexts. However, previous research on the influence of writing mode has revealed contradictory results for different populations (e.g., Berninger et al., 2009; Connelly et al., 2007; Goldberg et al., 2003, for a meta-analysis; Wollscheid et al., 2016, for a meta-anaylsis). Second, clinical observations of the patient from Study 1 revealed that writing modes do not have to be equally affected in acquired writing impairments; handwriting for letters was specifically impaired whereas typing on the computer keyboard was well preserved (Jung et al., 2015). Third, computers are supposed to reduce writing barriers for developmental writing difficulties such as in dyslexia. However, recent studies have neglected the influence of writing mode in children with developmental writing disorders (i.e., in the context of dyslexia). However, unambiguous evidence is particularly important as computers have high potential to increase participation of students with dyslexic writing problems in school (e.g., MacArthur, 2009). However, the mode effect, if not carefully understood, can (negatively) affect the evaluation of writing skills in any learning context.

5.2.1. Study 3: *Mode effect: an issue of perspective? The influence of writing mode in children with and without dyslexia*

Study 3 addressed these aspects by implementing a combined within- and between-subject design to provide first empirical evidence on how writing performance is affected by different writing modes in typically and atypically developing children. In a novel two-step framework, handwriting and typing skills were evaluated using a writing-to-dictation task. In the first step, children's spellings were assessed as either correct or incorrect as it is the case in many learning assessments. In the second step, spelling rule-specific analyses were applied. Thereby, different spelling rules were evaluated separately for each test word to avoid generalizations of a potential mode effect from a specific spelling rule to others. The rule-specific analysis allowed the separation of central spelling processes from peripheral writing processes, and hence, a detailed sequential analysis of handwriting and typing. Accordingly, Study 3 opens up new perspectives on the analysis of mode effects and provides future directions for (computerized) evaluation of writing in this relatively new field of research.

However, in view of the different findings in the literature, comparability between handwriting (i.e., paper-based testing) and keyboard typing (i.e., computerized testing) is not self-evident. Empirical studies are needed to shed light on the impact of media beyond mere writing processes. In this respect, one aim of Study 4 was to provide empirical data on spelling and basic arithmetic skills of children in a computerized testing. These data are compared with findings from previous studies using traditional paper-pencil tasks. This comparison allows conclusions as to whether processes beyond writing are also influenced by the writing mode. Results are particularly important for the development of up-to-date diagnostic and intervention approaches.

5.2.2 Study 4: Closing the gap: Bounded and unbounded number line estimation in secondary school children

In Study 4, the development of basic numerical and arithmetic skills is assessed and associated with spelling performance. Previous studies using traditional paper-pencil diagnostics have shown that children with poor literacy skills underperformed also in arithmetic tasks, while children without language difficulties did not (e.g., De Smedt et al., 2010; for a further discussion see Moll et al., 2009). In this vein, it has been shown that spelling and arithmetic performance in children correlate at a medium to a high level (e.g., a higher spelling performance usually goes along with higher arithmetic performance). However, these conclusions were drawn on the basis of traditional test procedures (i.e., paper-pencil tests). With regard to the mode of testing, Study 4 attempted to clarify whether similar correlations can be found when spelling and arithmetic skills are evaluated by means of a computerized assessment. These findings may indicate whether test results from traditional and digital approaches are comparable.

5.3 Section 3: Writing and beyond - assessment and training of orthography and numeracy

Section 3 is devoted to literacy (i.e., basic reading and writing) and numeracy. Both, literacy and numerical education have gained increasing research interest due to the PISA studies and their results. Remediation research, however, still lags behind. This is particularly serious, because recent studies have clearly indicated that poor literacy and numeracy increases the risk for school dropout and unemployment as well as for lower self-esteem and health prospects (e.g., Daniel et al., 2006, for literacy ; Parsons & Brynner, 2005, for numeracy). These consequences are not only serious for the individual but can also be socially significant (Beddington et al., 2013). Thus, the development of effective training programs to benefit students with learning difficulties (i.e., dyslexia and/or dyscalculia), is a strong requirement for research and practice.

Nowadays, for children the latest computer and communication technologies (such as smartphones, tablets, etc.) are very attractive. Thus, interest in using digital games to enhance education (for a review, see Boyle et al., 2016) and for remediation approaches is growing. In the domain of literary and numeracy, research on game-based learning is still in its infancy. However, recent studies have examined promising learning outcomes (e.g., Kast et al., 2011; Ninaus et al., 2016). Against this background, this section presents two different studies evaluating a novel assessment for (Study 5) and training of (Study 6) orthography and arithmetic.

5.3.1 **Study 5:** *Die TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)*

In Study 5 motivational aspects and barriers for digital assessment and learning games have been reviewed in order to provide a coherent picture of potential benefits of digital game-based learning. A recent approach how to tailor the assessment of orthography skills to the individual's learning level has been presented. This approach followed the rule-specific analysis presented in Studies 3 and 4. Analyzing spelling performance following the rule-based approach enables to identify specific writing difficulties which is essential for the development of effective interventions meeting individual needs.

Finally, digital learning games were evaluated with respect to their effectiveness in conveying orthography and arithmetic skills. By integrating findings from Study 5 and the following study, a comprehensive picture is provided regarding the use of digital media in different learning contexts.

5.3.2 **Study 6:** *Behavioral and Neurocognitive Evaluation of a Web-Platform for Game-Based Learning of Orthography and Numeracy*

Study 6 evaluated the application of the digital learning games presented in Study 5 in a training following a cross-over design. Both spelling and arithmetic performance were assessed at three time points to track learning progress: at the beginning, then after three and six training sessions. Intervention effects of the learning games were assessed. The assessment proceeded at two levels. On a general level, overall effectiveness of the games was evaluated. On a specific level, the study design allowed for conclusions regarding specific trainings effects for both domains.

PART II: EMPIRICAL STUDIES

Section 1:

Writing letters and digits: similar but not the same

Study 1:

What letters can "learn" from Arabic digits -
fMRI-controlled single case therapy study of peripheral agraphia.

Stefanie Jung, Katja Halm, Walter Huber, Klaus Willmes, Elise Klein

Abstract

Research on (hand-)writing has revealed that *Exner's area* subserves transferring linguistic impulses into writing programmes.

We report on a patient with a lesion affecting Broca's and Exner's area suffering from severe peripheral agraphia for letters but not for Arabic digits. Analogous to semantic (magnitude) information in numbers, we developed a specifically tailored writing training: additional mental imagery based semantic information was attached to letters. The training resulted in significant improvements. Imaging data revealed stronger fronto-parietal network activity including perilesional activation around Exner's area and precuneus for writing letters to dictation than for writing letters corresponding to their mental image expressions. Follow-up testing showed not only stable training effects but also an activation shift into the left angular gyrus.

Results document neuronal correlates of a successful intervention by attaching additional meanings to letters in order to retrieve their grapho-motor patterns. These findings contribute to understanding the impact of Exner`s area.

1. Introduction

Writing as a major cultural accomplishment is vital to cope with the requirements in a literate society. An acquired impairment or loss of writing may lead to significant restrictions in everyday life. Hence, it is crucial to understand writing processes and their underlying neural substrates to develop effective interventions.

Insights from brain lesion studies of patients with acquired writing impairments (e.g., Exner, 1881, Keller & Meister, 2014) and recent research in healthy participants (e.g., Beeson et al., 2003; Longcamp, Anton, Roth, & Velay, 2003; Vinckier, Dehaene, Jobert, Dubus, Sigman, & Cohen, 2007) revealed critical areas of the brain involved in writing processes by means of functional neuroimaging techniques. A recently published meta-analysis of neuroimaging studies on handwriting by Planton, Jucla, Roux and Démonet (2013) identified a left-hemispheric parieto-frontal network in the brain to be involved in writing processes. This network includes middle frontal gyrus, superior frontal sulcus, inferior parietal lobule, superior parietal lobule (SPL), intraparietal sulcus (IPS) and the right cerebellum. Supplementary (SMA) and pre-supplementary (pre-SMA) motor areas, thalamus and putamen were related to rather unspecific motor processes during handwriting. Ventral premotor areas and both inferior and posterior temporal areas were associated with linguistic processing. Importantly, the frontal "writing center" of Exner, already postulated by Exner in 1881 and recently revived by Roux, Draper, Köpke and Démonet (2010), in the posterior part of the middle frontal gyrus (MFG, Brodmann Area [BA] 6) - subsequently called "Exner's area" - appeared in this meta-analysis "as the area of strongest and most reliable activity during handwriting" (Planton et al., 2013; p. 10). Planton and colleagues (2013) redefined Exner`s area as a region bridging the gap between graphemic abstract representations and motor execution programmes of handwriting. They agreed with findings based on clinical data from Binkofski and Buccino (2004), who described Exner's area as the final pathway transferring linguistic impulses into writing programmes (i.e., grapheme formation and their temporal sequencing). Furthermore, they corroborated earlier statements by Roux and colleagues (2009), who referred to Exners`s area as "graphemic/motor frontal area (GMFA)" - an interface between orthographic (abstract) representations and the 'allographic' specification of the grapho-motor programmes - which is located between the central and the peripheral stages of handwriting (Roux et al., 2010). Although descriptions of patients show that brain lesions around this area may result in writing difficulties (Exner, 1881; Tohgi et al., 1995), the status of Exner`s area as writing-specific brain region has been and is still the subject of controversial debate. In particular, its crucial role during handwriting is still unclear, because: i) language organisation in the brain shows high intra-individual variability (Lubrano, Roux, & Démonet, 2004); ii) pure lesion data have the potential to provide causal structure-functional relationships; but only, when combined with functional imaging one can identify, which parts of the frequently larger structural lesions in some patients are still functionally active as compared to other patients with similar pathology or whether other areas may be involved compensating lost functions due to functional reorganisation in chronic patients; iii) although functional imaging studies are informative regarding potential areas functionally involved in writing, they do not provide causal structure-functional relationships but only correlations. The observed brain activity could also reflect unspecific co-activation of the respective cortex areas (Klein et al., 2013). Thus, drawing general conclusions regarding the role of Exner`s area requires further research.

In this vein, we present a combined structural and functional single case brain lesion study of a patient (CU) showing a single lesion in parts of Broca's Area (Area 44/45) extending superiorly into Exner`s area (BA 6) after a left hemispheric stroke. The patient suffered from severe peripheral agraphia.

Agraphia is generally categorised into central writing impairments, affecting spelling in various output domains (e.g., oral spelling, handwriting, typing) and peripheral writing impairments, which mainly affect only one domain - mostly handwriting (Beeson & Rapcsak, 2004). In various models of written language processing multiple stages are postulated to be involved in peripheral writing processes (e.g., Ellis, 1988; Hillis & Caramazza, 1989; Rapcsak & Beeson, 2002). Although literature is still lacking consensus regarding the precise characterisation of these stages (Purcell, Turkeltaub, Eden, & Rapp, 2011), the following components are assumed to be indispensable for handwriting: First, at the grapheme level (Ellis, 1988) - resp. the graphemic buffer - abstract representations of letter identities have to be activated. Second, abstract representations of letters are converted into specific letter shapes or allographs, generating a concrete representation of letter forms (e.g., case: lower or upper and style: print or cursive). Third, at the level of graphomotor patterns, selected allographs are specified (e.g., order, position, direction). Finally, information encoded in graphomotor patterns is translated into motor commands to execute writing movements (Ellis, 1988; Rapcsak & Beeson, 2002).

Insights from patients with neurological disorders provide a large but also diverging picture of peripheral agraphia and its accompanying disorders (Delazer, Lochy, Jenner, Domahs, & Benke, 2002, for an overview); but only a few cases have revealed dissociations between impaired writing of alphabetic scripts and relatively preserved writing of Arabic digits after lesions in Exner's area (Anderson, Damasio, & Damasio, 1990; Delazer et al., 2002; Keller & Meister, 2014, Starrfeld, 2007).

In the present case, the patient was well able to type in words and pseudo-words on a computer keyboard upon request during dictation. Crucially, spontaneous hand-writing and writing to dictation of words and even single letters was impossible, while copying of words/letters was preserved. Writing of numbers to dictation appeared largely unaffected at least up to three-digit numbers. This is of particular interest, as the most influential model of numerical cognition, the triple code model, assumes that after encoding numbers - even single digits - are linked to semantic information about numerical quantity (e.g., Dehaene & Cohen, 1997), even when the numerical quantity is irrelevant or disturbing for solving a task at hand (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003; Klein, Moeller, Nuerk, & Willmes, 2010). This automatically activated representation of numerical magnitude is supposed to be subserved by areas in the intraparietal sulcus (IPS), bilaterally (Dehaene, Piazza, Pinel, & Cohen, 2003). Numbers are also processed via a verbal representation of numbers in language-mediated operations like number naming, counting or fact retrieval (Delazer et al., 2003; Grabner et al., 2009). This verbal representation is associated with left perisylvian language areas such as middle (MTG) and superior temporal gyrus (STG), supramarginal gyrus (SMG), Broca's area, and the angular gyrus (AG) (Arsalidou & Taylor, 2011). Though, both, letters and numbers are linked to verbal linguistic information, single letters or graphemes - considered to be the smallest distinguishable graphic units of the writing system - do not carry semantic information comparable to the numerical magnitude conveyed in numbers. Hence, semantic magnitude information of numbers seems to facilitate their processing. This assumption can account for the observation of impaired letter processing with simultaneously preserved number processing in patients after brain lesions.

Based on these considerations, we wondered whether it would be possible to associate some kind of semantic cue to each letter of the alphabet in order to support retrieval of letter forms during writing. We decided to use mental imagery for two main reasons: On the one hand, letter forms allow for association to real world objects, which the patient was able to draw on request. On the other hand, mnemonic effects of imagery on the recall of verbal representations have been well-established in the Dual Coding Theory initially proposed by Paivio (1971). This theory builds on two functionally independent, but interacting, associative networks of verbal and non-verbal (imagery) information linked to sensory input and response output systems. Thus, visual, auditory as well as motor properties of language and objects can be processed and stored in memory. The retrieval of verbal memory contents can be enhanced by mental images associated to and evoked by word material.

Based on this theoretic account, we developed a specific training approach together with the patient, in which he learned to associate specific mental images of objects, which resemble the visual word form of the letters in one to three words (e.g., "Y" looks like a "cocktail glass"). These mental images provided cues for activating additional semantic information together with the letters. Since we wanted to elicit and strengthen mental imagery, the mental images were neither presented pictorially nor printed. The patient only imagined the mental images.

1.1. Study rationale and working hypotheses

The aim of the present study was to evaluate a specific training approach for re-establishing motor images of letter production in a patient with peripheral agraphia. Similar to numbers, which always convey additional semantic information assumed to facilitate their processing, the idea was to associate additional semantic (visual) information with letters by means of mental imagery. In particular, we expected that writing numbers should be well preserved. The patient, however, was expected to fail in writing letters to dictation but not in copying letters, since several case studies on peripheral agraphia have revealed dissociations in writing performance when comparing both tasks (Delazer et al., 2002, for an overview). Short-term and long-term training effects were expected for writing to dictation.

With respect to brain activation patterns we expected that i) IPS activation should be observed for numbers. Peri-lesional activation next to Exner`s area should be present for writing numbers and letters similarly. When writing letters to dictation, retrieval of mental object images should yield additional activation in the precuneus. This area has been proposed to be involved in visual-spatial mental imagery as well as in memory-related visual imagery, self-processing and consciousness (e.g., Cavanna and Trimble, 2006 for a review; Ganis, Thompson, & Kosslyn, 2004). This activation pattern would be taken to indicate that the patient indeed relied on the retrieval of mental images. ii) The direct contrast of writing letters to dictation compared to writing letters after the auditory presentation of the mental images for letters should reveal still stronger activation in (the vicinity of) Exner`s area and mental imagery related precuneus activations. Both areas subserve writing letters compared to writing mental images for letters, but letters have to be associated with the corresponding mental images first, implying additional processing load. iii) We expect brain activity changes eight months later at follow-up, presumably due to generalisation processes.

2. Methods

2.1. Patient

CU was a 52-year-old right-handed former civil engineer suffering from severe peripheral agraphia due to a left hemispheric stroke after spontaneous dissection of the internal carotid artery. Initially, CU also showed severe mixed transcortical aphasia and right hemiparesis.

The present study was conducted 3 years and 4 months post-onset during CU`s rehabilitation stay at the Aachen Aphasia Rehabilitation Ward in the University Hospital of the RWTH Aachen. On the ward patients receive intensive language treatment (duration of 7 weeks, more than 8h a week) with individually defined treatment goals according to an established therapy regimen (Huber, Springer, & Willmes, 1993). Relatively spared attention, memory, non-verbal, and verbal learning capabilities are required to profit optimally from this intensive treatment. MRI scanning during the stay revealed a brain lesion mainly affecting triangular and opercular parts of Broca's Area (Area 44/45) extending superiorly into Exner`s area (BA 6). Furthermore, the lesion involved the insula, the internal capsule, as can be inferred from the well visible Wallner degeneration of the fibers in the left peduncle of the brain stem (Fig. S1 in the supporting online material (SOM)), parts of the lentiform nucleus, the ventral (EC/EmC) as well as parts of the dorsal (arcuate fascicle) pathway of the dual loop model of language processing and brain organisation more generally (Weiller, Bormann, Saur, Musso, & Rijntjes, 2011). This cortical lesion pattern leads to white matter degeneration from frontal to inferior parietal lobe.

The patient gave his written informed consent to participate. The study was approved by the local Ethics Committee of the Medical Faculty of the RWTH Aachen University Hospital (protocol number: EK094/07) and performed in accordance with the latest version of the Code of Ethics of the World Medical Association (Declaration of Helsinki).

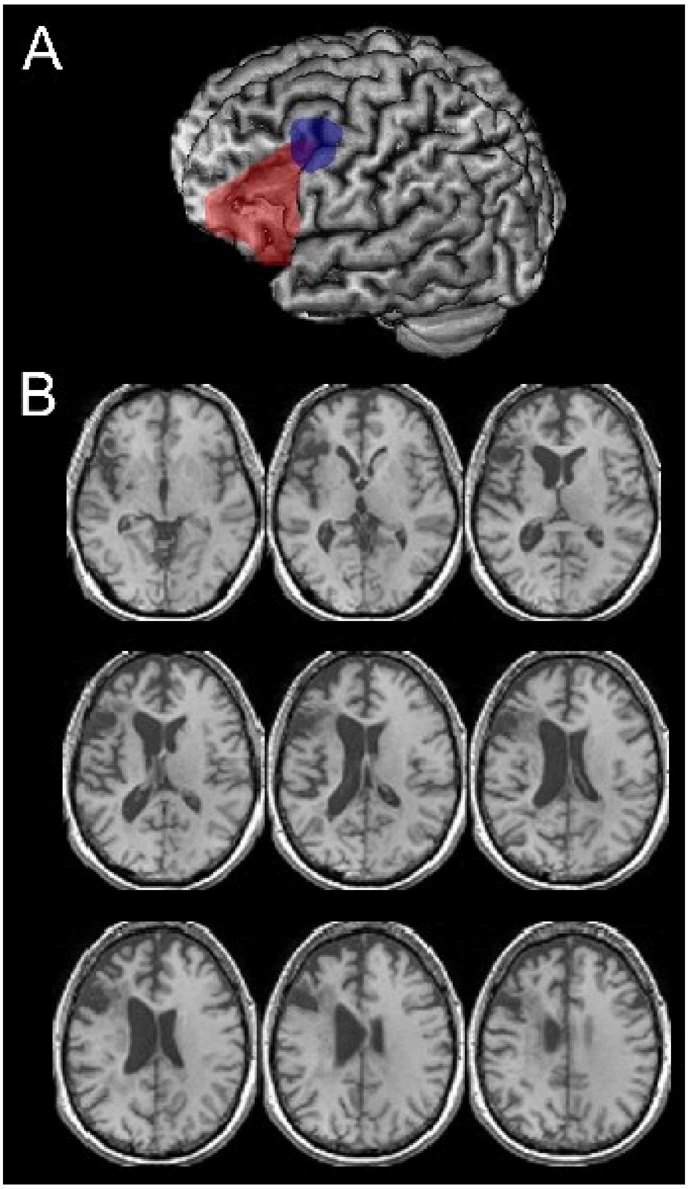


Fig. 1: Panel A depicts a 3-D reconstruction of patient CU's brain, depicting the frontal lesion seven years after the incident. The red tag depicts the infarct area, blue indicates Exner`s area (BA6). B: Patient CU's lesion affecting Broca's area and extending into Exner's area due to a left-hemispheric MCA stroke after ACI dissection seven years before the scan depicted on transverse slices of a high-resolution T1-weighted structural fMRI scan.

2.2. Neurolinguistic and neuropsychological assessment

CU's linguistic performance was assessed several times using the Aachen Aphasia Test (AAT; Huber, Poeck, Weniger, & Willmes, 1983; for a description of the English version of the test see Miller, Willmes, & De Bleser, 2000). The results of all neurolinguistic and neuropsychological assessments are summarised in Table 1. In the course of recovery, the clinical aphasic syndrome assignment changed from mixed transcortical aphasia to amnesic aphasia. According to the AAT spontaneous speech ratings before onset of the intensive treatment period, spontaneous speech was characterised by few phonemic paraphasias and articulatory errors. Mild lexical and syntactic difficulties were also apparent. In formal assessment with the AAT subtests, repetition of phonemes and words was nearly accurate but errors occurred with increasing length of utterances. Language comprehension (auditory and written) was mildly impaired. Reading aloud as well as composing words and phrases, respectively, from cards with graphemes resp. morphemes was well preserved. Writing to dictation of words and single letters was severely affected. Additional written language diagnostic assessments revealed that CU was, nevertheless, able to orally spell and copy words with his non-paretic left hand and to type in both words and pseudo-words on a computer keyboard. Moreover, he was capable to transpose pseudo-words from small to capital letters and vice versa (Fig. 2A and Fig S2 in the SOM).

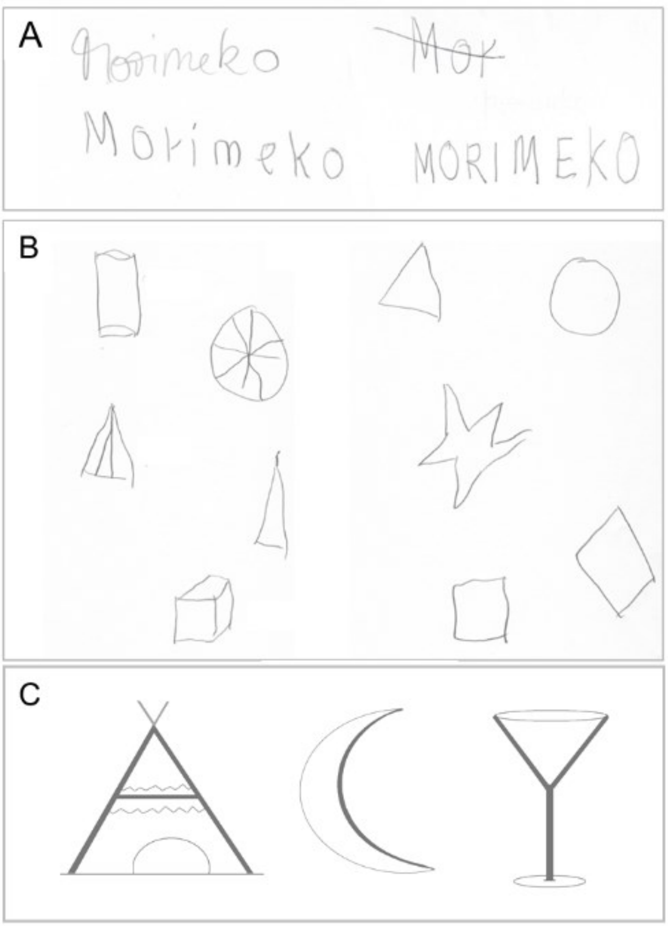


Fig. 2: Neurolinguistic and neuropsychological assessment: A) CU's performance on copying and transposing the pseudo-word "Morimeko". B) CU's drawings from memory of geometric figures. C) Training material: Illustration of mental images associated with the letter forms of the upper case Latin alphabet (examples).

The neuropsychological diagnostic examination (cf. Table 1) focused on attention, working memory, non-verbal intelligence, verbal, and non-verbal learning as well as on spatial-constructive abilities. Attention functions (tonic and phasic alertness, selective and divided attention) were in the average range for healthy participants as assessed with the “WAF- Perception and Attention Functions” test battery (Sturm, 2006). Working memory, measured with the German version of the ‘Wechsler Memory Scale-R’ (WMS-R, Wechsler, 1987) was severely impaired for digit span forward and backward, whereas CU`s performance in a visual working memory task was only slightly below average (Block-Tapping-Test; Schelling, 1997). Recognition of spatial rotations was below average, but abstract reasoning was preserved as assessed with the Performance test system for the 50 to 90-year-old (Leistungsprüfsystem LPS50+; Sturm, Willmes, & Horn, 1993). The below average results of the verbal and non-verbal learning test (VLT and NVLT; Sturm & Willmes, 1999) could not be interpreted unambiguously, because the patient’s visual discrimination abilities could not be considered non-impaired. The assessment of spatial-constructive abilities measured among others with the subtest Block Design from the Hamburg-Wechsler Intelligence Scale (HAWIE-R; Tewes, 1991) and the Rey complex figure revealed non-impaired performance. The patient was able to draw geometric figures from memory upon request with his non-paretic left hand (cf. Figure 2b). Both, assessments of apraxia (modified after DeRenzi, Motti, & Nichelli, 1980; Poeck & Kerschensteiner, 1975) and neglect (Neglekt Test, NET; Fels & Geisner, 1996) revealed no impairments. Moreover, the Edinburgh Handedness Inventory (Oldfield, 1971) was used to determine handedness before stroke (Lateralisation Quotient (LQ): L = +00; decile R.10). No case of left-handedness was described in the family.

In addition, CU`s numerical skills were examined with a focus on his ability to write numbers in comparison to writing letters. The Numerical Processing and Calculation test battery (NPC; Delazer, Girelli, Graná, & Domahs, 2003) and some additional numerical tasks were administered to get a comprehensive overview about most of the relevant basic number processing and calculation abilities. Assessment with the NPC revealed unimpaired counting (verbal and written) and parity judgement performance. Reading aloud of Arabic numerals (e.g., 34), number words (e.g., thirty-four), and transcoding from tokens to Arabic numerals (e.g., tokens carrying different values such as □□□□□□□ with □=10 and □=1) and vice versa were unimpaired. The same was true for writing Arabic numerals to dictation, at least up to three-digit numbers. Handling larger numbers in this dictation task led to permutations, substitutions or omission errors (e.g., four-thousand-twenty-eight □ 4280 or eight-thousand-two-hundred-seventy □ 8720). These errors occurred very likely due to increased (verbal) working memory load, since stimulus length also influenced the patient`s performance in word repetition; but also syntax errors are plausible.

Calculation of simple arithmetic facts of the NPC revealed a different pattern: whereas addition and multiplication were unimpaired, CU had considerable difficulties in subtraction and particularly division. Operation specific rule knowledge (e.g., n×1= n), however, was well preserved. Both mental and written calculation was rather difficult for the patient. Interestingly, CU achieved more correct results in written calculation, but still below the cut-off score.

These findings indicate that CU`s numerical skills and even his ability to write numbers were, although not totally unimpaired, far better than his writing abilities. In sum, CU showed severe peripheral agraphia for letters, but not for Arabic digits. Since brain lesions in Exner`s area can affect the selection of grapho-motor patterns, we hypothesised that the impairment was due to impaired pre-motor selection or deficient recall of allographic information of the specific grapho-motor patterns. For this reason, intrinsic grapho-motor planning for writing letters and words, respectively, was severely impaired for the patient.

2.3. Training

CU was trained to associate specific mental object images with letters according to an individually tailored training procedure. The mental object images and their accompanying one to three word utterances served as cues, resp. overlearned key words to facilitate the retrieval of letter forms by means of mental imagery. These mental images used by the patient were never "externalised" as (printed) drawings or figures.

In aphasia naming therapy it is very common to use phonological cues to facilitate the retrieval of a target word (e.g., apple: The target word starts with an [æ]) (Nickels, 2002, for a review). In primary German schools during first language writing acquisition initial sound tables are commonly applied, encouraging children to independently explore letter-sound combinations (e.g., "a" like in apple). In line with these techniques, we designed a collection of mental images as a semantic cue of the letter shape. For that purpose, patient and therapist associated the 26 cardinal letters and the three umlauts (Ä,Ö,Ü) of the German alphabet with specific mental images of objects that described the visual word form of the letters in one to three words (e.g., "Y" looks like a "cocktail glass" as shown in Figure 2C). A listing of the key words can be found in Table S1 in the SOM.

Table 3: Overview of the neurolinguistic and neuropsychological assessment

Test	Subtest	Performance			
LANGUAGE					
Aachen Aphasia Test	**Duration post onset**	**0.4**	**1.1**	**2.3**	**3.4**
Spontaneous Speech rating scales	Communication	2	3	3	4*
	Articulation and prosody	3	4	4	4
	Formulaic speech	2	3	3	5**
	Semantic structure	2	3	3	4*
	Phonemic structure	4	4	4	4
	Syntactic structure	3	4	4	4
	Communication	2	3	3	4*
Subtests (Raw scores/Percentiles)	Token Test	49/5	25/53*	29/47*	21/63**
	Repetition	140/86	142/88	141/87	140/86
	Written Language	36/41	28/35	52/53**	62/63**
	Reading aloud (words/phrases)		28/30	28/30	28/30
	Composing words/phases		0/30	23/30	26/30
	Writing words/phrases to dictation		0/30	1/30	7/30
	Confrontation Naming	72/46	70/44	89/62**	96/74*
	Comprehension	95/75	88/64	101/86	87/62
ATTENTION, WORKING MEMORY INTELLIGENCE & LEARNING					
		Test Scores		**Percentiles**	
WAF- Perception and Attention Functions	Intrinsic Altertness	T = 39		PR 14	
	Phasic Altertness	T = 40		PR 17	
	Go/No-Go-Test	T = 53			
Performance test system LPS50+	Abstract reasoning	T = 42			
	Mental rotation	T = 25°			
	Registration of forms and shapes	T = 29°			
Hamburg-Wechsler Intelligence Test	Block Design	WP = 8			
Rey complex figure		accurate			
Wechsler Memory Scale-R	Digit Span forward	5 numbers			
	Digit Span backward	3 numbers°			
Corsi Block Tapping Test		5		PR 37	
Verbal and Non-verbal learning test		T = 20		PR 0	
NEGLECT AND APRAXIA					
3D Apraxia Test		accurate (130s/ 3 figures)			
Neglect Test	2D – Test	accurate			
	3D – Test	minimal impairment			
Buccofacial Apraxia		30/30 correct			
Ideomotor Apraxia		53/72 correct			
NUMBER PROCESSING					
Number Processing and Calculation Battery	Counting	5/6	correct		
	Comprehension of Numbers	55/56	correct		
	Numerical Transcoding	40/56	correct °		
	Reading	28/28	correct		
	Writing	12/28	correct °		
	Calculation of simple arithmetic facts	125/152	correct °		
	Addition/Multiplication	60/66	correct		
	Subtraction/ Division	38/50	correct °		
	Multiple Choice	27/36	correct °		
	Mental calculation	3/20	correct °		
	Written calculation	7/20	correct °		
	Arithmetic Principles	28/30	correct		

Neurolinguistic assessment: all, except the first, language assessments were conducted at the Aachen Aphasia Ward; * significant improvement as compared to the first investigation 4 months post onset; ** significant improvements as compared to the preliminary investigation. *Neuropsychological assessment*: ° performance below average as compared to healthy subjects.

The training was administered daily (five days a week) over a period of three weeks including one weekly monitoring assessment (15 training sessions in total). For the training the letters of the alphabet and the corresponding written mental image were printed on cue cards. All 29 cardinal letters were practised using training words. Only one new letter/mental image expression was introduced per training word. The training schedule was arranged in a cumulative manner; the number of training words (and thus the number of letters/mental images) accumulates, whereas the number of new letters introduced per training session decreased (sessions 1 to 10: 2 × 5 letters, 1 × 4 letters, 4 × 3 letters, 3 ×1 letter). Sessions 11 to 15 were used for stabilisation.

Each training session was divided into two phases; a first phase of associating and learning isolated mental images for letters and a second practice phase, in which the associated mental images were used for writing at the word level. The association phase had three steps: i) the therapist presented a new training word (e.g., "BALL", B and A were already practised letters) and, thus, one new letter (e.g., "L"). ii) The patient repeated and spelled out orally the training word (e.g., "B-A-L-L"). iii) Then every single letter was worked on in writing. In particular, the patient had to remember the already trained mental images (B associated with "glasses", A associated with "Indian tent"), or was asked to associate the shape of the newly introduced letter with another real world object (L associated with "corner"). In this way, a new mental image was established. In the practice phase, CU wrote the target word in capital letters (e.g., "BALL") using the mental images and read it aloud. When the word was correctly written, he had to transpose it into small letters (e.g., "BALL" into "ball"). In case of misspellings the word was corrected with support from the therapist and re-written. The patient got homework in order to stabilise the training progress. To this end, he used the cue cards and a poster in his room on the ward showing both letters and mental image expressions. The homework was always discussed with the therapist and corrected if necessary at the beginning of the next training session.

2.4. Behavioural data analysis

Behavioural performance was assessed in four different ways: i) first, training progress was strictly monitored. Monitoring consisted of three steps: In the first step, the therapist dictated a letter and the patient had to name the corresponding mental image. In the second step, the therapist named a mental image and the patient was asked to write the corresponding capital letter. The third step concerned the word level. The patient was given either an already trained or a non-trained word in order to write it down. This step included the tasks spelling out the word, naming the mental images, and writing down the word in capital letters. In all steps performance was assessed either as correct or incorrect ; ii) second, three intermediate tests were conducted after each week of training, assessing the retrieval of mental images as correct/incorrect when the corresponding letter was presented auditorily; iii) third, The subtest “dictation” from the German test for word production (Wortproduktionsprüfung; Blanken, Döppler, & Schlenck, 1999) served as a diagnostic instrument for pre-, post- and follow-up testing eight months later; iv) finally, writing performance during the functional magnetic resonance imaging (fMRI) sessions recorded via video camera was analysed off-line by two speech-and-language therapists with respect to correctness, response (writing) time (RT) and its standard deviation across items (SD). Response time measurement started after stimulus presentation was completed. We want to draw the readers’ attention to the fact that variations in the measurement between the raters were directly compared per item and in case of differences > 300ms measurement was repeated. In cases of differences < 300ms response latencies were averaged. Thus, this analysis does not provide perfectly reliable writing time. For statistical analyses SPSS18.0 software (IBM, Armonk, NY, USA) was used.

2.5. Imaging procedure and experimental paradigm

Examination of the patient’s writing performance after the intensive training was conducted using a combined behavioural (writing) and fMRI paradigm. Stimulus material consisted of 10 letters (A,C,H,I,L,O,S,T,W and Y), the equivalent trained mental image expressions (“tent”, “half-moon”, “parallel bars”, “stick”, “corner of a house”, “hole”, “snake”, “table”, “wave” and “cocktail glass”) and 10 numbers (0 - 9) presented in a block design. These letters/mental image expressions were selected for the fMRI experiment because they were best solved in the intermediate tests, and thus, potential noise can be avoided in brain activation due to insufficient mastery. The 10 letters (A, C…Y) and numbers (0, 1…9) were associated (0-A, 1-C … 9-Y), while maintaining the relative ordinal position in either sequence. Writing performance was assessed twice. All stimuli were presented both auditorily and visually, in that order, except for the mental image expressions, which did not exist in pictorial format, either in a writing-to-dictation task or a copying task. Therefore, each task (resp. session) was divided into two blocks of ten stimuli in randomized order (trials lasting 6s each) each by a 15s off-condition. Each block lasted 135s. Afterwards a 21s pause was introduced by presenting a black screen.

The stimuli were presented via headphones or with MR-compatible video goggles (Resonance Technology, Inc. Northridge, CA, USA) designed to meet MR requirements using the Presentation software package (Neurobehavioral Systems, Inc., Albany, CA, USA). A scanner-compatible, head-coil mounted video camera was employed to analyse the patient's writing process during the fMRI experiment. Foam pads were used for immobilisation of the patient's head. Before fMRI scanning the patient was familiarised with the task, task sequences, and scanning procedure using a separate practise item set. Each task instruction was repeated via headphones before starting a new session. The patient was asked to write down the items with the tip of his left non-paretic index finger on a soft foam pad placed on his left thigh. To minimise body movements during writing, CU was instructed to move his hand as slightly as possible and to execute writing only on the pad. This writing task was performed without visual control for the patient due to his supine position in the scanner. The patient was scanned twice, directly after the training and in a follow-up testing eight months later to control for short and long-term training effects.

2.6. Image acquisition parameters and analyses

A high-resolution T1-weighted anatomical scan was acquired with a 3T Siemens Magnetom TrioTim MRI system (Siemens AG; Erlangen, Germany) equipped with a 12–channel head matrix coil (TR = 19 s, matrix = 256 x 256 mm, 190 slices, voxel size = 1 × 1 × 1 mm3; FOV = 256 mm, TE = 4.9 ms; flip angle = 25°). The anatomical scan was performed at the end of the experimental session. Functional T2*-weighted images were obtained using gradient-echo Echo planar imaging (EPI; TR = 2400 ms; TE = 30 ms; flip angle = 90°; FOV = 220 mm, 88 x 88 matrix; 42 slices, voxel size = 2.5 × 2.5 × 2.5 mm3, gap = 10%). Total scanning time was approximately 20 minutes. A baseline (rest) condition was accomplished comprising the last four surplus scans, as the patient was lying in the scanner without any specific instruction in the pause lasting 21s.

fMRI data analyses were performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm). In a first step, the fMRI time series was motion corrected and realigned to the patient's first image (Ashburner & Friston, 2003). Imaging data was then normalised into standard stereotaxic Montreal Neurological Institute space (MNI, McGill University, Montreal, Canada). Image data were not smoothed in the spatial or time domain. Images were resampled every 2.5 mm to correct for the gap using 4th degree spline interpolation. The data were high-pass filtered (128s) to remove low-frequency signal drifts and corrected for autocorrelation, assuming an AR(1) process. Brain activity was convolved over all experimental trials with the canonical haemodynamic response function (HRF) and its derivative.

The main goals of the fMRI analysis were to evaluate i) whether CU indeed relied on the retrieval of letter images in order to get access to the grapho-motor patterns of letters during dictation. This issue was addressed by comparing brain activity differences in the contrast writing letters vs. writing mental images (letters > mental images); ii) whether the fMRI data acquired in the follow-up study revealed brain activity diverging from the post-test (follow-up > post-test), and iii) whether modality specific differences in the neural correlates of writing processes occurred after auditory or visual presentation (dictation > copying) for letters and numbers. The SPM Anatomy Toolbox (Eickhoff et al., 2005) available for all published cytoarchitectonic maps from www.fz-juelich.de/ime/spm_anatomy_toolbox was used for the anatomical localisation of effects. Activations were thresholded at an uncorrected p-value of < .001 at the voxel level with a cluster size of k = 10 voxels and were reported when they remained significant following family-wise error correction (FWE) at the cluster-level with pcluster-corr < .05. Complex contrasts were masked inclusively at a threshold of p ≤ .05.

3. Results

3.1. Behavioural data

Behaviourally, the patient benefitted from the writing training. Training progress was observed in the three intermediate tests (T1 – T3) conducted after each week of training. Here, the retrieval of mental image expressions was assessed when the corresponding letter was presented auditorily. As documented in Table S2A in the SOM, the patient not only produced more correct responses in the course of training, but latencies for retrieving the mental images during assessments became shorter as well. From T1 to T2 the number of test items was increased from 15 to all 26 items (umlauts were not assessed), and thus, the test became more demanding. However, training performance improved significantly, as revealed by the exact sign test ($p < 0.001$, one-tailed). Figure S3 shows the patient's performance in writing words and pseudo-words at the end of therapy.

Moreover, significant improvements were apparent in the pre- (0/60 correct trials) and post- (32/60 correct trials) test (sign test, $p < 0.0001$, one-tailed), as well as for follow-up testing eight months after therapy (52/60 correct trials, $p < 0.0001$, one-tailed), assessed with the German test for word production. Comparison of post-test and follow-up revealed that all 32 correct trials were again correctly solved during follow-up. This result accentuates the stability of CU's writing performance (Figure S4 in the supporting online material shows a postal card written by the patient from holidays in Denmark about 6 months after therapy). Crucially, in about half of the cases in the follow-up testing the patient did not need to name mental images to be able to solve the task.

The behavioural data assessed during the fMRI measurement revealed significantly higher response latencies for auditory than for visual presentation (mean difference of RTs for numbers: 234ms and letters: 777ms) of the mental image expressions (Wilcoxon signed ranks test, exact $p < 0.005$). Additionally, the data indicated a modality specific difference in the processing of letters and numbers, with writing letters to dictation being specifically impaired. Significantly longer writing times (exact Mann-Whitney-Test, $p < .05$) for letters as compared to numbers were found in the fMRI dictation task. Accuracy was almost the same for both tasks. An overview of the behavioural results is provided in Table S2B.

3.2. Imaging data: Data analysis was based on all trials.

Writing to dictation vs. copying of numbers: Writing numbers to dictation was associated with more activation around the intraparietal sulcus (IPS), bilaterally, as well as precentral peri-lesionally in the vicinity of Exner's area (Figure 3A). Furthermore, activation differences were found in the right precuneus, bilateral angular and supramarginal gyri, left retrosplenial cortex, and bilateral hippocampus. Further activation differences were observed in bilateral temporal areas, right insula and thalamus, bilateral postcentral and middle frontal gyrus, right inferior frontal gyrus and in the occipital lobe (for detailed results see Table S3 in the SOM).

Writing to dictation vs. copying of letters: For letters the comparison between writing to dictation and copying revealed bilateral activation differences in the precuneus as well as peri-lesionally in the left middle frontal, precentral and inferior frontal gyrus (Figure 3B). Additionally, temporal activation differences were observed in the bilateral superior temporal gyrus, left middle temporal gyrus and right medial temporal pole, as well as in left insula, bilateral basal ganglia including putamen, thalamus and nucleus caudatus, left hippocampus and retrosplenial cortex, right parahippocampal gyrus, the supplementary motor area and the occipital lobe (cf. Table S4).

Writing to dictation, letters vs. mental object image labels after the training: Stronger peri-lesional activation was observed for letters in the vicinity of Exner`s area (precentral gyrus [Area 6], middle frontal gyrus, inferior frontal gyrus [Area 44]) as well as mental imagery related precuneus activation for comparing writing letters to dictation and writing mental image expressions to dictation (Figure 3C). Further clusters of more strongly activated voxels were found in the supplementary motor area and the anterior cingulate cortex (ACC), left retrosplenial cortex and hippocampus, right parahippocampal gyrus, bilateral inferior frontal cortices and thalamus, right caudate nucleus, superior, middle, and medial temporal gyrus as well as in occipital cortex (Table 2).

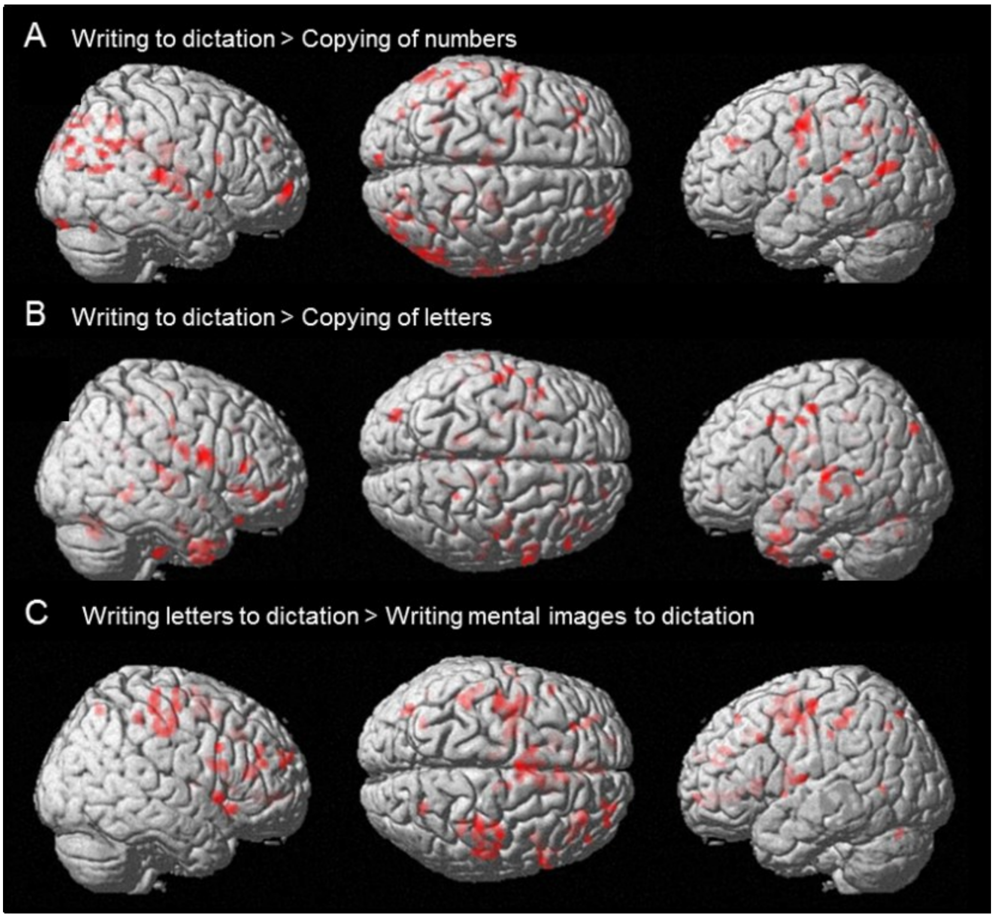


Fig. 3: : Panel A) depicts stronger fMRI activation when comparing writing to dictation of numbers to the copying of numbers at the beginning of the training, Panel B) shows activation associated with writing to dictation of letters compared to copying of letters at the beginning of the training. Activation in the IPS is observed only in the number dictation task; C) Comparison of writing letters to dictation vs. writing mental object image labels to dictation post training: stronger perilesional activation observed for letters in the vicinity of Exner`s area, as well as mental imagery related precuneus activation indicating association of letters with corresponding mental images (all at $p_{cluster\text{-}corr} < .05$, cluster size of k = 10 voxels). Activation threshold at an uncorrected p-value of < .001 reported when significant following family-wise error correction (FWE) at the cluster-level with p < .05. Complex contrasts were masked inclusively at a threshold of p ≤ .05, a minimal cluster size of k = 10 voxels.

Table 4: Cortical regions more strongly activated for dictation of letters vs. mental object image labels after the training.

Contrast	Brain Area	MNI (x, y, z)			Cluster size	z
Writing to dictation:	LH precuneus	-2	-64	28	11	3.55
letters >	RH superior parietal lobule	12	-36	44	12	3.71
mental images	LH precentral gyrus (Area 6)	-50	6	44	11	4.33
	RH precentral gyrus	66	4	22	32	4.26
	RH precentral gyrus (Area 3b)	52	-8	22	70	4.32
	RH supplementary motor area (Area 6)	12	-20	60	15	4.02
	RH anterior cingulate cortex	2	46	20	21	3.90
	LH retrosplenial cortex	-4	-46	30	16	4.30
	LH hippocampus	-28	-16	-22	144	5.12
	RH parahippocampal gyrus	26	-10	-26	45	4.62
	LH inferior frontal gyrus (Area 44)	-40	6	24	16	4.05
	RH inferior frontal gyrus (Area 44)	56	8	16	78	5.17
	RH inferior frontal gyrus (Area 45)	54	30	12	25	4.22
	LH thalamus	-16	-16	14	19	4.22
	RH thalamus	20	-22	0	37	3.87
	RH caudate nucleus	18	16	10	12	3.72
	RH superior frontal gyrus	22	56	0	13	3.72
	LH middle frontal gyrus	-44	14	42	18	3.82
	LH middle frontal gyrus	-30	10	46	14	3.75
	RH inferior frontal gyrus (p. orb)	30	28	-20	17	4.47
	RH inferior frontal gyrus (p. orb)	40	38	-6	89	4.34
	LH superior temporal gyrus	-68	-24	6	49	5.09
	RH superior temporal gyrus	54	-24	10	21	4.70
	LH middle temporal gyrus	-40	6	-32	40	4.18
	LH middle temporal gyrus	-50	-38	-2	41	5.57
	LH middle temporal gyrus	-66	-46	8	12	3,79
	LH medial temporal pole	-22	8	-40	65	3.93
	RH medial temporal pole	42	8	-38	91	5.06
	LH insula	-34	2	10	27	4.74
	LH insula	-36	2	-6	10	3.72
	RH fusiform gyrus	22	8	-46	110	4.29
	RH middle orbital gyrus	22	56	-12	15	5.07
	LH postcentral gyrus	-52	-14	50	27	4.31
	LH postcentral gyrus	-58	-8	42	15	4.23
	LH middle occipital gyrus	-30	-82	38	29	4.20

$p_{cluster\text{-}corr}$ < .05 (k = 10 voxels); MNI: Montreal Neurological Institute coordinates

<u>Writing letters to dictation, post training vs. rest condition:</u> Writing letters immediately after the training was significantly associated with bilateral activation of the precuneus, peri-lesional activation in the left precentral and superior frontal gyrus. Further more strongly activated voxels were observed in bilateral IPS, bilateral PSPL, bilateral supramarginal gyrus, right angular gyrus, left retrosplenial cortex and hippocampus, right parahippocampal gyrus (Figure 4A), bilateral superior and middle temporal gyrus as well as the right inferior temporal gyrus and the left medial temporal pole. Also stronger left insula, bilateral thalamus, right pallidum and putamen, right

caudate nucleus as well as bilateral supplementary motor area and occipital activation was observed (cf. Table S5).

Writing letters to dictation, follow-up vs. rest condition: Eight months after training, writing letters to dictation was still associated with peri-lesional activation in the left precentral, superior frontal, middle frontal and inferior frontal gyrus (Area 45, Figure 4B) as well as in the left precuneus. However, also activation in bilateral angular gyrus (PGa and PGp) and bilateral supramarginal gyrus (PFm) was observed as well as in the bilateral IPS (hIP1 and hIP3) and PSPL. Moreover, left hippocampus as well as right parahippocampal gyrus were found active. Additionally, again various clusters of activated voxels in the temporal cortex were observed comprising the bilateral superior, middle, and inferior temporal gyrus as well as right insula, bilateral thalamus, bilateral supplementary motor area and right ACC and occipital areas (Table S6).

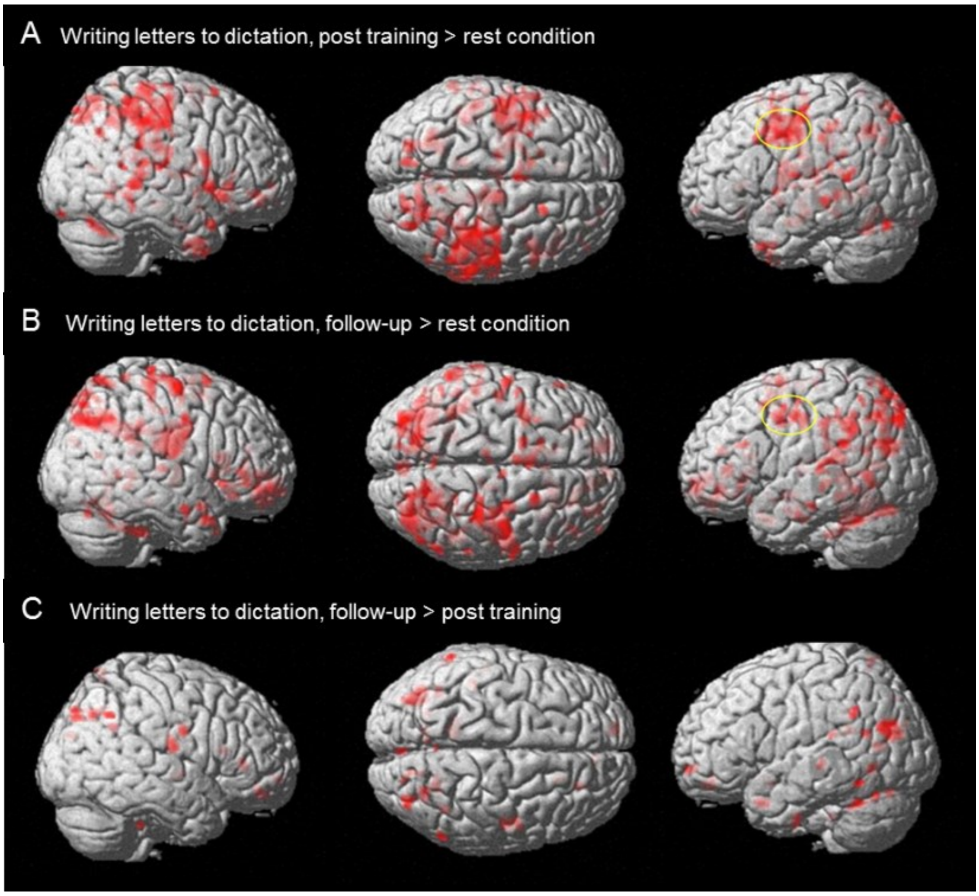


Fig. 4: Comparison of writing letters to dictation vs. rest across the training period; A) immediately after training strong perilesional activation in the vicinity of Exner`s area as well as mental imagery related precuneus activation; B) eight months later at follow-up; again, perilesional activation for letters in the vicinity of Exner`s area as well as mental imagery related precuneus activation; C) comparison of follow-up > post training (masked inclusively with follow up); primarily differential activation in the left angular gyrus indicating generalisation processes. (all at $p_{cluster\text{-}corr} < .05$, cluster size of k = 10 voxels).

Writing letters to dictation, follow-up vs. immediately after therapy: The direct contrast of writing letters to dictation during follow-up minus the examination immediately after training (Figure 4C) revealed more activation primarily in the bilateral angular gyrus (PGa and PGp) as well as the right supramarginal gyrus (PFm). Furthermore, activation differences in the left superior, middle and inferior temporal gyrus, the temporal pole, left retrosplenial cortex, left hippocampus, right thalamus, right caudate nucleus, left fusiform and lingual gyrus were observed as well as in occipital areas (Table 3).

Contrasting post-training vs. follow-up, no suprathreshold activation was observed.

Table 5: Cortical regions more strongly activated in the follow up study compared to immediately after the training.

Contrast	Brain Area	MNI (x, y, z)			Cluster size	z
Letters:	LH angular gyrus (Area 39, PGa)	-57	-56	31	15	3.15
Follow up >	LH angular gyrus (Area 39, PGp)	-36	-70	22	162	4.57
post training	RH angular gyrus (Area 39, PGa)	40	-54	30	15	3.46
	RH angular gyrus (Area 39, PGp)	38	-82	42	31	3.83
	RH supramarginal gyrus (BA 40, PFm)	62	-54	38	13	3.50
	LH superior temporal gyrus	-44	-36	18	14	3.27
	LH middle temporal gyrus	-52	-28	-2	10	3.42
	LH inferior temporal gyrus	-60	-54	-16	12	4.07
	LH temporal pole	-26	14	-28	35	3.92
	LH retrosplenial cortex	-10	-48	34	16	3.48
	LH hippocampus	-26	-11	-24	13	3.43
	RH thalamus	16	-8	-2	19	3.44
	RH caudate nucleus	20	22	14	13	3.24
	LH fusiform gyrus	-34	-12	-38	20	3.45
	LH lingual gyrus	-26	-62	-4	33	5.27
	RH middle orbital gyrus	6	46	-14	50	4.06
	LH superior orbital gyrus	-14	60	-6	30	3.61
	RH postcentral gyrus	50	-6	28	29	3.81
	LH cuneus (Area 18)	2	-84	26	11	3.04
	RH cuneus	16	-76	42	12	2.83
	RH superior occipital gyrus	28	-68	38	58	4.32

$p_{cluster\text{-}corr}$ < .05 (k = 10 voxels); MNI: Montreal Neurological Institute coordinates.

4. Discussion

In reporting the case of CU, we described an individually tailored, highly effective writing training for an aphasic patient with peripheral agraphia. The aim of the training was to enable CU to write manually again. Therefore, additional semantic information - similar to the numerical magnitude information automatically activated when processing Arabic digits - was associated to letters by means of mental images. After the training the patient indeed associated specific mental images with letters, which he then recalled in order to retrieve the grapho-motor patterns of letters. These effects were stable even at follow-up eight months after training, although CU did no longer seem to have to rely on the mental images entirely.

fMRI data corroborated these findings at the neural level. We found primarily peri-lesional activation adjacent to Exner`s area for all writing conditions directly after the training and in the follow-up examination. Whereas number processing yielded more activation in the intraparietal sulcus (IPS) and the angular gyrus (AG), increased fMRI signal was found for writing letters to dictation in the precuneus, an area assumed to be involved in visual mental imagery (Cavanna & Trimble, 2006; Ganis et al., 2004). This is in line with our hypothesis that the patient used mental imagery in order to gain access to the grapho-motor pattern of letters. As expected, the increase of precuneus activity was larger in the letter writing condition than when the patient was presented the mental image expression auditorily. While writing letters to dictation still resulted in some precuneus activation during follow-up, writing letters to dictation at follow-up compared to immediately after therapy revealed no differential precuneus activations. Instead, primarily more activation in the bilateral angular gyrus was found. This is in line with the literature reporting increased activation in the angular gyrus following intensive training processes, probably due to retrieval from memory (Delazer et al., 2005; Grabner et al., 2009). Therefore, we assume the change of activation pattern to be directly linked to retrieval processes from long-time memory through consolidation processes involving meaningful association and rehearsal. This assumption is further corroborated by the simultaneous involvement of the left hippocampus, a region typically subserving processes involved in coordinating long-term memory retrieval (e.g., Montaldi & Mayes, 2010).

4.1. Specific writing training by means of mental images analogous to number inherent semantic information

The relatively good performance in writing numbers as compared to letters can be explained by the additional semantic information conveyed by numbers, which is activated automatically (Eger et al., 2003; Klein et al., 2010), probably facilitating their processing. With respect to kinematics, Lochy, Zoppoth, Domahs and Delazer (2003) showed that digits require fewer changes in direction during hand writing than letters, which may also contribute to the likely integrity of digits in comparison to letters after stroke. Additionally, there are fewer digits (0 to 9) than letters.

However, the current results show that three weeks of intensive writing training yielded significant improvement in patient CU's writing abilities. Most remarkably, writing performance assessed with the German test for word production remained not only stable eight months after the training but even improved at follow-up with a reduced need to apply the mental images strategy. CU was used to name aloud the mental image before writing the corresponding letter, as he was taught during the training; but he did no longer employ this recommended strategy. Months later, the mental images were no longer necessary for CU to solve the task, which is also documented by reduced response times and changes in neuroimaging findings, suggesting that processing got more and more automatic in the end.

So far, a few studies have reported cases of peripheral agraphia with (Toghi et al., 1995) or without associated acalculia (Anderson et al., 1990; Delazer et al., 2002; Keller & Meister, 2014; Starrfelt, 2007) or indicated the possible impact of distinct lesions in Exner`s area (Exner, 1881; Anderson et al., 1990). Only Delazer and colleagues (2002) reported that intensive writing practice as well as techniques such as the imagination of letters or the mental association of letter shapes with real objects, as employed during their training study, improves writing performance. As the latter approach resembles our training by means of mental images, our study strongly supports these findings. Furthermore, the present study provides first evidence that it is possible to enhance writing performance in a case of peripheral agraphia even years after stroke, and thus, our study is an argument for the success of systematic and symptom-oriented language treatment even in chronic aphasia.

4.2. Hand-writing related neuroimaging findings

The following discussion takes into account how the writing tasks were conducted in the MR scanner. Because the patient wrote in rather large movements with his left index finger on his left thigh, he actually had to retrieve abstract motor programmes rather than fine writing movement details, when using a pen on a sheet of paper.

For all writing tasks we found peri-lesional activation in the vicinity of Exner's area (BA 6), mainly covering parts of the precentral, inferior, middle and superior frontal gyrus. The middle frontal gyrus was found to be primarily specific for writing, whereas areas in the precentral gyrus are involved in preparing and executing hand movements (Planton et al., 2013, for a meta-analysis). The inferior frontal gyrus (Area 44), which is also supposed to be involved in language processing (comprising "Broca's area" together with Area 45, see Broca, 1861; Grodzinsky & Santi, 2008), was found to be active even during imagery of hand movements as well (Rizzolatti, Fogassi, & Gallese, 2002). These frontal lobe areas were stronger activated in the contrast writing letters to dictation vs. rest condition directly after training as compared to follow-up. We suggest that less writing specific activation at follow-up reflects reduced effort in retrieving grapho-motor patterns and, thus, indicates stabilisation of the writing process.

Additionally, general limb movement related activity was found in the supplementary motor area (SMA), thalamus and putamen as well as in the insula. Increased activity in the SMA and the basal ganglia can account for articulatory and phonological processing, as the patient repeated the target letter and the corresponding mental image verbally. However, CU was engaged in active movements as well as in the repetition of words in a phonological cortico-subcortical loop. Thus, we assume that this observed activation most probably reflected activation related to both processes.

4.3. Writing-training related neuroimaging findings

fMRI results clearly corroborated the presence of two distinct networks for writing numbers and letters to dictation. Whereas ventral and dorsal pathways for speech production and comprehension have been identified and framed in the “dual loop model”, also referred to as “dual stream model”, in recent years (DeWitt & Rauschecker, 2012; Friederici, 2009; Gow, Keller, Eskandar, Meng, & Cash, 2009; Hickok & Poeppel, 2004, 2007; Parker et al., 2005; Saur et al., 2008; 2010; Turken & Dronkers, 2011; Ueno, Saito, Rogers, & Lambon Ralph, 2011, Weiller et al., 2011), to the best of our knowledge the organisation of the connections within the writing network has not been investigated separately so far. Nevertheless, using tracking knowledge about the dual loop model for various domains (Rijntjes, Weiller, Bormann, & Musso, 2012) as well as tracking data for number processing, the difference between writing numbers to dictation and writing letters to dictation might be explained as follows:

When listening to the auditory presentation of number words or letters, in a first step primary and secondary auditory cortex is involved, reflected by temporal activation. In particular, the superior temporal gyrus was shown to be responsive to auditorily presented speech and speech perception (Buchsbaum, Hickock, & Humphries, 2001). CU repeated both numbers and letters verbally, probably inducing not only further temporal activation but also activation of the phonological loop (including basal ganglia such as thalamus, caudate nucleus, etc.). According to the cortical network model of written language production by Scarone and colleagues (2009) the phonological loop might be also included in writing processes. Furthermore, the patient tried to retrieve information from long-term memory during recollection processes, as reflected by the hippocampal (supported by parahippocampal regions) activation (Baddeley, 1966; Rugg & Yonelinas, 2003, for a review). From this processing step onward, the typical processing streams before and during the training might be distinguished as follows:

(i) In case of numerical processing, numerical magnitude is automatically retrieved, reflected by IPS activation (Eger et al., 2003; Klein et al., 2010), which was found bilaterally in both number writing tasks. This retrieval requires connections between temporal auditory cortex and the intraparietal cortex, probably subserved by the inferior longitudinal fascicule (ILF; e.g., Seghier, 2013). In healthy participants, the additional information about numerical magnitude is then relayed both ventrally (via the external/extreme (EC/EmC) capsula system) and dorsally (via the superior longitudinal fascicule, SLF II/III) to Area 44 and Exner's area, situated in the direct superior vicinity to Area 44 (see Klein et al., 2013, Klein et al., 2015). However, the ventral connection terminates considerably more inferior in the frontal cortex than the SLF (see also Seghier, 2013; Caspers, Zilles, Eickhoff, Schleicher, Mohlberg, & Amunts, 2013). Since the lesion is located in this inferior part of the frontal location, numerical magnitude information may most probably have been transmitted via the dorsal connection in CU. This may explain CU's ability to write numbers to dictation by relying on number magnitude information, which is conveyed dorsally. The strong perilesional activation in the upper part of Exner's area may account for the retrieval of grapho-motor patterns of digits. Additionally, angular gyrus and retrosplenial cortex activation were found for writing numbers. The AG has been frequently suggested to subserve and/or mediate arithmetic fact retrieval (Dehaene et al., 2003; Grabner et al., 2009, for a review), while the retro-splenial cortex has repeatedly been associated with the recognition of familiar objects, faces, voices, or procedures (Shah et al., 2001; Vann, Aggleton, & Maguire, 2009, for a review).

(ii) In case of writing letters to dictation, again temporal areas and basal ganglia were activated via the phonological loop. However, between temporal cortex, basal ganglia and Exner's area, mostly ventral connections are observed (Weiller et al., 2011), as well as dorsal connections from the thalamus. Both pathways of the dual loop model of language (EC/EmC, parts of arcuate fascicle) are disconnected. Furthermore, these pathways enter Exner's area from inferior – the brain tissue which is impaired. Therefore, before the training, it was not possible for patient CU to retrieve the grapho-motor patterns of letters. During training additional semantic information was attached to the letters. This semantic information was retrieved by mental imagery, presumably reflected by additional precuneus activation. Connections pass from both the temporal cortex (Seghier, 2013) as well as the thalamus (via dorsal parts of the corona radiata) to the precuneus. We suppose, that once additional activation is retrieved from the precuneus, the connectivity is highly similar to the one suggested above for numbers (via the SLF II/III to the upper [spared] part of Exner's area). Subsequently, the grapho-motor patterns of letters could be accessed by CU.

Further intensive writing training might most likely result in a similar shift of activation into the AG as previously observed for drill learning of mathematical tasks (e.g., Delazer et al., 2003; Ischebeck et al., 2006).

More particularly, after the intensive writing training specific activation was found in the left precuneus for writing letters to dictation as compared to copying letters, and more interesting, when compared to the auditorily presented mental image expressions. The left precuneus and its co-activated right homologue area have been shown to support visual-spatial mental imagery during retrieval from memory (Ganis et al., 2004; Cavanna & Trimble, 2006) and are provisionally indicative of specific recall of mental images. Even though the precuneus is assumed to be included in the default network, increased brain activity found in the precuneus for writing letters to dictation is not associated with default network recruitment, since in the fMRI analysis all complex contrasts were masked inclusively at a threshold of $p \leq .05$. Thus an artefact differential activation due to less activation of this brain area compared to the rest condition can be ruled out. Increased precuneus activation might be also linked to increased task demand, which is the case for writing letters to dictation as compared to writing letters after presentation of the mental image expressions. However, precuneus activation is reduced in the follow-up, although task demands can be assumed to be the same as compared to after therapy. This finding suggests that precuneus activation might be mental image specific.

When comparing brain activation at follow-up vs. immediately post training, indeed, much larger activation of the bilateral angular gyrus (AG) was observed, which was significantly *activated*, not deactivated to a lesser degree than before. This activation pattern may be explained as follows: apart from its role in learning processes and fact retrieval (Delazer et al., 2005; Grabner et al., 2009), the AG has been proposed to be engaged in mapping processes (Ansari, 2008) and in the integration of cross-modal information. We suggest that increased AG activity accounts for long-term learning effects of the writing training because, on the one hand, fMRI data revealed strongly reduced precuneus activation, when comparing brain activation in the follow-up with the rest condition. This result is in line with the behavioural data (decreased response times and reduced mental image utilisation), showing that CU did not rely on the mental images anymore for writing the required letters. On the other hand, successful mapping of a letter with its corresponding mental image, stored in the AG and accessed by the hippocampal formation, indicates long-lasting training effects.

Additionally, we found increased activation in the fusiform gyrus. This activity is related to the recognition of various types of visual stimuli such as letter strings of language-specific orthographic structures (Binder et al., 2006; Dehaene et al., 2005), but recently also to more high-level integration of visual processing (Price, 2012). In the present task of writing to dictation the fusiform gyrus may have rather been co-activated, since the task was to retrieve the graphical form of letters.

4.4. CU and his contribution to the concept of Exner's area

Patient CU suffered from a brain lesion mainly affecting triangular and opercular parts of Broca's Area (Area 44/45) extending superior into Exner`s area (BA 6). We attribute his pronounced writing impairment for words and even letters specifically to the "strategic" localisation of this lesion: i) First, following the models of written language processing, central processes seemed unimpaired, because CU was able to orally spell and type auditorily presented words and pseudo-words correctly on a computer keyboard. Indeed, he failed to write them manually, showing that Exner's area might subserve specific grapho-motor planning for handwriting, but not typing. This finding contradicts recent results of a meta-analysis of Purcell (2011, p.13) who argued that "it [Exner's area] may play a role in the conversion of graphemic representations to motor commands regardless of whether the word is handwritten or typed". Crucially, since CU was able to type auditorily presented pseudo-words correctly on the computer, we cannot confirm Exner's role in phoneme-grapheme-conversion as suggested by Omura, Tsukamoto, Kotani, Ohgami, and Yoshikawa (2004) and recently by Keller and Meister (2014). The acute stroke patients reported in the latter study were able to write letter shapes correctly, suggesting that they had allographic knowledge. The type of errors committed by the latter patients, for instance, addition or omission of graphemes, rather indicates a deficit in orthographic working memory, where names and sequences of letters are retrieved for actual writing. Such deficits caused by lesions in the left inferior fontal gyrus have been frequently related to deficits in orthographic working memory (Purcell, Turkeltaub et al., 2011, for a meta-analysis). However, the patients` writing deficits shown by Keller und Meister (2014) seem to affect central but not peripheral writing processes. ii) Second, letter identification as well as copying were preserved. In conclusion, impairment at the stage of visual graphic representation of letters can be excluded to be the cause of his writing problems (Ellis, 1988; Delazer et al., 2002). iii) Third and much more crucial, however, is the fact that transposing words and pseudo-words was almost unimpaired, whereas spontaneous writing was severely affected before therapy. In the transposition task, visual forms of either lower or upper case letters were presented to the patient. Once the patient saw a letter, and thus, had access to the allographic information of this specific majuscule or minuscule, he was able to write manually the corresponding lower resp. upper case letter. Thus, processing at the allograph level, where graphic styles (e.g., lower versus upper case, italics versus block letters) are stored before selected allographs are specified at the level of graphomotor planning (Roux et al., 2009), seems to function properly. Sufficient mastery in the transposition task before the writing training, however, might most likely be explained by disrupted functional connections to the allographic level. Due to this disconnection, the patient cannot access more peripheral representations from abstract graphemes. With respect to the functional role of Exner's area we provisionally hypothesise that multiple components of peripheral writing processes - abstract grapheme level, allograph level and graphomotor planning - are connected by this area of the brain before pre-motor induced execution of handwriting. With this assumption we are in line with the current definition of the role of Exner's area for handwriting by Planton and colleagues (2013).

iv) Finally, we suggest that Exner's area does not selectively subserve letter and word writing because increased Exner's area activity was found during writing numbers in Arabic digits to dictation as well. Number related pathways, at least the dorsal pathway connecting IPS and Exner's area, were not affected by the lesion. Therefore, the patient could rely on magnitude related sematic information about numbers, when retrieving grapho-motor information for digits; a process which Exner's area is probably also involved in. Thus, we suggest that this area of the brain might rather be insensitive to the input modality and be involved in processing information related to, both, letter and digits. We admit that this hypothesis deserves further investigations with healthy controls to be substantiated. However, both letters and digits are complex scripts, which have to be retrieved grapho-motorically. While digits always provide additional semantic information for processing due to automatic access to numerical magnitude information, which is normally relayed to Exner's area via ventral and dorsal pathways, letters in isolation do not contain additional semantic information. A lesion in Exner's area, and thus, impairment in the fronto-temporal interaction, can lead to substantial writing impairments.

5. Conclusion

The present study demonstrates that it is possible to enhance hand writing performance in a chronic aphasic patient suffering from severe peripheral agraphia even years post onset. Even though isolated cases of peripheral agraphia are rather rare or probably masked by other aphasic symptoms, various studies are on record indicating that this topic is still a matter of debate (Anderson, Damsion & Damasio, 1990; Delazer et al., 2002; Keller & Meister, 2014; Toghi et al., 1995).

Recently, Marangolo and colleagues (2011) combined intensive language training with anodal tDCS over the left inferior frontal gyrus (Broca's area) in aphasic patients with speech apraxia. The authors reported, apart from long-lasting improvement of articulatory gestures, unexpected improvements in writing, which was not practised during the training. They hypothesised that stimulation of Broca's area could have co-activated premotor brain regions, specifically Exner's area, and thus facilitated writing processes. These results demonstrate that new techniques such as tDCS might enhance recovery of cognitive functions following stroke, and thus, may be suitable to supplement speech and language therapy (see also Baker, Rorden & Fridriksson, 2010; Harris-Love & Cohen, 2006, for a review), in particular when fiber tracks connecting relevant cortical processing areas are not impaired.

In summary, the present results contribute to our understanding of the functions subserved by Exner's writing center during the translation from more abstract graphomotor patterns to the execution of hand movements. Most importantly, they suggest an unconventional but very effective, easy to administer, and reliable approach to the treatment of peripheral agraphia.

Acknowledgements

This research was supported by the German Research Foundation (DFG) to Klaus Willmes (Wl 1804/6-1) and Hans-Otto Karnath (KA 1258/11-1). Stefanie Jung was supported by the Science Campus Tuebingen (Cluster 1, TP8 in the third phase) providing funding to Juergen Heller. Elise Klein was supported by the Leibniz-Competition Fund (SAW-2014-IWM-4) as well as the Margarete von-Wrangell Fellowship of the Ministery Science, Research and Arts Baden-Wuerttemberg and the European Social Fonds (ESF). We are indebted to Georg Eder from the "Brain Imaging Facility of the Interdisciplinary Centre for Clinical Research in the Faculty of Medicine at the RWTH Aachen University" for his precious help in acquiring the MRI/fMRI data. Furthermore, we wish to express our gratitude to Annette Weimer for conducting the writing training and the anonymous reviewers for very helpful comments on an earlier version of the paper.

References

Anderson, S. W., Damasio, A. R., & Damasio, H. (1990). Troubled letters but not numbers: Domain specific cognitive impairments following focal damage in frontal cortex. Brain, 113(3), 749-766.

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. Nature Reviews Neuroscience, 9(4), 278-291.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. Neuroimage, 54(3), 2382-2393.

Ashburner, J., & Friston, K. J. (2003). Spatial normalization using basis functions. Human Brain Function, 655-672.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *The Quarterly Journal of Experimental Psychology*, *18*(4), 302-309.

Baker, J. M., Rorden, C., & Fridriksson, J. (2010). Using transcranial direct-current stimulation to treat stroke patients with aphasia. *Stroke*, *41*(6), 1229-1236.

Beeson, P., Rapcsak, S., Plante, E., Chargualaf, J., Chung, A., Johnson, S., & Trouard, T. (2003). The neural substrates of writing: A functional magnetic resonance imaging study. *Aphasiology*, *17*(6-7), 647-665.

Beeson, P. M., & Rapczak, S. Z. (2004). Agraphia. In R. D. Kent (Ed.), The MIT encyclopedia of communication disorders (pp. 233–236). Cambridge, MA: MIT Press.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage*, *33*(2), 739-748.

Binkofski, F., & Buccino, G. (2004). Motor functions of the Broca's region. *Brain and Language*, *89*(2), 362-369.

Blanken, G., Döppler, R., Schlenck, K. J., & Bautz, M. (1999). *Wortproduktionsprüfung*. Hofheim: NAT-Verlag.

Broca, P. (1861). Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bull Soc Anthropol*, *2*, 235-238.

Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, *25*(5), 663-678.

Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(3), 564-583.

Caspers, J., Zilles, K., Eickhoff, S. B., Schleicher, A., Mohlberg, H., & Amunts, K. (2013). Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. *Brain Structure and Function*, *218*(2), 511-526.

Crary, M. A., & Heilman, K. M. (1988). Letter imagery deficits in a case of pure apraxic agraphia. *Brain and Language*, *34*(1), 147-156.

Crosson, B., Benefield, H., Cato, M. A., Sadek, J. R., Moore, A. B., Wierenga, C. E., ... & Briggs, R. W. (2003). Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*, *9*(07), 1061-1077.

De Bastiani, P., & Barry, C. (1989). A cognitive analysis of an acquired dysgraphic patient with an “allographic” writing disorder. *Cognitive Neuropsychology*, *6*(1), 25-41.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*(1), 83-120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. Cortex, *33*(2), 219-250.

Dehaene, S., Piazza, M., Pinel, P., Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487-506.

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic—an fMRI study. *Cognitive Brain Research, 18*(1), 76-88.

Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C. M., ... & Felber, S. (2005). Learning by strategies and learning by drill—Evidence from an fMRI study. *Neuroimage*, *25*(3), 838-849.

Delazer, M., Girelli, L., Graná, A. & Domahs, F. (2003). Number processing and calculation - Normative data from healthy adults. *The Clinical Neuropsychologist*, *17*(3), 331-350.

Delazer, M., Lochy, A., Jenner, C., Domahs, F., & Benke, T. (2002). When writing 0 (zero) is easier than writing O (o): a neuropsychological case study of agraphia. *Neuropsychologia*, *40*(12), 2167-2177.

De Renzi, E., Motti, F., & Nichelli, P. (1980). Imitating gestures: A quantitative approach to ideomotor apraxia. *Archives of Neurology*, *37*(1), 6.

DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences, 109*(8), E505-E514.

Eger, E., Sterzer, P., Russ, M.O., Giraud, A., Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, *37*, 1 – 20.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, *25*(4), 1325-1335.

Ellis, A. W. (1988). Normal writing processes and peripheral acquired dysgraphias. *Language and Cognitive Processes*, *3*(2), 99-127.

Exner, S. (1881). *Untersuchungen über die Localisation der Functionen in der Grosshirnrinde des Menschen...* Wilhelm Braumüller.

Fels, M., & Geissner, E. (1996). *Neglect-Test (NET).* Göttingen: Hogrefe.

Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences, 13*, 175–181.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research, 20*(2), 226-241.

Gow, D. W., Keller, C. J., Eskandar, E., Meng, N., & Cash, S. S. (2009). Parallel versus serial processing dependencies in the perisylvian speech network: A Granger analysis of intracranial EEG data. *Brain and Language, 110*, 43–48.

Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., & Neuper, C. (2009). Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. *Human Brain Mapping*, *30*(9), 2936-2952.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*(12), 474-480.

Harris-Love, M. L., & Cohen, L. G. (2006). Noninvasive cortical stimulation in neurorehabilitation: a review. *Archives of Physical Medicine and Rehabilitation*, *87*(12), 84-93.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67–99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*, 393–402.

Hillis, A. E., & Caramazza, A. (1989). The graphemic buffer and attentional mechanisms. *Brain and Language*, *36*(2), 208-235.

Huber, W., Poeck, K., Weniger, D., & Willmes, K. *Der Aachener Aphasie Test*, Hogrefe, Göttingen.

Huber, W., Springer, L., & Willmes, K. (1993). Approaches to aphasia therapy in Aachen. In *Aphasia Treatment* (pp. 55-86). Springer US.

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstätter, F., Benke, T., Felber, S., & Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. *Neuroimage*, *30*(4), 1365-1375.

Keller, C., & Meister, I. G. (2014). Agraphia caused by an infarction in Exner's area. *Journal of Clinical Neuroscience*, *21*(1), 172-173.

Klein, E., Moeller, K., Nuerk, H. C., & Willmes, K. (2010). On the neuro-cognitive foundations of basic auditory number processing: an fMRI study. *Behavioral and Brain Functions*, 6(42), 1-13.

Klein, E., Mann, A., Huber, S., Bloechle, J., Willmes, K., Karim, A.A., Nuerk, H.-C., & Moeller, K. (2013). Bilateral bi-cephalic tDCS with two active electrodes of the same polarity modulates bilateral cognitive processes differentially. *PlosONE*, *8*(8): e71607.

Klein, E., Suchan, J., Moeller, K., Karnath, H.-O., Knops, A., Wood, G., Nuerk, H.-C., & Willmes K. (2015). Considering structural connectivity in the triple code model of numerical cognition - Differential connectivity for magnitude processing and arithmetic facts. *Brain Structure & Function*..

Lambert, J., Viader, F., Eustache, F., & Morin, P. (1994). Contribution to peripheral agraphia... a case of post-allographic impairment?. *Cognitive Neuropsychology*, *11*(1), 35-55.

Lambert, J., Giffard, B., Nore, F., de La Sayette, V., Pasquier, F., & Eustache, F. (2007). Central and peripheral agraphia in Alzheimer's disease: from the case of Auguste D. to a cognitive neuropsychology approach. *Cortex*, *43*(7), 935-951.

Lochy, A., Zoppoth, S., Domahs, F., & Delazer, M. (2003). Peripheral processes in writing: Why are numbers so often preserved? A comparative study of letters vs. digits kinematics. *Le Langage et l'homme*, *38*(2), 25-46.

Lubrano, V., Roux, F. E., & Démonet, J. F. (2004). Writing-specific sites in frontal areas: a cortical stimulation study. *Journal of neurosurgery*, *101*(5), 787-798.

Marangolo, P., Marinelli, C. V., Bonifazi, S., Fiori, V., Ceravolo, M. G., Provinciali, L., & Tomaiuolo, F. (2011). Electrical stimulation over the left inferior frontal gyrus (IFG) determines long-term effects in the recovery of speech apraxia in three chronic aphasics. *Behavioural Brain Research*, *225*(2), 498-504.

Menichelli, A., Rapp, B., & Semenza, C. (2008). Allographic agraphia: A case study. *Cortex*, *44*(7), 861-868.

Miller, N., Willmes, K., & De Bleser, R. (2000). The psychometric properties of the English language version of the Aachen Aphasia Test (EAAT). *Aphasiology*, *14*(7), 683-722.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology*, *16*(10-11), 935-979.

Montaldi, D., & Mayes, A. R. (2010). The role of recollection and familiarity in the functional differentiation of the medial temporal lobes. *Hippocampus*, *20*(11), 1291-1314.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Omura, K., Tsukamoto, T., Kotani, Y., Ohgami, Y., & Yoshikawa, K. (2004). Neural correlates of phoneme-to-grapheme conversion. *Neuroreport*, *15*(6), 949-953.

Paivio, A. (1971). *Imagery and verbal processes.* Hillsdale, N.J.: Erlbaum.

Parker, G. J. M., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A. M., Ciccarelli, O. & Lambon Ralph, M. A. (2005). Lateralization of ventral and dorsal auditorylanguage pathways in the human brain. *NeuroImage, 24*, 656–666.

Planton, S., Jucla, M., Roux, F. E., & Démonet, J. F. (2013). The "handwriting brain": a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex*, *49*(10), 2772-2787.

Poeck, K., & Kerschensteiner, M. (1975). Analysis of the sequential motor events in oral apraxia. In *Cerebral localization* (pp. 98-111). Springer Berlin Heidelberg.

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62(2), 816-847.

Purcell, J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, *2*, 239.

Rapcsak, S.Z., Beeson, P.M. (2002). Neuroanatomical correlates of spelling and writing. In A.E. Hillis (Ed.) *The Handbook of Adult Language Disorders*. p. 71-99. New York: Psychology Press.

Rapp, B., & Caramazza, A. (1997). From graphemes to abstract letter shapes: Levels of representation in written spelling. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 1130.

Rijntjes, M., Weiller, C., Bormann, T., & Musso, M. (2012). The dual loop model: its relation to language and other modalities. *Frontiers in evolutionary neuroscience*, *4*.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current Opinion in Neurobiology*, *12*(2), 149-154.

Roux, F. E., & Démonet, J. F. (2013). The "handwriting brain": a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex*, *49*(10), 2772-2787.

Roux, F. E., Draper, L., Köpke, B., & Démonet, J. F. (2010). Who actually read Exner? Returning to the source of the frontal "writing centre" hypothesis. Cortex, *46*(9), 1204-1210.

Roux, F. E., Dufor, O., Giussani, C., Wamain, Y., Draper, L., Longcamp, M., & Démonet, J. F. (2009). The graphemic/motor frontal area Exner's area revisited. *Annals of neurology*, *66*(4), 537-545.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, 7(7), 313-319.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., ... & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, *105*(46), 18035-18040.

Saur, D., Schelter, B., Schness, S., Kratochvil, D., Kupper, H., Kellmeyer, P., et al. (2010). Combining functional and an atomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage, 49*, 3187–3197.

Scarone, P., Gatignol, P., Guillaume, S., Denvil, D., Capelle, L., & Duffau, H. (2009). Agraphia after awake surgery for brain tumor: new insights into the anatomo-functional network of writing. *Surgical neurology*, *72*(3), 223-241.

Schelling D. (1997). *Block-Tapping-Test*. Frankfurt: Swets Test Service GmbH.

Seger, C. A., & Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *The Journal of Neuroscience*, *25*(11), 2941-2951.

Seghier, M. L. (2013). The angular gyrus multiple functions and multiple subdivisions. *The Neuroscientist*, *19*(1), 43-61.

Shah, N. J., Marshall, J. C., Zafiris, O., Schwab, A., Zilles, K., Markowitsch, H. J., & Fink, G. R. (2001). The neural correlates of person familiarity A functional magnetic resonance imaging study with clinical implications. *Brain*, *124*(4), 804-815.

Starrfelt, R. (2007). Selective alexia and agraphia sparing numbers—A case study. *Brain and Language*, *102*(1), 52-63.

Sturm, W. (2006). *WAF-Testbatterie (Wahrnehmungs-und Aufmerksamkeitsfunktionen).* Schuhfried, Wien.

Sturm, W., & Willmes, K. (1999). *Verbaler und nonverbaler Lerntest (VLT/NVLT).* Hogrefe, Göttingen.

Sturm, W., Willmes, K., & Horn, W. (1993). *Leistungsprüfsystem für 50-90jährige (LPS 50+): von Walter Sturm, Klaus Willmes, Wolfgang Horn*. Hogrefe Verlag für Psychologie.

Tewes, U. (1991). *Hamburg-Wechsler Intelligenztest für Erwachsene. Revision.* Bern.

Tohgi, H., Saitoh, K., Takahashi, S., Takahashi, H., Utsugisawa, K., Yonezawa, H., ... & Sasaki, T. (1995). Agraphia and acalculia after a left prefrontal (F1, F2) infarction. *Journal of Neurology, Neurosurgery & Psychiatry*, *58*(5), 629-632.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience, 5*, 1.

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal–ventral language pathways. *Neuron, 72*, 385–396.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do?. *Nature Reviews Neuroscience*, *10*(11), 792-802.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., ... & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage*, *30*(4), 1414-1432.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, *55*(1), 143-156.

Wechsler, D. (1987). *WMS-R: Wechsler Memory Scale-Revised: Manual*. San Antonio: Psychological Corporation.

Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the ventral pathway got lost–and what its recovery might mean. *Brain and Language*, *118*(1), 29-39.

Study 2:

Differing connectivity of Exner's area for numbers and letters.

Elise Klein, Klaus Willmes, K., Stefanie Jung, Stefan Huber,
Lucia W. Braga, Korbinian Moeller

Abstract:

There is a growing body of evidence indicating a crucial role of *Exner's area* in (hand-)writing symbolic codes such as letters and words. However, a recent study reported a patient with a lesion affecting Broca's and Exner's area, who suffered from severe peripheral agraphia for letters but not for Arabic digits. The authors suggested a speculative account postulating differential connectivity of Exner's area for numbers and letters in order to explain this dissociation.

In the present study, we evaluated this account, employing atlas-based tractography for the patient's anatomy, deterministic fiber-tracking as well as an automated toolkit to investigate the connectivity of Exner's area in healthy adults. In particular, fiber pathways connecting Exner's area with areas associated with language processing (e.g., the arcuate fascicle, ventral pathways encompassing the External/Extreme capsule system) reached the inferior part of Exner's area, while fronto-parietal fibers (e.g., the superior longitudinal fascicle) connected the upper part of Exner's area with the intraparietal sulcus typically involved in number processing. Our results substantiated the differential connectivity account for Exner's area by identifying the neural connections between fiber tracts and cortex areas of interest. Our data strongly suggest that white matter connectivity should be taken into account when investigating the neural underpinnings of impaired and intact human cognition.

1. Introduction

Is there a difference between handwriting a letter or a single-digit number? Consider a person who is able to read aloud both, words ("mountain") and numbers ("352"), to copy them, and to type them to dictation into a computer keyboard. Could one nevertheless expect a difference between writing an "R" or a "4" by hand?

Indeed, such a specific dissociation was recently described by Jung, Halm, Huber, Willmes and Klein (2015). In their single case study, patient CU was well able to type in letters, words, and pseudo-words on a computer keyboard on dictation. Also copying of words and isolated single letters was not a problem. However, spontaneous hand-writing and writing to dictation of words and even single letters was impossible, indicative of agraphia for letters/words. However, and even more interestingly, writing down single- and multi-digit numbers to dictation was largely unaffected. To evaluate the origin of this particular impairment, Jung et al. (2015) considered CU's brain lesion. CU was a 52-year-old patient suffering from severe peripheral agraphia following a left hemispheric stroke. He had a single lesion in parts of Broca's Area (Brodmann Area [BA] 44/45) extending superiorly into tissue referred to as "Exner's area" (BA 6). Already in 1881, Sigmund Exner isolated agraphic symptoms as a distinct syndrome and postulated that lesions of the posterior part of the middle frontal gyrus (MFG) may lead to specific writing impairments (Exner, 1881).

A meta-analysis of neuroimaging studies on handwriting corroborated the idea that Exner's area has a crucial role in handwriting (Planton, Jucla, Roux, & Démonet, 2013, see also Roux, Draper, Köpke & Démonet, 2010). While the meta-analysis revealed a left-hemispheric parieto-frontal network to be involved in writing, Exner's area in the posterior part of the MFG was the strongest and most reliable area subserving handwriting (Planton et al., 2013). From a theoretical point of view, Planton and colleagues (2013) argued that Exner`s area is a region connecting abstract graphemic representations with the motor execution programmes for handwriting. This is in line with findings based on clinical data reported by Binkofski and Buccino (2004), who described Exner's area as the final pathway, in which linguistic impulses are transferred into writing programmes (i.e., grapheme formation and their temporal sequencing). Furthermore, the results of Planton et al. (2013) support an earlier interpretation by Roux and colleagues (2009) that Exner's area serves as an interface between the allographic specification of the grapho-motor programmes and (abstract) orthographic representations. Thereby, Exner's area integrates central and peripheral stages of handwriting (see also Roux et al., 2010).

Thus, a lesion in Exner's area, as observed in CU, would be responsible for the general pattern of behavioural problems observed. If Exner's area is crucial for connecting abstract graphemic representations with motor execution programmes of handwriting, CU should be impaired in handwriting any shapes of letters, words, etc. to dictation, while the ability to copy shapes, to read them or to type them into a computer keyboard should be preserved. But how does this fit with CU's preserved ability to handwrite numbers to dictation? Could there be a different neural substrate for handwriting numbers?

Accordingly, Jung et al. (2015) argued that the encoding of numbers - even of single digits – may automatically activate semantic information about the numerical quantity information they denote (e.g., Dehaene & Cohen, 1997; Dehaene, Piazza, Pinel, & Cohen, 2003). There is accumulating evidence corroborating this view, even when numerical quantity information is irrelevant or detrimental for solving the task at hand (e.g., Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003; Klein, Moeller, Nuerk, & Willmes, 2010). Thus, Jung et al. (2015) suggested that activation of this semantic meaning might have helped the patient to handwrite numbers to dictation. Nevertheless, the question remains how such an activation of semantic information might be of any help when Exner's area is lesioned – as in CU.

Considering this fact, Jung et al. (2015) indicated that – like many other cortical areas – Exner's area may be connected to other cortical areas via different dorsal and ventral fiber pathways. Furthermore, the authors suggested that the activation of semantic information about a number's magnitude, which is associated with the bilateral intraparietal sulcus (IPS; Dehaene et al., 2003; Arsalidou & Taylor, 2011), might have led to additional input into preserved upper parts of Exner's area via a dorsal connection in CU. In contrast, the specific process of phoneme-to-grapheme conversion is traditionally viewed as a dorsal pathway function. Mostoften it was associated with the arcuate fascicle (Kemmerer, 2014), which connects Broca's and Wernicke's area (Catani & ffytche, 2005). Moreover, Andrews (2016) proposed that with increasing automation of reading and writing processes the dorsal pathway may be used less. Instead, a more ventral pathway implying no sub-vocal accompagniment may be preferred (Andrews, 2016). This is in line with the fact that letters and words are typically processed in left-hemispheric language areas (cf. Vigneau et al. 2006 for a meta-analysis), which should be connected to Exner's area encompassing ventral pathways as well.

Unfortunately, a direct reconstruction of the white matter affected by CU's lesion was impossible because we could not acquire diffusion tensor imaging data (DTI) in him. Thus, the suggestion of Jung et al. (2015) that differential connectivity of Exner's area with brain areas associated with the processing of numbers and language remains speculative unless it is evaluated by means of fiber-tracking. Importantly, however, the hypothesis of Jung et al.'s (2015) should be based on general anatomical realities. The assumption should hold for healthy subjects in general, independent of age. Thus we find it reasonable to evaluate to evaluate this account without DTI data acquired in CU.

In the current study, we pursued this idea by means of three different approaches, namely by (i) atlas-based tractography for the brain imaging data of CU, (ii) fiber-tracking in a sample of young healthy adults, and (iii) using an automated toolkit to create disconnectome maps. In case the propositions of Jung et al. (2015) were correct in principle, fiber-tracking results should substantiate differing connectivity of Exner's area for the processing of numbers and language. Thereby, the aim of the present study was to find neuro-functional and neuro-structural evidence for the dissociation observed between handwriting numbers and letters in CU.

2. Methods

2.1. Atlas-based tractography for CU

Patient: CU was a 52-year-old right-handed former civil engineer suffering from severe peripheral agraphia due to a left hemisphere stroke after spontaneous dissection of the internal carotid artery 3 years and 6 months before the study. Initially, CU also showed severe mixed transcortical aphasia and right hemiparesis. At the time of the study, reading aloud as well as composing words or phrases from cards showing graphemes or morphemes, respectively, was well preserved[8].

MRI Data acquisition: For CU, a high-resolution T1 anatomical scan was obtained (TR = 19 s, matrix = 256 x 256 mm^2, 190 slices, voxel size = 1 × 1 × 1 mm^3; TE = 4.9 ms; flip angle = 25°) with a 3T Siemens Magnetom Trio Tim MRI system (Siemens AG; Erlangen, Germany) at the RWTH Aachen University Hospital. The anatomical scan was acquired at the end of experimental sessions, the results of which have been reported elsewhere (Jung et al., 2015). The patient gave his written informed consent to participate. The study was approved by the local Ethics Committee of the

[8] For a more detailed description of the neuropsychological profile of patient CU see Jung et al. (2015)

Medical Faculty of the RWTH Aachen University Hospital and conducted in accordance with the latest version of the Declaration of Helsinki.

Atlas-based analysis: MRI data analysis from patient CU was performed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). Imaging data was normalised into standard stereotaxic Montreal Neurological Institute space (MNI, McGill University, Montreal, Canada). Image data were not smoothed in the spatial or time domain. In a second step, different fiber tracts as provided in the most recent DTI atlas available (Rojkova et al., 2015) were overlayed on CU's anatomy in MNI space using MRIcron (http://www.mccauslandcenter.sc.edu/mricro/mricron/) software, to show in a standardized space which tracts were connected to lesioned or unaffected tissue, respectively.

2.2. Fiber-tracking in healthy adults

Participants: DTI data were collected from a sample of 20 right-handed healthy volunteers (10 females, mean age: 23 years, range 18-26 years). Participants were scanned with the approval of the local Ethics Committee of the Medical Faculty of the Eberhard Karls University in Tuebingen. All participants gave their written informed consent. All participants had no neurological or psychiatric history and were not taking any psychoactive medication.

DTI Data acquisition: MRI and DTI data were acquired on a 3T Siemens TIM Trio scanner (Siemens, Erlangen, Germany) in Tuebingen. For DTI a total of 68 scans with 69 slices was acquired using a diffusion sensitive spin-echo EPI sequence with CSF-suppression [61 diffusion encoding gradient directions (b-factor = 1000 s/mm^2) and 8 scans without diffusion weighting (b value = 0 s/mm^2), voxel size = 2 × 2 × 2 mm^3, matrix size = 104 × 104 $pixel^2$, TR = 11.8 s, TE = 96 ms, TI = 2.3 s]. Furthermore, an additional high-resolution T1 anatomical scan was obtained (160 slices, voxel size = 1 × 1 ×1 mm^3, TR = 2.2 s, TE = 2.6 ms, matrix = 256 x 256 mm^2).

For each slice, raw diffusion-weighted data was registered and corrected simultaneously for subject motion and eddy-current induced geometrical distortions using ExploreDTI v. 4.8.5 (http://www.exploredti.com, see Leemans & Jones, 2009). Constraint spherical deconvolution (CSD) was performed to estimate multiple orientations in voxels containing different populations of crossing fibers (Alexander, 2006). Afterwards whole-brain deterministic tractography was employed using an interpolated streamline algorithm that propagates from voxel to voxel, following a step length of 1 mm and a maximum angle threshold of 35°. Fractional anisotropy (FA), a scalar value that captures the degree of diffusion anisotropy, was computed from the eigenvalues of the diffusion tensor along the defined segments (Basser & Pierpaoli, 1996). Voxels showing FA values below 0.2 were excluded from tractography (Jones et al., 2002; Jones, 2003; 2004). The motion-corrected whole-brain tractography was then imported to TrackVis (http://www.trackvis.org; Wedeen et al., 2008). This fiber-tracking software allows amongst other features for the identification of the tracts and their visualization in 3-dimensional space.

Fiber-tracking analysis: We aimed at evaluating the connectivity of Exner's area with the two most important activation sites reported in Jung et al. (2015), i.e., (i) intraparietal cortex, which is associated with processing numbers in general (without the need to write them down, e.g., Arsalidou & Taylor, 2011, for a meta-analysis) and (ii) the superior and middle temporal gyri, which are associated with writing down letters/words to dictation (although probably more at the input stage before the actual hand-writing is executed). Nevertheless, as we were interested in the processing pathways from more basic input- and content-related processing to actual hand-writing, these three areas were defined as seed regions. After co-registering the T1 anatomical scan to the b0 image, peak coordinates for left IPS, left STG and left MTG as well as Exner's area were taken from Jung et al. (2015). In particular, coordinates for the left IPS were taken from Table S3, reflecting the specific contrast dictation vs. copying of numbers (MNI: -44, -46, 58). Coordinates for the left superior (MNI: -68, -24, 6) and left middle (MNI: -50, -38, -2) temporal gyrus were taken from Table 2, reflecting relatively stronger activation when writing down letters to dictation as compared to writing down the mental images of the letters. Finally, the center of the lesion within the MFG was determined (MNI: -36, 4, 54).[9] Then all coordinates were transferred from MNI space to the native space of each participant's DTI data, using the inverse normalization parameters obtained during segmentation of the T1 anatomical scan (unified segmentation as implemented in SPM12) and enlarged to a sphere with a radius of 4 mm to reach white matter with simultaneous avoidance of bias due to manual hand-drawing (e.g., Kreher et al., 2008, for a similar procedure in probabilistic tracking; Willmes, Moeller & Klein, 2014 for deterministic tracking). These spheres defined the seed regions for the fiber-tracking procedure. To substantiate the tracking results, we switched seed and target regions in our analysis. This means that each seed region also served as target region and the other way around. Only fibers connecting region A to region B but also region B to region A were considered meaningful and are reported in the results section.

After having acquired whole-brain tractography for the 20 individual participants' data sets, these tracts were spatially normalized and averaged using a method similar to the one previously described by Jones et al. (2002) and Catani & Thiebaut de Schotten (2008). The three main fiber

[9] The lesion covered also the peak coordinate identified for Exner's area in the ALE meta-analysis by Planton et al. (2013, MNI: -22, -8, 54).

pathways described in the results section below were identified in each of the 20 participants. Similar to Catani & Thiebaut de Schotten (2008) one representative data set was used to perform virtual dissections (Figure 2).

2.3. Automated toolkit for disconnectome maps

Using the software BCBtoolkit (www.brainconnectivitybehaviour.eu/) the localization of CU's lesion in MNI 152 space was compared in an automated way with the connectivity pattern of 10 healthy participants provided by the toolbox. In a next step, disconnectome maps were computed allowing the identification of fibers probably disrupted by the lesion.

3. Results

3.1. Atlas-based results for CU

For the atlas-based results, patient CU's anatomy was depicted in standard MNI space on transverse slices of a high-resolution T1-weighted structural fMRI scan (Figure 1A). As already outlined in Jung et al. (2015), CU's lesion mainly affected Broca's area but also extended into Exner's area. For a better visualization of affected tissue, also fluid-attenuated inversion recovery (FLAIR) sequence images are provided (Figure 1B). In the (white) gliosis zones, it seems possible that additional white matter tissue was damaged. However, in gliotic regions typically only a small fraction of neurons (with their respective axons) is affected, which is then replaced by gliotic connective scar tissue. It can be seen that large parts of Exner's area and the ventral pathway were replaced by a pseudocyst following a colliquative necrosis. On the other hand, only parts of the SLF II might be affected by some gliosis.

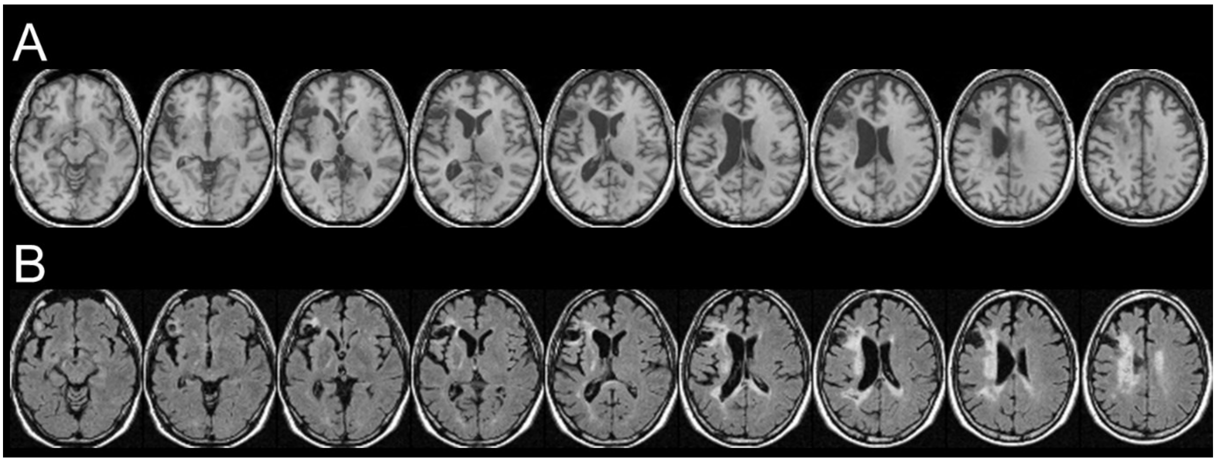


Fig. 1: Patient CU's anatomy in standard Montreal Neurological Institute (MNI) space on transverse slices. (A) Depicts a high-resolution T1-weighted structural MRI, while (B) shows a fluid-attenuated inversion recovery (FLAIR) scan. In most cases, T2-weighted images such as FLAIR scans show more extensive injury (e.g., gliosis) that is difficult to detect on T1-weighted scans. While large parts of Exner's area and the ventral pathway are replaced by a pseudo cyst following a colliquative necrosis, parts of the superior longitudinal fascicle (SLFII), might be affected by some gliosis (white tissue).

According to the dual loop model of language processing in the brain (e.g., Weiller, Bormann, Saur, Musso, & Rijntjes, 2011), Broca's area is connected to other language areas such as Wernicke's area (located within the left superior temporal gyrus) both dorsally (via the arcuate fascicle) and ventrally via the external/extreme (EC/EmC) system, which is encompassed by the inferior-fronto-occipital-fascile, IFOF). Therefore, we first considered the overlay of CU's lesion with the arcuate fascicle. As can be taken from Figure 2B, the arcuate fascicle overlapped to a considerable extent with CU's lesion, especially in the upper part of Broca's as well as in the main part of Exner's area. Second, we evaluated the overlay of CU's lesion with the IFOF, encompassing the EC/EmC system ventrally, thereby connecting temporal language areas ventrally with inferior frontal areas such as Broca's area. Figure 2C illustrates that CU's lesion in Broca's area fully overlaps with the IFOF making any ventral connection to Exner's area via Broca's area unlikely. Furthermore, it can be seen that the IFOF does not reach Exner's area and, thus, does not seem to connect Exner's area with any other cortical region in CU.

Finally, we inspected the overlay of CU's lesion with the superior longitudinal fascicle (SLF II), which reflects a dorsal connection of frontal and parietal cortex areas including the IPS. Figure 2D depicts that the SLF II indeed reaches into the (intact) upper part of Exner's area in CU connecting Exner's area with parietal cortex, the IPS in particular.

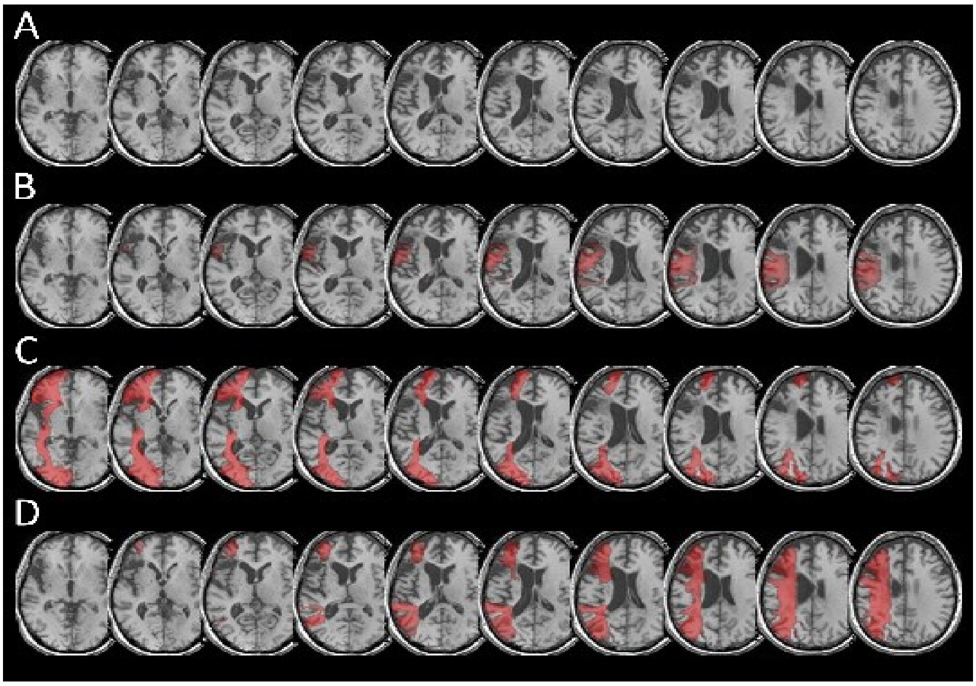


Fig. 2: (A) Patient CU's anatomy in standard MNI space on transverse slices of a high-resolution T1-weighted structural fMRI scan. The lesion affects Broca's area and extends into Exner's area due to a left-hemispheric MCA stroke after ACI dissection. **(B)** Overlay of CU's brain with the arcuate fascicle, which connects Broca's area (BA 44 and BA 45) with Wernicke's area. As can be taken from the figure, the fibers of the arcuate fascicle overlap to a considerable extent with patient CU's lesion, especially in the upper part of Broca's as well as in the lower part of Exner's area. **(C)** Overlay of CU's anatomy with the inferior-fronto-occipital fascicle (IFOF), which in the temporal lobe encompasses ventrally the external/extreme capsule (EC/EmC) system, thereby connecting language areas ventrally with inferior frontal areas such as Broca's area. As can be taken from the figure, the lesion in Broca's area fully overlaps with the IFOF, so there is no ventral connection to Exner's area via Broca's area. Furthermore, it can be seen that the IFOF does not have contact with Exner's area. **(D)** Overlay of patient CU's anatomy with the superior longitudinal fascicle (SLF II). As can be seen clearly, the SLF II reaches into the upper part of Exner's area connecting Exner's area with parietal cortex.

3.2. Fiber-tracking results

First, separate left-hemispheric trackings were run between Exner's area and the intraparietal cortex (IPS) as well as Exner's area and the superior temporal cortex (including Wernicke's area) in all 20 participants. Interestingly, both connections were found for the fiber tracking in all 20 participants (see Figure 3). As can be taken from Figure 3, the fibers connecting Exner's area and Wernicke's area ran both, dorsally (via the arcuate fascicle) as well as ventrally via the external/extreme capsule (EC/EmC) system. However, both, the ventral connection as well as the dorsal arcuate fascicle reached Exner's area considerably inferior to the SLF II (connecting Exner's area with the intraparietal cortex). Therefore, in case the lower part of Exner's area is affected by a lesion (as was the case in CU), there may nevertheless be a preserved dorsal fronto-parietal connection to the IPS via the SLF II.

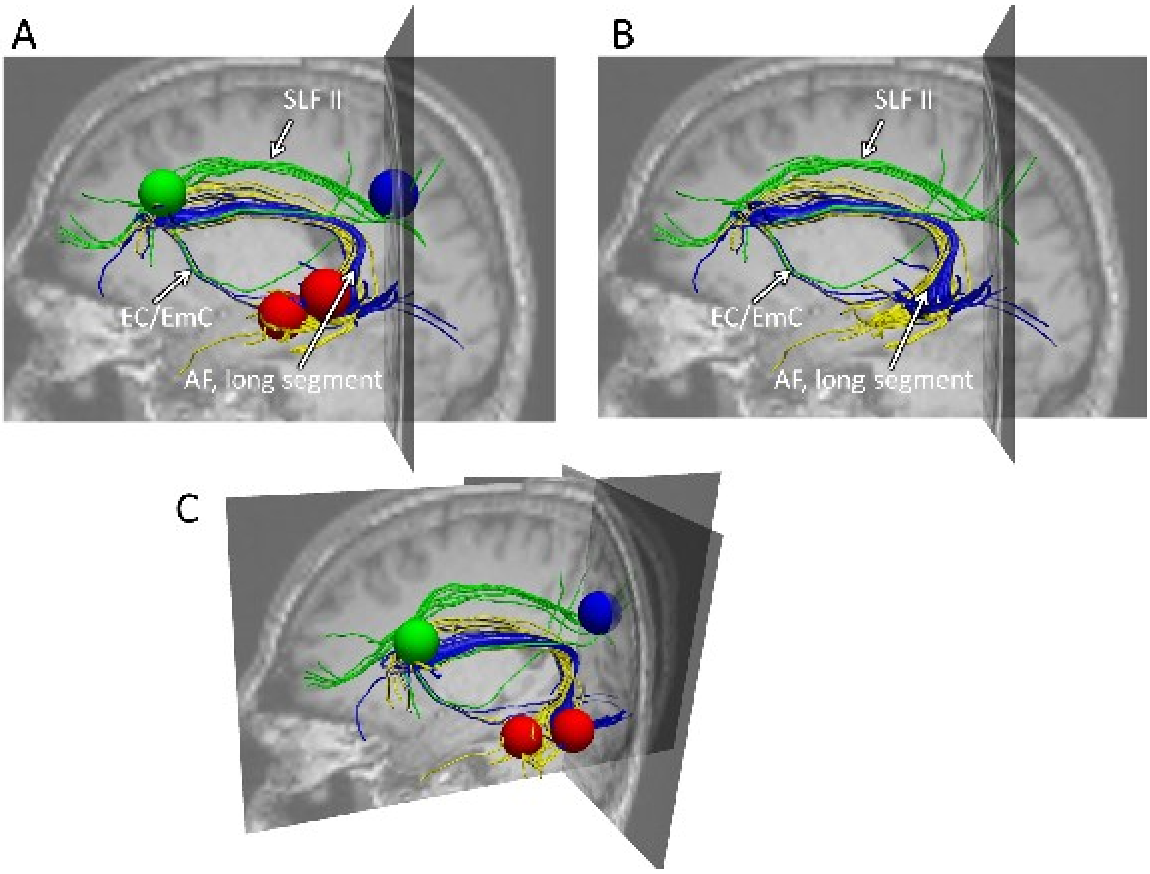


Fig. 3: Example of typical fiber tracking results within healthy participants. **(A)** Depicts the sagittal view of both fiber trackings: (i) the connections between Exner's area (blue sphere) to the intraparietal cortex (green sphere) via the superior longitudinal fascicle (SLF II; green fibers), (ii) the connections between Exner's area (blue) and the superior temporal gyrus (upper red sphere) given in blue fibers (AF, longitudinal segment), and (iii) the connections between Exner's area (blue) and the middle temporal gyrus (lower red sphere) given in yellow fibers. As can be seen, the blue fibers connect these two areas, both, dorsally (via the arcuate fascicle) as well as ventrally via the external/extreme capsule (EC/EmC) system. Importantly, both, the ventral connection as well as the dorsal arcuate fascicle reach Exner's area considerably inferior to the SLF II. Therefore, in case the lower part of Exner's area is affected by a lesion, there may be a dorsal fronto-parietal connection left, nevertheless. **(B/C)** Illustrate the same findings from a perspective a bit more superior and tilted towards the reader, providing also the coronal and transversal slices. **(B)** This view is given with the seed regions used, while in **(C)** only the fibers within the brain context are depicted. Note that following Jones et al. (2002) DTI analyses were based on data averaged across all participants, whereas segmented fibers of a representative participant are shown here for illustration purposes.

3.3. Automated toolkit results

Further substantiating the results of the two previous approaches, disconnectome maps indicated that the SLF II still seemed to be intact, while both, arcuate fascicle as well as ventral fibers seemed to be disconnected by the lesion of CU. This is indicated by the missing fiber structures in Figure 4.

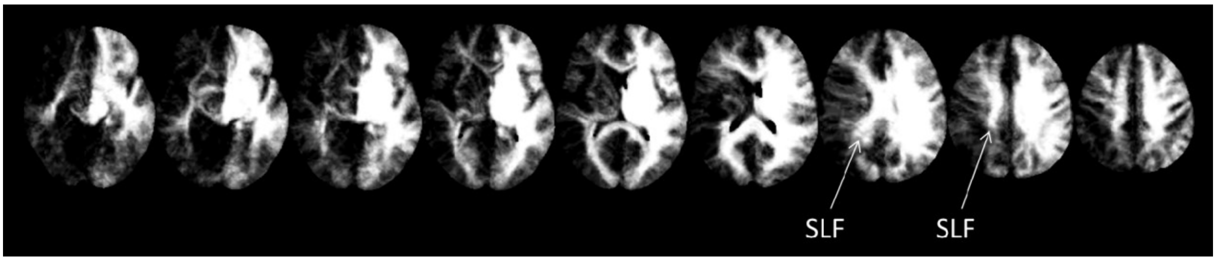


Fig. 4: Disconnectome maps as provided by BCB toolkit. The SLF II seemed to be intact, while both, the anterior segment of the AF as well as ventral fibers seemed to be affected by the lesion.

4. Discussion

Recently, Jung et al. (2015) had reported the single case of CU, whose writing of numbers to dictation appeared largely unaffected, while spontaneous hand-writing and writing to dictation of words and even single letters was severely impaired. The present study set off to provide first neuro-structural evidence for the differing connectivity account regarding Exner's area proposed by Jung and co-workers to explain this dissociation.

To this end, we employed an atlas-based analysis of the lesion data of CU, fiber-tracking in healthy adults with an intact Exner's area, and an automated toolkit to evaluate disconnectome maps. The results of all these analyses corroborated the conclusions of Jung et al. (2015) consistently. In the tracking analysis, fiber pathways connecting Exner's area with areas associated with language processing (e.g., the arcuate fascicle) primarily reached the lesioned inferior part of Exner's area in CU, while fronto-parietal fibers (e.g., the SLF II) connected the upper (partly intact) part of Exner's area with the IPS in CU. It is well possible that in healthy adults the processing of words encompasses the long segment of the arcuate fasciculus to involve Broca's/Exner's area, while the processing of numbers encompasses the anterior segment of the arcuate fasciculus as well. So, functions of the SLF II might be taken over by the anterior segment of the AF. This possible substitution is important for the argument, as parts of the SLF II (Figure 1B) as well as of the anterior segment of the AF (Figure 4) may be affected by gliosis in patient CU. However, at least the disconnectome maps calculated by the automated toolkit rather suggested that the SLF II should be more probably intact/less affected than the anterior segment of the AF. Taken together, all these results are well compatible with the speculative explanation of Jung et al. (2015) that (partly) intact connections with cortex areas subserving the semantic processing of number magnitude (either by the SLF II or by the anterior segment of the AF) may have led to spared handwriting of numbers in CU.

Patient CU had to handwrite either numbers or letters/words upon oral dictation. For the case of numbers there is evidence that numerical magnitude information is automatically activated when one encounters a number, which is reflected by IPS activation (e.g., Eger et al., 2003; Klein et al., 2010). This activation requires connections between temporal auditory cortex and the intraparietal cortex, probably subserved by the inferior longitudinal fascicule (ILF; e.g., Seghier, 2013). Once the IPS is involved, the semantic information may then be transferred via the SLF II to the upper part of Exner's area (which was partially intact in CU) and may trigger the retrieval of the respective digit's shape (see Figure 1A and 1D). The fiber-tracking results for healthy adults as well as the disconnectome maps further corroborate these anatomical relations: The SLF II (depicted in green in Figure 2A-C) runs much more superior and medial than the arcuate fascicle (given in yellow in Figure 2A-C) before entering into Exner's area from superior. This substantiates the idea of a preserved dorsal connection of the upper part of Exner's area and the IPS in CU which may account for his ability to write numbers to dictation by means of relying on number magnitude information. In line with this argument, there was strong perilesional activation observed by Jung et al. (2015) in the upper part of Exner's area in CU when handwriting a verbally dictated number, which may reflect the retrieval of grapho-motor patterns for digits.

Yet, based on the present data from healthy adults it cannot be decided whether the semantic magnitude information of a verbally presented number might also be relayed ventrally via the EC/EmC system connecting Broca's to Exner's area, for instance, in case the superior longitudinal fascicule (SLF II) connecting the IPS directly to Exner's area is lesioned.

However, for letters a ventral connection encompassing the EC/EmC system and running via Broca's area to Exner's area situated in the direct vicinity to Area 44 seems to exist (Exner, 1881; Planton et al. 2013). Importantly, however, in the case of CU the ventral connection terminates considerably more inferior in the frontal cortex than the SLF (see also Seghier, 2013) and his lesion affected this inferior part of the frontal cortex. Therefore, numerical magnitude information from the IPS must have been transmitted most probably exclusively via the dorsal connection in CU.

This argument is further corroborated by the fiber-tracking results for the temporal language areas. Writing letters to dictation is typically associated with mostly ventral connections encompassing the EC/EmC system (Weiller et al., 2011) as well as the dorsal arcuate fascicle. Importantly, both these pathways of the dual loop model of language processing enter Exner's area from inferior and were disrupted by CU's lesion (Figure 1A). Therefore, it was not possible for CU to retrieve the grapho-motor patterns of letters while those for numbers may be accessible by the dorsal connections to the IPS. Of course, one might argue that the written production of numbers and words does not necessarily require a direct monosynaptic connection between posterior temporal/parietal and anterior frontal areas. The networks involved may be polysynaptic and indirect through long and short association fibers. Indeed, numerous connections to the frontal lobes have been described only recently (e.g., Catani et al., 2012). Additionally, U-fibers exist connecting Exner's area and both inferior and superior frontal gyri. However, this does not affect our argument: the inferior part of Exner's area of CU was lesioned and so should have been the respective U-fibers. In contrast, our data suggest the upper part of Exner's area to be intact, so superior U-fibers might further facilitate preserved cognitive operations. Last but not least, we wish to clarify that we aimed at evaluating the account suggested by Jung et al. (2015). Future studies are needed to substantiate our results or provide evidence for another account explaining the dissociation observed in CU.

This interpretation is further substantiated by the results of a writing training, which was successfully completed by CU and reported by Jung et al. (2015). The authors trained CU to associate specific mental images with each letter (e.g., a cocktail glass for Y). After a month of intensive training, patient CU was able to write most of the letters of the alphabet. A follow-up test 6 months later revealed that he was still able to do so. Jung et al. (2015) argued that due to the training additional pictorial semantic information was attached to the letters by means of mental imagery. On the neural level mental imagery is associated with the precuneus (e.g., Ganis, Thompson, & Kosslyn, 2004; Cavanna & Trimble, 2006 for a review). Interestingly, there are white matter connections from the temporal cortex associated with language processing (Seghier, 2013) to the precuneus, which is directly neighbouring the IPS (in fact, the sphere of 4 mm around the IPS in the fiber tracking study included parts of the precuneus as well). Thus, once the mental images associated with letters are established, their retrieval may involve the precuneus, which is connected to Exner's area very similar to the IPS. Connections run via the SLF II to the upper part of Exner's area, which was spared in CU. This may also explain why the mental imagery intervention was successful, allowing CU to associate grapho-motor patterns with the mental images reflecting letters.

Taken together, these results lend strong support to the hypothesis claiming differing connectivity of Exner's area for numbers and letters. Nevertheless, it needs to be considered that our account based on fiber pathway analyses relies on fMRI data, which do not necessarily indicate that activated areas are functionally critical for the task. This uncertainty remains even though we carefully selected activations from differential contrasts as seed regions for our trackings.

Therefore, further empirical evidence would be desirable whether the dissociation between letter and number writing observed in patient CU is explained sufficiently by our account of differing connectivity.

5. Conclusion

Our results corroborate the conclusion of Jung et al. (2015) that the dissociation between handwriting of numbers vs letters is due to differing connectivity of Exner's area with cortex sites subserving number and language processing. These data indicate that domain-specificity in the human brain may evolve from the specific combination of activated cortex areas and their interconnections (see also Willmes et al., 2014). In particular, the present study indicates how informative the investigation of specific white matter connectivity can be when investigating the neural underpinnings of impaired or unimpaired human cognition.

Acknowledgements

The current research was supported by the Leibniz-Competition Fund (SAW) providing funding to Elise Klein, supporting Stefan Huber (SAW-2014-IWM-4) as well as by the German Research Foundation (DFG) providing funding to Klaus Willmes (Wl 1804/ 6-1) and Korbinian Moeller (MO 2525/ 2-1) as well as the Margarete von-Wrangell Fellowship of the European Social Fonds (ESF) and the Ministry of Science, Research and Arts Baden-Wuerttemberg, supporting Elise Klein. Stefanie Jung was supported by the Science Campus Tuebingen (Cluster 1, TP8 in the third phase) providing funding to Juergen Heller.

References

Alexander, D.C. (2006). *An introduction to computational diffusion MRI: the diffusion tensor and beyond.* In: Visualization and processing of tensor fields. Berlin: Springer. pp. 83–106.

Andrews, D. (2016). *Language Disorders.* In: Neuropsychology. From Theory to Practice. London & New York: Taylor & Francis. pp. 487-488.

Arsalidou, M., & Taylor, M.J. (2010). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage. 54,* 2382-2393.

Basser, P.J., & Pierpaoli, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance Series B, 111,* 209–219.

Binkofski, F., & Buccino, G. (2004). Motor functions of the Broca's region. *Brain and Language*, *89*(2), 362-369.

Catani, M., Dell'acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., & Thiebaut de Schotten, M. (2012). Short frontal lobe connections of the human brain. *Cortex, 48(2),* 273-91.

Catani, M., & ffytche, D. (2005). The rise and fall of disconnection syndromes. *Brain, 128,* 2224- 2239.

Catani, M., & Thiebaut de Schotten, M. (2012). *Atlas of Human Brain Connections* (pp. 56–74). Oxford: Oxford University Press.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex, 44,* 1105–1132.

Cavanna, A. E. & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129(3),* 564-583.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33,* 219–250.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20,* 487-506.

Eger, E., Sterzer, P., Russ, M.O., Giraud, A., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, *37*, 1 – 20.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research, 20(2),* 226-241.

Jones, D.K. (2003). Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magnetic Resonance in Medicine, 49,* 7–12.

Jones, D.K. (2004). *Diffusion MRI - Theory, Methods, and Applications*. Oxford: University Press.

Jones, D.K., Griffin, L.D., Alexander, D.C., Catani, M., Horseld, M.A., Howard, R., & Williams, S.C. (2002). Spatial normalization and averaging of diffusion tensor MRI data sets. *Neuroimage, 17,* 592–617.

Jung, S., Halm, K., Huber, W., Willmes, K., & Klein, E. (2015). What letters can "learn" from Arabic digits - fMRI-controlled single case therapy study of peripheral agraphia. *Brain and Language, 149*,13-26.

Kemmerer, D. (2014). Part VI: Other modalities of language use. In: Cognitive Neuroscience of Language. London & New York: Taylor & Francis.

Kreher, B.W., Schnell, S., Mader, I., I`lyasov, K.A., Hennig, J., Kiselev, V., & Saur, D. (2008). Connecting and merging fibres: Pathway extraction by combining probability maps. *Neuroimage, 43,* 81–89.

Klein, E., Moeller, K., Glauche, V., Weiller, C., & Willmes, K. (2013). Processing pathways in mental arithmetic-evidence from probabilistic fiber tracking. *PLoS ONE, 8(1).* e55455.

Klein, E., Moeller, K., Nuerk, H. C., & Willmes, K. (2010). On the neuro-cognitive foundations of basic auditory number processing: an fMRI study. *Behavioral and Brain Functions*, *6(42),* 1-13.

Leemans, A., & Jones, D.K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magnetic Resonance in Medicine, 61,* 1336–1349.

Planton, S., Jucla, M., Roux, F. E., & Démonet, J. F. (2013). The “handwriting brain”: a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex, 49(10),* 2772-2787.

Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell'Acqua, F., & Thiebaut de Schotten, M. (2015) Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. *Brain Structure and Function.*

Roux, F. E., Draper, L., Köpke, B., & Démonet, J. F. (2010). Who actually read Exner? Returning to the source of the frontal “writing centre” hypothesis. *Cortex, 46(9),* 1204-1210.

Roux, F. E., Dufor, O., Giussani, C., Wamain, Y., Draper, L., Longcamp, M., & Démonet, J. F. (2009). The graphemic/motor frontal area Exner's area revisited. *Annals of Neurology*, *66(4),* 537-545.

Rijntjes, M., Weiller, C., Bormann, T., & Musso, M. (2012). The dual loop model: its relation to language and other modalities. *Frontiers in Evolutionary Neuroscience,* 4:9.

Seghier, M. L. (2013). The angular gyrus multiple functions and multiple subdivisions. *The Neuroscientist*, *19(1),* 43-61.

Vigneau, M., Beaucousin, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage, 30,* 1414-32.

Wedeen, V.J., Wang, R.P., Schmahmann, J.D., Benner, T., Tseng, W.Y., Dai, G., Pandya, D.N., Hagmann, P., D'arceuil, H.E., & De Crespigny, A.J. (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage, 41,* 1267–1277.

Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the verntral pathway got lost - And what its recovery might mean. *Brain and Language, 118,* 29-39.

Willmes, K., Moeller, K., & Klein, E. (2014). Where numbers meet words – A common ventral network for semantic classification. *Scandinavian Journal of Psychology, 55*, 202-211.

Section 2:

Handwriting and typing:

Modal influences on spelling and arithmetic

Study 3:

Mode effect: an issue of perspective?
The influence of writing mode in children with and without dyslexia

Stefanie Jung, Korbinian Moeller, Elise Klein & Juergen Heller

Abstract

Digital technology has an increasing influence on writing habits and processes. Previous studies comparing differences between handwriting and typing on the computer keyboard (i.e., the mode effect) mainly assessed typically developing children and revealed inconsistent results across different age groups and tasks. So far, data examining mode influences on writing in children with developmental writing disorders are scarce.

Accordingly, the present article reports a study investigating the influence of writing mode (i.e., handwriting vs. typing) on spelling accuracy, writing speed and self-corrections in a writing-to-dictation task in typically developing children and children with dyslexia from two different levels. Results show that, on a general level, writing mode was found to influence overall spelling accuracy (i.e., more spelling errors while typing at the computer keyboard), writing speed (i.e., writing speed decreased during typing), and self-corrections (i.e., more self-corrections while typing) comparably in both populations. On a more rule-specific level, outcomes for writing speed and self-corrections substantiated these findings. However, writing accuracy may not be subject to a general mode effect. Instead, a mode effect was only apparent for capitalization. For all other spelling rules, writing mode did not affect spelling accuracy suggesting that additional underlying processes related to writing (e.g., phonological awareness, rule and morphologic knowledge assessed by different spelling rules) were resistant to mode influences.

As such, our findings suggest that the mode effect is present for specific aspects of writing only rather than reflecting a general influence. We therefore recommend that the evaluation of media influences on children's writing and spelling abilities, whether with or without dyslexia, should consider additional underlying processes related to writing as well (i.e., according to separate spelling rules).

1. Introduction:

Digital technology has an important influence on the fields of both (school) education and promotion of children with special educational needs. Today, a lot of written requirements from our everyday life are processed digitally (i.e., by computers, tablets or mobile phones) and no longer handwritten. On the one hand, this gradual replacement of handwriting has triggered an ongoing debate about the impact of changes in writing habits on general cognitive processes such as memory (Longcamp et al., 2008; Longcamp, Zerbato-Poudou, & Velay, 2005) but also on writing processes in adults and in children (Goldberg, Russell & Cook, 2003; Mangen & Velay, 2010; Wollscheid, Sjaastad, & Tømte, 2016, for an overview). On the other hand, digital media have changed the traditional procedures of standardized assessments in school (e.g., OECD, 2016, PISA, Jerrim, 2016 for a discussion on how the results of computerized tests compare to paper-pencil tests), but also in the field of diagnostics within therapeutic contexts (e.g., Anderson, 2005). This was attributed to numerous advantages of digital media: for instance, computerized tests are characterized by simple test administrations, new test design options and simplified evaluation of test performance. In addition, learning success and progress as well as improvements within the scope of interventions can be monitored, evaluated and assessed more easily with the computer as compared to paper-pencil methods (e.g., Pellegrino & Quellmalz, 2010).

The increasing use of computers in different learning and assessment contexts has raised the question of whether the actual test mode affects writing performance. The accordingly named *mode effect* (i.e., differences between handwriting and typing on the computer) during writing acquisition has been subject of many investigations in children of different age levels (e.g., Gaskill & Marshall, 2007; Lottridge, Nicewander, Schulz, & Mitzel, 2008, for an overview). The vast majority of these studies focused on typically developing children. So far, the mode effect has hardly been evaluated in children with writing difficulties in the context of dyslexia. Furthermore, although existing studies examined various parameters such as accuracy rates, writing times and revisions (i.e., prevailing in the form of numbers of self-corrections), the question of if and to what extent writing mode influences additional processes related to writing/spelling (e.g., phonological awareness, rule and strategy knowledge) has largely been neglected. However, consideration of these additional processes may allow to draw conclusions on how different writing modes may be used to promote specific (impaired) writing processes.

The present article investigated handwriting and typing in typically developing children and in children with dyslexia by evaluating the influence of writing mode. In the following, we first report the current state on research on how writing processes change when we switch from pen(cil) to keyboard in typically developing children. Subsequently, we explore specificities of handwriting in children with developmental dyslexia. For both populations, we then discuss effects of computer use and how the writing mode affects writing (i.e., usually spelling) performance.

1.1. Handwriting and typing in typically developing children

Writing, as we learn it traditionally with a pen(cil) in the hand on a sheet of paper, is an active and highly embodied process. It requires integration of visual, graphomotor and tactile information (Mangen & Velay, 2010, for a discussion of haptics in writing). Like the trace a pencil leaves on a paper, the visual form of a letter (or grapheme) acquired in childhood leaves a neuronal representation of the specific letter's graphomotor program in the brain (e.g., James & Gauthier, 2006; Longcamp, Anton, Roth, & Velay, 2003). In turn, the association between graphemes and their specific graphomotor program affects recognition and storage of letters (i.e., Longcamp et al., 2008, 2005; but see Vaughn, Schumm, Gordon, Vaughn, & Schumm, 1992). Writing on a keyboard, however, eliminates this graphomotor component. It is replaced by a spatial process of learning different locations of letters on a computer keyboard. Furthermore, typing requires divided attention between screen and keyboard, while handwriting requires a one-directed focus on the tip of the pen (Mangen & Velay, 2010). These are some examples of how writing processes change when we switch from pen(cil) to keyboard. In view of these changes, it can be assumed that different writing modes lead to differences in writing performance (depending on individual media experience). These media-dependent differences, which might affect the comparability of computer- and equivalent paper-based writing assessments, are investigated in mode effect studies.

The mode effect has been predominantly examined in reading, mathematics and multiple choice tasks. In a five-year longitudinal study on secondary school children, Gaskill and Marshall (2007) found that paper-based testing yielded significantly better results for reading (but see Wang et al., 2008, for a meta-analysis). In contrast, however, mode of testing did not affect mathematics performance (e.g., Wang et al., 2007, for a meta-analysis). Yet, very few studies have addressed writing. These studies mainly assessed writing/typing time (i.e., speed) on different language levels (i.e., grapheme, word and/or sentence level; e.g., Berninger, Abbott, Augsburger, & Garcia, 2009), but they reported less frequently on writing performance in terms of spelling accuracy. They also considered text production, such as quality of essay writing and number of self-corrections (e.g., Goldberg et al., 2003; Wollscheid et al. 2016, for reviews on older children and primary school children). The results of these studies show opposite patterns for younger and older children: younger children were faster and produced longer essays by pen (e.g., Berninger, Abbott, Augsburger, & Garcia, 2009; Connelly, Gee, & Walsh, 2007; Wollscheid, Sjaastad, & Tømte, 2016, for a meta-anaylsis), but wrote more automatically on alphabetic (or grapheme) level by keyboard (Berninger et al., 2009). Older children, instead, usually wrote faster and produced longer texts by keyboard, but tended to make more self-corrections when typing (Goldberg et al., 2003, for a meta-analysis). Interestingly, these self-corrections were performed online during the typing process rather than at the end of the written passage as it is usually the case with handwriting (Goldberg et al., 2003). These diverging results for the different age groups were not surprising as computer experience and keyboarding speed was found to predict writing performance (e.g., Russell, 1999; Russell & Plati, 2002). Nowadays, older children are more likely to have more experience with computers. In contrast, spelling accuracy in fifth grade children did not differ between handwriting and typing (Frahm, 2013). A lack of differences between handwriting and typing was also reported by Russel (1999; see also Russel & Plati 2002) who evaluated quality of essay writing of eighth-graders.

Despite the increasing number of studies, however, it seems hardly possible to draw conclusions about the impact of the mode on children's writing skills in one direction or another. Unambiguous conclusions are even more difficult because many of these studies have methodological limitations. First and foremost, the study design is noteworthy: in many studies, either handwriting or typing (e.g., Martlewm, 1992; Sumner, Connelly, & Barnett, 2013) was assessed only which prevented direct comparisons. Furthermore, those studies, which investigated both handwriting and typing (Goldberg et al., 2003), mainly used between-subject designs (Lottridge et al., 2008 for a review). Thereby, intra-individual mode effect differences could not be considered. Moreover, between-subject designs require larger sample sizes to reliably detect small effects – but only few studies met this requirement (e.g., Wollscheid, Sjaastad, & Tømte, 2016, for a fruther discussion). Finally, from a diagnostic point of view, there is little empirical evidence on the mode effect in children with developmental writing disorders in the context of dyslexia. In this context in particular, more research would be desirable as in recent years computers have increasingly been used for diagnostics and interventions for writing difficulties (e.g., Jung et al., 2016; Mejía, Diaz, Jiménez, & Fabregat, 2012; Nicolson, Fawcett, & Sheffield, 2008), but also to compensate for dyslexic writing problems in school (e.g., Anderson, 2005).

1.2. Handwriting and typing in the context of developmental dyslexia

Children with developmental dyslexia have deficits in reading and/or writing despite unimpaired intelligence. Dyslexia is defined either by a discrepancy criterion between IQ and reading/writing performance (International Statistical Classification of Diseases and Related Health Problems, ICD-10, established by the World Health Organization, 2015), or by an IQ greater than or equal 70 (Diagnostic and Statistical Manual of Mental Disorders (DSM-V; American Psychiatric Association, 2014). The symptomatology of dyslexia is very heterogeneous. For various reasons (Döhla & Heim, 2016, for an overview of the causes of dyslexia and dysgraphia), these children struggle with different handwriting aspects, such as graphomotor planning and grapheme transcription (e.g., Kandel, Lassus-Sangosse, Grosjacques, & Perret, 2017), correct spelling (e.g., for children: Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008; Cidrim & Madeiro, 2017, for a review; for adults: Coleman, Gregg, McLain, & Bellair, 2009), and writing fluency (e.g., Sumner, Connelly, & Barnett, 2013; but see Martlewm, 1992 for contrary results) as compared to their typically developing peers. They also have difficulties in recognizing and correcting errors (e.g., Horowitz-Kraus & Breznitz, 2011, for reading). Accordingly, these difficulties have a strong impact on children's motivation to write (e.g., Berninger, Winn, et al., 2008), and thus on children's learning, self-esteem, and educational achievement (e.g., Alexander-Passe, 2006).

Computers seem to offer considerable potential to reduce at least some of the writing barriers caused by dyslexia. In addition to more general advantages such as incorporation of multisensory elements and hence increased motivational aspects (e.g., Warschauer, 1996), integrated spell-checker/word processing software can provide spelling assistance and printed block letters can enhance legibility (e.g., MacArthur, 2009, for a discussion on different computer applications to promote writing outcomes). Although these media-specific features are considered beneficial for children with dyslexia, there is little evidence from empirical research corroborating these claims. In one of these few studies, Morken and Helland (2013) evaluated typing performance of ten to eleven-year-old children with and without dyslexia in a writing-to-dictation task. Results showed that children with dyslexia wrote slower and less accurately using the computer keyboard but corrected their own writing as much as typically developing children. However, the study left open whether and to what extent the writing medium (i.e., the computer) influenced these findings or whether the same results can be expected for handwriting. On a broader level, MacArthur (2009) discussed research on media, which is used to support writing in children with learning disabilities. He noted that, although research on assistive media usage for writing is limited, some application (e.g., word processing, speech recognition) were identified as beneficial at least for some children.

Nevertheless, empirical studies are needed to substantiate the potential benefits of media and its impact on writing.

Importantly, computers are not only used for compensatory purposes, but increasingly also for diagnostics and interventions for developmental writing difficulties (e.g., Mejía, Diaz, Jiménez, & Fabregat, 2012; Nicolson, Fawcett, & Sheffield, 2008). However, the question of their comparability to (handwritten) paper-pencil tests has largely been neglected so far. To the best of our knowledge, there is currently no study that investigated writing and/or spelling performance in both, children with and without developmental writing difficulties in the context of dyslexia applying both paper-pencil and computerized diagnostic procedures. Therefore, a direct comparison of handwriting and typing is needed to evaluate the comparability of paper-pencil and computer-based methods for diagnostics (and intervention) of written language difficulties, but also learning assessments in school. Moreover, with regard to therapeutic approaches, it is imperative to gain deeper insights into the extent to which mode effects influence additional processes of writing by more specific analyses (i.e., by assessing different spelling rules in a test word separately).

1.3. Overview of the two-step approach in the study

In the current article, we evaluated handwriting and typing skills in typically developing children and children with dyslexia following a two-step procedure. In a first step, we directly compared the mode effect (i.e., the difference between handwriting and typing) between children with and without writing difficulties in the context of dyslexia. For this purpose, we investigated handwriting and typing in a combined within- and between-subject design to provide first empirical evidence on how writing performance in typically and atypically developing written language acquisition is affected by different writing modes. We used a writing-to-dictation instead of an essay writing task to avoid planning and text generation processes (cf. Morken & Helland, 2013). In addition, a dictation task has the advantage that test words can be easily parallelized to ensure comparable test conditions. We evaluated the mode effect overall spelling accuracy, writing times and self-corrections. Therefore, children's spellings were assessed as either correct or incorrect as it is the case in many learning assessments (Deimel, 2002, for an overview of German spelling assessments).

In the second step, we aimed for a more detailed evaluation of the influence of writing mode on additional underlying processes related to writing (e.g., phonological awareness, rule and morphologic knowledge) in the same samples. Therefore, spelling rule-specific analyses were applied: by evaluating different spelling rules separately for each test word (i.e., capitalization, consonant doubling, lengthening, final obstruent devoicing, and rule words) we aimed at capturing processes necessary for these specific spelling rules. For instance, realization of consonant doubling and lengthening signs requires phonological skills to monitor vowel length whereas detection of final obstruent devoicing requires more morphological knowledge. These spelling rules are taught in schools according to German curricula (e.g., Bildungsplan Sekundarstufe I, Ministerium für Kultus, Jugend und Sport, Baden Württemberg, 2016). We chose this approach because we expected that some spelling rules or additional (underlying) writing processes (e.g., initial capitalization requires an additional key to be pressed) might be more sensitive towards mode-effects.

2. First step: Consideration at general level

2.1. Objectives and hypotheses

On the basis of the previous findings, we have derived the following hypotheses:

For typically developing children, we expected longer writing times and more self-corrections for typing as compared to handwriting (Goldberg et al., 2003; Wollscheid et al., 2016). Additionally, we expected these differences to decrease with age and experience in computer use. We did not have specific hypotheses with respect to spelling accuracy because results of previous studies were inconsistent on this point (cf. Berninger et al., 2009; Frahm, 2013). Results of the present study may help to further clarify this issue.

For children with dyslexia, we expected longer writing times and lower spelling accuracy as compared to the control group in both conditions (Berninger et al., 2009). In line with findings of Morken and Helland (2013), we hypothesized no differences in the number of self-corrections between dyslexic and control children.

2.2. Methods

2.2.1. Participants:

Fifty-two children participated in the study, 22 children with dyslexia (experimental group: 12 males, mean age=11.45 years, SD=1.22 years) and 30 typically developing children matched for chronological age (control group: 13 males, mean age 11.33 years, SD=1.02 years). Only monolingual German-speaking children without any developmental disorder (except dyslexia) were included.

Children with dyslexia were initially diagnosed with dyslexia either by their local education authority, a psychologist or an educational therapist using different cognitive and linguistic measures, e.g., the Hamburg-Wechsler-Intelligenztest für Kinder – IV (HAWIK-IV, Petermann & Petermann, 2010), the Coloured Progressive Matrices (Raven, 1962), the Hamburger Schreibprobe (May, 2002), and the Lese- und Rechtschreibtest SLRT – II (Moll & Landerl, 2010). Dyslexia was moderate to severe according to the standards of the respective test procedures. In most dyslexic children (n=17), impaired writing occurred combined with reading difficulties while some children had isolated writing difficulties (i.e., dysgraphia, n=5).

Written informed consent was obtained from parents prior to the study besides children's verbal assent before actual testing. The study was approved by the local ethics committee (LEK 2014/19).

2.2.2. Procedure:

Data was collected in an individual setting in a quiet room either at our facility or at an institute for learning therapy. Children were tested in two sessions for about 60 minutes each. All children were given a writing-to-dictation task for real words and pseudowords as well as a copy-task that had to be completed in both a computer keyboard and a handwriting condition. Order of conditions (i.e., writing mode and lexicality) was counterbalanced across participants. As many German spelling assessments use gapped sentences in the dictation task (e.g., Deimel, 2002), this format was chosen. The copying task served as control task to control for writing speed (i.e., writing time) and always followed the dictation task.

In the typing (i.e., computer keyboard) condition, a 15.6-inch Lenovo ThinkPad T530 laptop with a resolution of 1024 x 768 pixels was used at normal viewing distance. Children solved the tasks by typing on a standard QWERTZ keyboard. Task relevant laptop functions (e.g., capitalization, deletion of incorrect entries) were introduced before the task, so that all children had the same prerequisites for coping with the tasks. All computerized tasks were programmed using C#.

In the handwriting condition, writing was performed on a digitizing tablet (Wacom Intuos 2) linked to a laptop and controlled by Eye and Pen 2 software (Alamargot, Chesnet, Dansac, & Ros, 2006; Chesnet & Alamargot, 2005, for a detailed description of the setup). Using a Wacom Ink Pen (inking pen), children wrote on a sheet of paper put on the surface of the tablet. All paper-pencil tasks were programmed with the Eye and Pen 2 software. This software controlled the experimental procedure, tracked all writing movements and recorded various writing parameters (e.g., position of the pen on the tablet's surface, temporal information on each individual event conveyed by the tablet). In addition to the writing tasks, children were assessed for their general cognitive ability and verbal working memory using the HAWIK-IV's Processing Speed and Digit Span subtests (Petermann & Petermann, 2010), as well as for their language skills using the Zürcher Lesetests-II (ZLT-II, Petermann, Daseking, Linder, Grissemann, & Weid, 2013). The ZLT-II also includes sub-tests assessing rapid automatized naming (RAN) and verbal working memory by capturing the syllable span for pseudowords. All diagnostic tests were administered according to the instructions described in their respective manuals.

2.2.3. Tasks:

2.2.3.1. Writing to dictation

In the writing-to-dictation task, children were required to complete gapped sentences. First, each sentence was presented visually and auditorily, followed by an auditory repetition of the target word. Children were instructed to read along the sentence during dictation, wait until the target word was repeated and then write or type in (depending on the condition) the respective target word. Overall, 34 real words as well as 20 pseudowords were administered using two parallel test versions. Tasks were preceded by two practice items in each condition. Items and sentences for both test versions and the respective translation in English are given in Appendix A1. Table 1 provides information regarding the parallelization of the test sheets.

Table 1: Overview of parallelized test items

Sheet	N	Lex.	Example	Word category			Number of syllables				Spelling rules				
				NOUN	VERB	ADJ.	1	2	3	>3	CAP	COD	LEN	DEV	RULE
Set A	34	Word	voll	16	9	9	9	15	8	3	34	16	7	5	8
	22	Pseudo	boff	10	8	4	7	9	4	0	20	4	4	2	0
Set B	34	Word	hell	16	9	9	7	15	8	3	34	16	7	6	8
	22	Pseudo	humm	10	8	4	7	9	4	0	20	4	2	2	0

Note: Test sheets were parallelized according to Lex=Lexicality (word or pseudoword), word category (noun, verb, ADJ.=adjectives), number of syllables, and spelling rules (CAP=Capitalization, COD= Consonant doubling, LEN=Lengthening, DEV=Final obstruent devoicing, RULE= Rule words.

2.2.3.2. Copying task

The copying task always followed the writing-to-dictation task. Children were asked to copy the given text as correctly as possible with their normal writing speed within a five-minute time limit. Children were allowed to revise within this time limit. Texts were counterbalanced across writing conditions and displayed on the laptop screen. Children wrote on lined paper with 2cm spacing between the lines in the handwriting condition to leave enough space for self-corrections and used the keyboard of the laptop for typing. Texts were matched as accurately as possible (see Table 2 for a content description of the text in English). Appendix A2 provides the German texts from which the parallelization emerges. In total, German text corpora consisted of 43 words each (11 nouns, 7 verbs and 5 adjectives) for both conditions, presented in five sentences.

Table 2: English translation of the parallelized texts for the *Copying Task*

Text A:	**Text B:**
Bello and the locomotive	Wuffi and the lawn mower
The mother asks her son Felix to go for a walk with Bello now.	The mother asks her son Lukas to mow the lawn today.
But Felix has no desire at all. He prefers to play with his new locomotive. What does Bello do? The poor dog sits impatiently with his leash in front of the door and waits.	But Lukas has no desire at all. He prefers to play with his new football. His faithful dog Wuffi comes into the garden. He sits down and waves his tail joyfully.

2.2.4. General level analysis:

2.2.4.1. Data processing

Data processing was performed separately for the keyboard typing and the handwriting condition. In the typing condition, all keystrokes were logged. Thus, individual letters of each test word were recorded with their corresponding writing times. In the same way, self-corrections (i.e., deleting letters and re-entering them) were documented. In the handwriting condition, writing duration for each grapheme in a test word was measured by determining the time at which the writing process of the grapheme was started and finished.

In summary, spelling accuracy, writing times and number of self-corrections as assessed in both handwriting and typing conditions were entered into the data analysis. In terms of accuracy, spelling was assessed as either correct (scored 1) or incorrect (scored 0) with respect to German spelling rules. Writing times were calculated by subtracting the times for corrections (i.e., the deletion of letters on the keyboard or the manual strike-through) from the overall writing time of the respective words. This procedure was chosen because times of self-corrections differ considerably between handwriting and typing per se. Subsequently, trimmed writing time for each test word was divided by the number of graphemes. This quotient (trimmed writing time/grapheme) entered the analysis as dependent variable for writing time.

2.2.4.2. Descriptive group comparison

In addition to dyslexia diagnostics already obtained for the participating children, between-group differences were examined with respect to general cognitive abilities, language skills as well as RAN. In a first step, the Shapiro Wilk test (Shapiro & Wilk, 1965) indicated that distributions for cognitive ability, working memory and language skills deviated significantly from normal distribution. Accordingly, we used the Wilcoxon sign-rank test (Wilcoxon, 1945) for group comparisons. Bonferroni-Holm procedure (Holm, 1979) was applied to correct for multiple testing. Finally, rank-biserial correlation was computed as a measure of effect size for the Wilcoxon sign-rank test (equivalent to the Mann–Whitney *U* test, Wendt, 1968). The Type I error rate was set to alpha = .05.

2.2.4.3. Mode effect in typical and atypical writing

Mode effect in children with and without dyslexia was investigated using (generalized) linear mixed models (G)LMM) on overall spelling accuracy, self-corrections and writing times using the R package *lme4* (Bates, Maechler, Bolker, & Walker, 2014).

As fixed effects, we entered group (i.e., typically developing and dyslexic children) and mode (i.e., handwriting or typing) at the subject level as well as lexicality (i.e., words and pseudowords) at the item level into the model. As random effects, random intercepts for both subjects and items were included. The following covariates were considered as control variables in the respective models: for both i) spelling accuracy and ii) self-corrections as dependent variables, children's sex and age were added as covariates on a subject level. For iii) writing times as dependent variable, writing duration in the copying task was entered as covariate in addition to sex and age. It is important to note that the copying time has been adapted to reading performance as lower reading performance will affect copy times. For this purpose, text reading performance as assessed by means of the ZLT-II was also added to the model as a covariate. This measure was employed to control for handling of handwriting and typing.

The model selection procedure was as follows: At first, a baseline model was defined controlling independently for variability due to subjects and items (i.e., only error terms for subjects and items were included in the model). In R syntax, the baseline model looked as follows:

```
Outcome ~ (1|subject) + (1|item)
```

Fixed effects were added stepwise to this baseline model depending on significance. Non-significant parameters were excluded from the models. Likelihood-ratio tests were conducted to compare which model describes the data best. Subsequently, covariates were included into the model as additive terms and finally compared with the full model considering all interactions of fixed effects and covariates. The same model selection procedure was repeated under inclusion of the covariates. This procedure was identical for the analysis of all variables of interest (i.e., spelling accuracy, writing times and self-corrections). Hence, the initial hypotheses were tested on the basis of model comparisons.

2.3. Results

2.3.1. Descriptive statistic

Children with and without dyslexia were assessed on general non-verbal intelligence (i.e., assessed by means of processing speed), verbal working memory and language skills (i.e., reading, hyphenation) and RAN. Children with dyslexia, as expected, scored significantly lower in the language tasks (i.e., reading and hyphenation) and in RAN (except for unknown material), but they did not differ significantly in processing speed (see Table 3). They also scored significantly below the control group in verbal working memory, most notably when assessed with lexical material. Table 3 presents group comparisons, statistical details, and effect sizes.

Table 3: Descriptive statistics per group on non-verbal intelligence and language skills.

	Dyslexia		**Controls**			
	Mean	**SD**	**Mean**	**SD**	**W**	**r**
Verbal working memory						
Digit span forward	5.05	1.36	6.16	1.05	505.50**	.53
Digit span backward	3.59	2.64	4.40	2.21	479.50*	.45
Syllable span (pseudowords)	3.36	0.65	4.92	0.80	605.00**	.83
General cognitive ability						
Processing speed (PR)	64.03	15.17	68.84	16.53	395.50	.19
Reading						
Word reading time (PR)	56.10	33.90	82.00	22.79	426.50*	.48
Word reading accuracy (PR)	20.20	22.32	81.33	14.58	643.00**	.94
Pseudoword reading time (PR)	48.45	33.01	81.73	21.90	527.50*	.59
Pseudoword reading accuracy (PR)	12.70	10.48	79.91	17.18	653.50**	.98
Text reading time (PR)	67.05	31.72	97.26	4.33	619.50**	.87
Text reading accuracy (PR)	10.25	10.77	68.83	20.09	653.00**	.97
Hyphenation						
Verbal (PR)	48.48	33.02	73.26	24.49	470.00*	.42
Written (PR)	29.63	23.39	77.16	18.91	607.50**	.84
Rapid automatized naming (RAN)						
Known material (PR)	67.50	27.35	88.43	11.76	477.50*	.44
Unknown material (PR)	90.86	22.44	95.86	11.40	377.50	.14
Letters (sec)	20.59	4.21	16.93	4.01	159.50**	.52
Numbers (sec)	33.27	4.60	27.70	5.46	149.50*	.54

Note: Table 3 provides mean percentage ranges (PR) of the standardized (sub-)tests for processing speed and language skills. Accuracy measures in the Zürcher Lesetest II are given in percentile bands. To compare performances of both groups, the mean of the respective percentile bands was calculated. Contrasts (e.g., verbal working memory, reading, etc.) were corrected for multiple testing using the Bonferroni Holm procedure. * $p<.05$ and ** $p<.001$.

2.3.2. Test sheet comparison

Prior to more detailed analyses, a two-way analysis of variance (ANOVA) was conducted to compare whether children's spelling performance (i.e., accuracy) differed between the mode of testing assessed by means of the two test sheets. Results indicated that the average number of correctly written words collected by hand and on the computer did not differ between test sheet A (handwriting: M=34.55, SD=9.58 and typing: M=32.77, SD=9.40) and test sheet B (handwriting: M=34.82, SD=10.60 and typing: M=35.97, SD=10.69), $F(2,101) = .34$, p=.71. Both test sets were thus regarded as equally difficult in the subsequent analyses. Furthermore, we used Bayesian methods as described in detail by Masson (2011) to evaluate null effects. Bayesian analysis for set as well as for mode indicated a posterior probability of .88 for set and .86 for mode. Thus there is positive evidence for the null hypothesis indicating comparable test material.

2.3.3. Mode effect in typical and impaired writing

An overview of mean group differences for writing mode and lexicality in terms of spelling accuracy, self-correction, and writing times is given in table 4.

Table 4: Descriptive statistics per group on writing and spelling performance.

	Mode	**Lexicality**	**Accuracy** (in %)	**Writing Times** (ms/grapheme)	**Self-Corrections** (quantity)
Dyslexia	Writing	Word	54.81 (49.80)	565.14 (249.49)	11.22 (31.59)
		Pseudo	41.59 (49.34)	537.65 (234.47)	8.63 (28.12)
	Typing	Word	54.67 (49.81)	958.06 (781.84)	20.58 (40.46)
		Pseudo	39.31 (48.90)	959.22 (722.73)	15.45 (36.18)
Controls	Writing	Word	81.64 (38.72)	554.64 (223.78)	3.72 (18.94)
		Pseudo	69.16 (46.21)	519.14 (223.54)	4.16 (19.99)
	Typing	Word	79.31 (40.52)	792.75 (635.02)	18.13 (38.55)
		Pseudo	64.51 (47.89)	855.45 (521.97)	12.66 (33.28)

Note: Table 7 provides group means and standard deviations in parentheses for spelling accuracy (in %), writing times per grapheme (in ms) and mean absolute number of self-corrections.

In the following, results of separate hierarchical (G)LMM are presented. Table 5 (Panel A-C) shows the overall terms considered in the models and their parameter values in the final models. Figure 1 (Panel A-C) shows mean fixed effects coefficients from the final model for each group as well as the mean group differences from the original data. Appendix A3 describes the model selection procedure in detail.

For *spelling accuracy,* likelihood-ratio tests revealed the best fit for the model including group, mode and lexicality as fixed effects as well as sex and age as covariates. Estimates, standard errors, Z-statistics and *p*-values of the final model are given in Table 5A. As can be seen in the table, group (β=-1.66) and mode (β=-0.16) significantly predicted lower spelling accuracy; children with dyslexia wrote less accurately compared to their peers. Moreover, typing at the computer keyboard resulted in more spelling errors compared to handwriting in all children. At the item level, writing words was significantly easier than writing pseudowords as indicated by a main effect of lexicality (β=0.89). No significant interactions were found between model coefficients. Furthermore, sex (β=-0.73) and age (β=0.25) were found to significantly influence spelling accuracy: in both populations, male performed less accurately than female children and older children spelled more accurately than younger children.

Table 5: Analysis of spelling accuracy, self-corrections and writing times (A-C): parameter values for fixed effects and covariates added to the final model

A)	**Spelling accuracy**	**Estimate**	**SE**	***Z***	***P***	
	Intercept	-1.37	1.15	-1.19	0.23	
	Sex (male)	-0.73	0.22	-3.30	<.001	
	Age	0.25	0.10	2.51	0.01	
	Group (dyslexia)	-1.66	0.22	-7.51	<.001	
	Lexicality (words)	0.89	0.27	3.26	<.001	
	Mode (typing)	-0.16	0.07	-2.20	0.03	
B)	**Self-Corrections**					
	(Intercept)	-6.20	1.73	-3.59	0.00	
	Mode (typing)	2.80	1.68	1.66	0.10	
	Group (dyslexia)	3.63	2.13	1.70	0.09	
	Lexicality (words)	0.33	0.10	3.21	0.00	
C)	**Writing times**	**Estimate**	**SE**	***t-value***	**Confidence interval**	
					Lower	Upper
	(Intercept)	702.10	488.58	1.44	-227.73	1631.60
	Copy time	0.68	0.18	3.86	0.34	1.04
	Text reading	-1.35	1.46	-0.92	-4.12	1.42
	Age	-13.36	41.06	-0.33	-91.55	64.75
	Mode (typing)	2066.63	277.86	7.44	1520.15	2608.68
	Age * Mode	-162.79	24.45	-6.66	-210.47	-114.70

Note: Only significant parameters resp. interactions from the preferred model are presented in the table.

For *self-corrections*, likelihood-ratio tests showed that the model that includes group, mode, and lexicality as fixed effects predicted the data best. Table 5B provides model estimates, standard errors, Z-statistics and *p*-values. The number of self-corrections was significantly influenced by lexicality; children self-corrected more frequently when writing words compared to when writing pseudowords (*β*=0.33). Mode (*β*=2.80) and group (*β*=3.36) tended to affect the number of self-corrections in such a way that typing was associated with more self-correction behavior. Children with dyslexia tended to revise more often. Again, no significant interactions were found between model coefficients. Adding age and sex as covariates did not improve model fit.

For *writing times*, linear mixed models that included group, mode, and lexicality as fixed effects as well as age and copying time as covariates showed a good fit to the data. The base model was defined by including the copy times adapted for text reading performance. Consideration of sex in the model did not improve the fit. Yet, the full model considering all interactions of these effects revealed a better fit. Crucially, for age and reading performance a four-way interaction was found with group and copy time. This finding suggested that there might be additional influences on writing times which have to be considered. However, interpretation of those interactions does not provide an unambiguous interpretation.

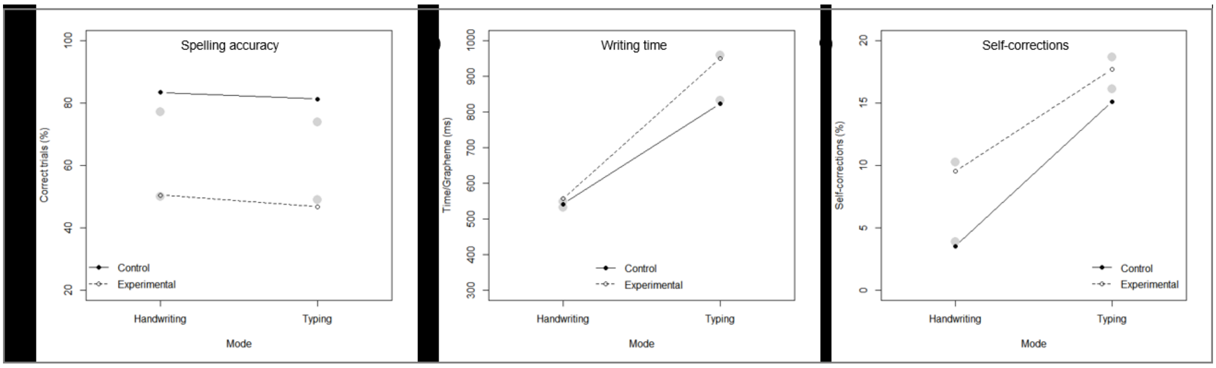


Fig. 1: Group differences for mean fixed effect coefficients and original data. *Note:* Group means from original data without consideration of covariates for spelling accuracy (A), writing times (B) and self-corrections (C) are depicted as gray dots.

Table 8C provides model estimates, standard errors, *t*-statistic and confidence intervals for the more parsimonious significant model according to the likelihood ratio tests. Again, significant effects are presented in the table. Model coefficients showed that typing mode significantly predicted longer writing times (*β*=2066.63) meaning that typing on the computer keyboard was on average about 2 seconds slower per letter than typing by hand. Writing and typing times did not differ significantly between groups, but children who worked more slowly on the copying task also showed longer writing times in writing-to-dictation (*β*=0.68). Additionally, a significant interaction was found between age and mode (*β*=-162.79). This interaction indicated that for older children the mode effect (i.e., increased writing times when typing on the computer keyboard) was smaller than for younger children.

2.4. Discussion

In the first step, our goal was to compare spelling performance of typically developing children and children with dyslexia within two different writing contexts: handwriting and typing on the computer keyboard. In the following, we will discuss the influence of writing mode on spelling accuracy, number of self-correction and writing times in turn.

Most importantly, our results substantiated that writing mode affected spelling accuracy in a writing-to-dictation task in both populations. Typing on the keyboard was more error prone than handwriting resulting in significantly reduced spelling accuracy for all children. These results are consistent with findings from Berninger and colleagues (2009) for children with dyslexia. Yet, they are inconsistent with previous results from Frahm (2013) who also used a writing-to-dictation paradigm to assess spelling performance in typically developing 5th graders but did not find any mode effect. However, the present study and the study by Frahm (2013) are not directly comparable as we used a within-subject design whereas Frahm (2013) used a between-subject design. Therefore, the two findings are not necessarily mutually exclusive: on the one hand, between-subject designs have less power to detect small effects as compared to within-subject designs (Lottridge, Nicewander, Schulz, & Mitzel, 2008). For this reason, it is conceivable that an existing mode effect could not be captured in the study design by Frahm (2013). On the other hand, qualitative error analysis as carried in the present study and the study by Frahm (2013) , revealed that some test words seem to be more responsive to mode effects than others (e.g., Frahm, 2013; Kröhne & Martens, 2011). Possibly, item selection in the two studies differed in this respect.

Apart from the influence of writing mode and also consistent with previous results, in both populations, male children were found to perform less accurately (e.g., Berninger & Fuller, 1992; Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008a). Furthermore, spelling errors significantly decreased with age. Older children spelled more accurately than younger children which again fits nicely with the previous literature (for children: Russell, 1999; for adults: Salthouse, 1984; Westerman & Davies, 2000). Additionally, for all children, it was more difficult to write pseudowords correctly as compared to real words. Pseudowords have structural properties similar to real words but do not access semantics (see Binder, Desai, Graves, & Conant, 2009, for a review on neurofunctional evidence). Therefore, unlike real words, their spelling cannot be generated lexically. Instead, spelling pseudowords relies on the knowledge of the correspondence between phonemes and graphemes (based on the dual-route model, e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). In German, this correspondence is rather clear (Landerl et al., 2013). Nevertheless, there is convincing evidence that sub-lexical processes are more error-prone than lexical retrieval (e.g., Landerl, 1997), also because pseudowords are more difficult to be auditorily encoded and thus pose higher cognitive demands (Newman & Twieg, 2001).

Finally, and in line with our expectations, children with dyslexia exhibited significantly more spelling mistakes reflecting symptoms of developmental dyslexia (e.g., Berninger, Nielsen, et al., 2008; Lyon, Shaywitz, & Shaywitz, 2003).

On the other hand, the number of self-corrections was not influenced significantly by writing mode. Nevertheless, there was a trend indicating that children tended to self-correct more often when typing on the computer keyboard. This trend is in line with results of the meta-analysis by Goldberg and colleagues (2003). Our results were also ambiguous regarding the differences between groups. Overall, the number of self-corrections did not differ significantly between groups (but again, there was a trend towards more self-corrections in children with dyslexia). Morken and Helland (2013) found similar results and concluded that although children with dyslexia revised in the same manner as typically developing children, the final product was of poorer spelling quality. Crucially, the number of self-corrections was influenced significantly by lexicality, i.e., children self-corrected real words more frequently than pseudowords. One explanation may again stem from different lexical and sub-lexical processing pathways: for real words, the spelled word can be compared to the corresponding lexical entry, and thus, spelling errors may be detected more easily. Furthermore, additional semantic information may facilitate retrieval processes (Howard & Gatehouse, 2006). For pseudowords, only a more demanding comparison with the auditory input is possible. Also, the lack of a correct word form may cause children to correct pseudowords less frequently.

Writing times in the present study were affected significantly by writing mode and specifically increased for typing. This increment decreased significantly with age, and thus with growing computer experience, which is in line with previous findings (Goldberg et al., 2003; Russell, 1999). Interestingly, there were no significant group differences in writing times for typing on the computer keyboard and handwriting. The latter result is consistent with previous findings in 8 to 10 years old children (Martlewm, 1992), but contradicts results from Sumner et al. (2013) who assessed primary school children (mean age: 9 year and 4 months). Both studies examined writing at different linguistic levels (word vs. grapheme vs. text level), which might explain the differing results. Furthermore, rather high standard errors in writing time were observed particularly in the typing condition. This seems to substantiate critical influences of computer experience. In our samples, some children received typing instructions in school while others did not. It was observed that different typing instructions can reinforce the mode effect on writing times (Connelly et al., 2007). For writing assessments on the word level, this finding is not as decisive as on other linguistic levels (e.g., text level). Nevertheless, previous typing instruction should be considered in subsequent studies.

Taken together, our data provided converging evidence with most previous studies that writing of typically developing children and children with dyslexia is equally affected by mode of writing. This effect impacted significantly spelling accuracy (but see Frahm, 2013) and writing times, but less the number of self-corrections (Berninger et al., 2009, Goldberg, Russell, Cook, 2003; Wollscheid, Sjaastad, & Tømte, 2016). It has also been shown that age alters mode effects on writing times. Apart from writing mode, lexicality predicted spelling accuracy and self-correction behaviour significantly in both populations. Crucially, qualitative error analyses revealed differences between test words and spelling rules in terms of error rates and the potential influence of mode. The latter observation prompted the second step of analyses.

3. Second step: Consideration at rule specific level

3.1. Objectives

In the first step of analyses, we observed that writing mode influenced writing and spelling performance of both typically developing children and children with dyslexia. However, it seemed that different test words and more specifically different spelling rules were not affected in a comparable way. Importantly, spelling rules used in our study put different demands on additional underlying cognitive processes in terms of correct spelling (i.e., application of rule and morphological knowledge, realization of consonant doubling or lengthening signs demands phonological awareness, writing of capital letters on the computer keyboard requires an additional motoric movement). In our second step of analyses, we aimed at evaluating whether writing mode modulates those additional writing processes differentially.

3.2. Methods

3.2.1. Data Processing

At a rule-specific level, the mode effect on children's writing performance was evaluated separately for the following spelling rules reflected by test words: capitalization (34 words), consonant doubling (16 words), lengthening (7 words), rule words (i.e., words for which the spelling cannot be derived by the phoneme structure; 8 words), and final obstruent devoicing (6 words). For each test word it was assessed whether the relevant spelling rule was correct or incorrect. For instance, 'voll' [engl. full] was assessed concerning (inner sentence) capitalization and consonant doubling, while the other spelling rules did not apply to this particular test word. With respect to consonant doubling, for instance, it was assessed whether the double consonant was realized correctly irrespective of whether the overall spelling was correct (i.e., whenever the test word 'voll' [engl. full] was written like 'foll', double consonant spelling was assessed as correct and scored 1, although the overall spelling was not correct, but a spelling such as 'vol' was considered incorrect and scored 0).

Similar (general) linear mixed model analyses were performed as in study 1. Again, spelling accuracy, number of self-corrections and writing times were examined for each specific German spelling rule. Only test words were considered for which the relevant rule was applicable (e.g., capitalization applied to all test words (n=34), whereas consonant doubling appeared in 16 test words).

The same stepwise data modeling procedure was applied: first, by defining a baseline model for each spelling rule; second, by including the fixed effects on subject and item level; and third, by including covariates into the model as additive terms. The final model was then compared with the full model considering all interactions of fixed effects and covariates. Appendix A4 shows the model selection procedure in detail.

3.3. Results

For *spelling accuracy*, the most striking result was that writing mode significantly predicted lower spelling accuracy only for capitalization (β=-0.61). Crucially, writing mode had no significant influence on all other spelling rules. This result was supported by two other findings for capitalization: first, a significant main effect for word class (nouns vs. other words [β=3.81]) indicated that writing noun words was significantly more difficult in both conditions. This effect was even more pronounced for pseudowords as suggested by the significant interaction of lexicality and word class (β=-1.38). Second, the significant interaction of mode with word class

(β=1.40) indicated that writing nouns was even more difficult when typing. Only for capitalization, the full model yielded a better fit.

Apart from that, similar results were found in the different spelling rules: Group significantly predicted lower spelling accuracy for all spelling rules (capitalization: β=-0.99, consonant doubling: β=-1.89, lengthening: β=-1.90 and rule words β=-1.26 except for devoicing: β=-0.61). Children with dyslexia wrote less accurately than control children. Furthermore, sex (capitalization: β=-0.59, consonant doubling: β=-1.02, lengthening: β=-0.86 and rule words β=-0.72) and age capitalization: β=0.19, consonant doubling: β=0.39, lengthening: β=0.36 and rule words β=0.28) were found to influence spelling accuracy, but not for devoicing: in both populations, male children performed less accurately, and older children spelled more accurately than younger children.

At the item level, writing words was significantly easier than writing pseudowords as indicated by the significant main effect of lexicality in all spelling rules (capitalization: β=1.82, consonant doubling: β=2.78, lengthening: β=3.28), again except for devoicing (β=0.74). Model estimates, standard errors, Z-statistics and *p*-values for each spelling rule are given in Table 6A.

For *self-corrections*, rule specific analyses revealed a significant main effect of mode on the number of self-corrections (i.e., self-corrections increased when typing on the computer keyboard) for all spelling rules (capitalization: β=0.98, consonant doubling: β=0.82, lengthening: β=2.80 and rule words β=1.01) except for devoicing (β=1.86). Children self-corrected more when typing on the computer keyboard. For capitalization two specific results were found: first, self-corrections behavior was significantly influenced by lexicality meaning that children self-corrected more frequently when writing words as compared to writing pseudowords (β=0.49). Second, children self-corrected less when writing non-nouns (β=-1.50). Interestingly, sex (β=9.23) was found to influence self-correction behavior in rule words: male revised more often than female children. Table 9B gives model estimates, standard errors, *t*-statistics and confidence intervals for the specific spelling rules.

Table 6: Rule specific analysis of spelling accuracy, self-corrections and writing times (A-C): significant parameter values for fixed effects and covariates added to the final models.

	Capitalization					Consonant doubling					Lengthening					Devocing					Rule words				
A) Spelling accuracy	**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***	
Intercept	-1.16	1.49	-0.78	0.44		-2.93	1.54	-1.90	0.06		-2.60	1.35	-1.92	0.05		3.54	0.67	5.32	0.00		-1.40	1.73	-0.81	0.42	
Sex (male)	-0.59	0.29	-2.06	0.04		-1.02	0.29	-3.55	0.00		-0.86	0.26	-3.31	0.00		-	-	-	-		-0.72	0.33	-2.19	0.03	
Age	0.19	0.13	1.48	0.14		0.39	0.13	2.98	0.00		0.36	0.12	3.12	0.00		-	-	-	-		0.28	0.15	1.88	0.06	
Mode (typing)	-0.61	0.16	-3.69	0.00		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
Group (dyslexic)	-0.99	0.28	-3.49	0.00		-1.89	0.29	-6.57	0.00		-1.60	0.32	-5.04	0.00		-0.61	0.62	-0.97	0.33		-1.26	0.33	-3.79	0.00	
Lexicality (words)	1.82	0.40	4.61	0.00		2.78	0.53	5.25	0.00		3.28	0.58	5.70	0.00		0.74	0.74	0.99	0.32		-	-	-	-	
Word class (non nouns)	3.81	0.47	8.08	0.00		-	-	-	-		-	-	-	-							-	-	-	-	
Lexicality * Group	-	-	-	-		-	-	-	-		-1.44	0.46	-3.13	0.00		-1.37	0.67	-2.05	0.04		-	-	-	-	
Lexicality * Mode	0.49	0.23	2.16	0.03		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
Lexicality * WordClass	-1.38	0.61	-2.27	0.02		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
WordClass * Mode	1.40	0.41	3.40	0.00		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
B) Self-corrections	**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***		**Est.**	**SE**	***Z***	***p***	
(Interc.)	-4.50	0.28	-16.37	0.00		-4.83	0.38	-12.83	0.00		-6.63	1.04	-6.38	0.00		-14.73	5.47	-2.69	0.01		-6.31	2.57	-2.45	0.01	
Sex (male)	-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		9.23	4.00	2.31	0.02	
Age	-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		0.21	0.22	0.95	0.34	
Sex * Age	-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		-0.82	0.35	-2.32	0.02	
Mode	0.98	0.20	4.80	0.00		0.82	0.38	2.17	0.03		2.80	1.00	2.80	0.01		1.86	1.63	1.14	0.25		1.01	0.35	2.91	0.00	
Lexicality (words)	0.49	0.21	2.34	0.02		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
Word class (non nouns)	-1.50	0.22	-6.73	0.00		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-	
C) Writing time	**Est.**	**SE**	***t***	***ConfInt***		**Est.**	**SE**	***t***	***ConfInt***		**Est.**	**SE**	***t***	***ConfInt***		**Est.**	**SE**	***t***	***ConfInt***		**Est.**	**SE**	***t***	***ConfInt***	
(Intercept)	1248.84	410.28	3.04	452.37	2045.91	1007.48	413.52	2.44	210.96	1803.64	949.32	374.19	2.54	222.72	1676.97	922.02	315.05	2.93	312.18	1532.90	1064.90	580.92	1.83	-60.11	2191.37
Copy time	1.33	0.23	5.69	0.87	1.79	1.32	0.23	5.64	0.87	1.78	1.58	0.42	3.75	0.76	2.44	1.20	0.33	3.69	0.56	1.86	1.25	0.60	2.09	0.08	2.46
Text reading	-1.69	1.31	-1.29	-4.23	0.86	-1.32	1.32	-1.00	-3.86	1.22	-1.49	1.08	-1.38	-3.61	0.61	-1.48	0.99	-1.50	-3.40	0.44	-1.71	1.81	-0.95	-5.22	1.80
Age	-66.17	32.77	-2.02	-129.81	-2.51	-51.96	33.63	-1.55	-116.73	12.80	-45.13	29.52	-1.53	-102.51	12.25	-46.02	25.23	-1.82	-94.86	2.87	-57.80	46.55	-1.24	-147.93	32.42
Mode (typing)	1936.94	243.27	7.96	1459.44	2412.75	2016.33	269.79	7.47	1486.39	2543.68	1562.86	353.38	4.42	867.05	2252.18	1438.95	262.49	5.48	921.91	1951.90	1525.50	483.99	3.15	571.88	2470.02
Age * Mode	-137.10	21.17	-6.48	-178.50	-95.54																				
Lexicality * Mode	99.16	33.24	2.98	33.95	164.19	-149.84	23.57	-6.36	-195.90	-103.54	-104.22	30.87	-3.38	-164.43	-43.44	-99.55	22.95	-4.34	-144.36	-54.34	-98.93	42.30	-2.34	-181.44	-15.57

Note: Only significant parameters resp. interactions from the preferred model are presented in the table.

For *writing times*, linear mixed model analyses revealed consistent results for the different spelling rules. Writing mode significantly predicted longer writing times (capitalization: β=1936.94, consonant doubling: β=2016.33, lengthening: β=1562.86, devoicing: β=1438.95, and rule words β=1525.50). Typing on the computer keyboard was on average about 1500 – 2000ms per letter slower than writing by hand. Writing and typing times did not differ significantly for the two participant groups and sex, but children who worked more slowly on the copying task also showed longer writing times in writing-to-dictation (capitalization: β=1.33, consonant doubling: β=1.32, lengthening: β=1.58, devoicing: β=1.20, and rule words β=1.25). Moreover, age tended to influence writing times as older children tended to write faster than younger children (capitalization: β=-66.17, consonant doubling: β=-51.96, lengthening: β=-45.13, devoicing: β=-46.02, and rule words β=-57.80). Additionally, a significant interaction was found between age and mode for capitalization (β=-137.13). This interaction indicated that for older children the mode effect (i.e., increased writing times when typing on the computer keyboard) was about 130 ms second smaller than for younger children. Finally, significant interactions between mode and lexicality were found. Crucially, these interactions differed in direction between capitalization (β=99.16) and the other spelling rules (consonant doubling: β=-149.48, lengthening: β=-104.22, devoicing: β=-99.55, and rule words β=-89.83) indicating that for real words the mode effect was longer for capitalization but shorter for the rest.

Regarding likelihood ratio tests, only for capitalization did the full model considering all interactions of the model terms yield a better fit. For all other spelling rules, model terms included in our simpler model might be sufficient to explain empirical data. Table 9C provides model estimates, standard errors, *t*-statistics, and confidence intervals for the specific spelling rules.

3.4. Discussion:

With the second step of analyses we aimed at evaluating the influence of writing mode on a more fine-grained analytic level. In order to provide a more detailed picture about the influence of writing mode on additional underlying processes related to writing (e.g., phonological awareness, rule and strategy knowledge assessed by different spelling rules), specific spelling-rules were analyzed separately. Results of our first step of analyses indicated a mode effect present for spelling accuracy and writing times: independent of impaired or unimpaired written language development, children made more spelling mistakes and took more time to write on the computer keyboard than when writing by hand.

In terms of spelling accuracy, the most striking result of our second step of analyses was that the mode effect observed in the first step was evident only for capitalization but not for the other spelling rules. Two explanations seem conceivable for this matter of affairs: First, capitalization of letters is one of the most significant difficulties of the German language and the most common source of errors (e.g., Günther & Nünke, 2005). Also, Frahm (2013) observed that in her study most spelling errors were case-sensitive. Second, writing capital letters on the computer differs significantly from all other spelling rules (i.e., by pressing the additional 'Shift' key). This additional key press might affect children's spelling performance. Consonant doubling and lengthening, in contrast, which are based especially on phonological awareness (Moll et al., 2009), were not affected by writing mode. Hence, spelling accuracy did not differ between handwriting and typing on the computer keyboard for all other spelling rules. The same was true for morphological and rule knowledge as assessed with final obstruent devoicing (i.e., requiring a derivation of the word) and rule words, respectively.

Furthermore, results replicated findings from the first step regarding spelling accuracy: in both populations, typing on the computer keyboard was more error prone than handwriting which resulted in significantly reduced spelling accuracy for all children. Children with dyslexia performed poorer with respect to each spelling rule, which is in line with previous findings (V. W. Berninger et al., 2009; Cidrim & Madeiro, 2017). In the analysis of self-corrections from a rule specific perspective, a mode effect was present for all spelling rules except for final obstruent devoicing. Children self-corrected more when typing on the computer keyboard than when writing by hand (Goldberg et al., 2003). Moreover, children with dyslexia did not differ from controls in their self-corrections behavior, which is in line with findings on typing from Morken & Helland (2013). Apart from that, an influence of sex and age on self-corrections was only apparent for rule words indicating that boys revised rule words more often than girls. However, this difference declined with age. Sex differences in language skill were found repeatedly. Neuroimaging results suggested that boys and girls rely on different brain areas for accurate performance on language tasks (Burman et al., 2008). The authors found that boys rely on a modality specific (i.e., visual or auditory) network whereas girls used supramodal language networks, representing a more abstract, conceptual knowledge of words. This distinction might be one explanation why boys made more self-corrections in rule words: writing of rule words requires a direct writing strategy, this means the retrieval of orthographic information stored in a long-term orthographic lexicon (Ellis, 1982). Sub-lexical writing strategies, for instance, by phoneme-grapheme correspondence, are not successful in most cases. Yet, they are used when there is no lexical entry for the target word, or to match the lexical entry with the spelling outcome (Günther & Ludwig, 1994). Different from girls, boys might rely on phoneme-grapheme correspondence when they are not entirely sure about the correct spelling of a target word. Subsequently, they self-correct more in order to match the word to the lexical entry. Additionally, older children revised less-frequently on rule words, probably because the number of lexical entries increases with age (e.g., Nippold, 2002).

In terms of writing times, the present rule-specific results corroborated previous findings for the mode effect (Goldberg et al., 2003; Wollscheid, Sjaastad, & Tømte, 2016). Writing times in all spelling rules were significantly affected by writing mode and specifically increased for typing. Again, this increment significantly decreased with age, and thus with growing computer experience (Goldberg et al., 2003; Russell, 1999). Comparable to the results of our first step of analyses, significant group differences occurred neither for typing on the computer keyboard nor for handwriting indicating that writing fluency of children with and without dyslexia is influenced by the medium used for writing in a comparable way.

4. General Discussion

In the current article, we evaluated handwriting and typing skills in typically developing children and children with dyslexia from a more general level and from a rule-specific level directly comparing the mode effect (i.e., the difference between handwriting and typing) between typically and atypically developing children.

At first glance, the two steps of analyses provided consistent results. Writing mode influenced writing time (i.e., speed) on both the holistic and rule specific level of evaluation. Children wrote more slowly on the computer keyboard than by hand, which corresponds to previous research (i.e., Berninger, Abbott, Augsburger, & Garcia, 2009; Connelly, Gee, & Walsh, 2007; Wollscheid, Sjaastad, & Tømte, 2016). This effect diminished with age and, probably, with growing computer experience (Goldberg et al., 2003; Russell, 1999). Crucially, this finding applied to both typically developing children and children with dyslexia alike. Writing mode also affected children`s self-corrections behavior with typing being more error-prone, as the number of self-corrections was increased for typing. Participating children showed a similar self-correction behavior as observed be Goldberg and colleagues (2003). They revised online during the writing process rather than at the end of the word. A significant impact of writing mode on children’s spelling performance was also apparent. In line with previous results, typing on the keyboard reduced spelling accuracy for typically developing children (but see Frahm, 2013) and children with dyslexia (e.g., Berninger et al., 2009). At second glance, however, rule-specific analysis revealed that this effect was only present for capitalization. Crucially, writing capital letters on the computer keyboard differs from writing of any other letter (i.e., by pressing the additional 'Shift' key). In all other spelling rules, writing mode did not affect spelling accuracy suggesting that additional processes related to writing (e.g., phonological awareness, rule and strategy knowledge as assessed by different spelling rules) were not affected significantly by medial influences.

In standard spelling assessments using a writing-to-dictation task, either in a school or a therapeutic context, each test word is assessed regarding capitalization. Thus capitalization has an enormous impact on the evaluation of overall spelling performance. In our first step of analyses, the significant mode effect may have been caused by the capitalization of letters: capitalization seemed to superimpose other rules which in turn might have led to an overall effect of writing mode on spelling accuracy. Only the rule-specific analysis of our second step made it possible to disentangle the mode effects on writing.

This result is of particular importance in two respects: first, writing accuracy from both typically developing children and children with dyslexia may not be subject to a general mode effect. Instead, the mode effect seemed specifically caused by the peculiarity of typing upper case letters on the computer keyboard. Furthermore, writing mode did not interact with other effects on spelling accuracy (i.e., such as with the presence of dyslexia, age, or the type of word material) suggesting a very specific influence. Second, and of particular importance for the application of digital media in school assessments, digital learning environment or for therapeutic aspects, computers and paper-and-pencil tests seem equally suitable to assess writing skills in children. This statement is based on the premise that spelling performance is taken into account in a differentiated approach, as for example from a rule-specific perspective. A rule-specific analysis of spelling accuracy avoids generalization from one spelling rule to others. This is particularly important for the development of intervention programs tailored to fit individual`s needs. Above that, the diagnostic effort is low, but the benefit of deriving specific learning contents in school and also in the intervention of writing difficulties is promising.

When interpreting these results, there are some limitations of the present study that need to be considered. First, in the present study, writing-to-dictation was assessed on the word level. As such, we did not evaluate text generation or planning processes. We are therefore cautious about the generalization of our results to the level of text generation, even though they are in line with the recent literature (e.g., Connelly et al., 2007; Wollscheid, Sjaastad, Tømte, et al., 2016; but see Sumner et al., 2013). As such, it would be desirable to further pursue the intra-individual effect of writing mode on these higher lexical levels. Second, we made conclusions based on the age of the children about their possible computer experience. The use of a standardized questionnaire (e.g., INCOBI-R, Richter, Naumann, & Horz, 2010) would have been more advantageous. However, conclusions on computer experience seem warranted as keyboard writing times decreased with children's age. Further intra-individual variables identifying influences of writing mode in using other linguistic tasks and additional control measures could provide further insights.

A final assumption might suggest that the mode effect varies in different languages. In German, where the use of capital letters is very common compared to other languages, this influence may be much more pronounced in comparison to other languages. Future cross-linguistic studies have the potential to address this issue.

5. Conclusion

Digital technology has changed writing habits to such an extent that the question arises whether it needs to be taken specifically into account when assessing writing and spelling skills in educational and therapeutic contexts. In this article, we investigated the influence of writing mode (i.e., handwriting vs. typing) on spelling performance in typically developing children and children with dyslexia at a general as well as a rule-specific level. At both levels of analysis, we found that children of the same age group (i.e., secondary school children) wrote comparably fast by hand and computer keyboard irrespective of whether or not they had writing problems. Additionally, both groups showed slower writing speed on the computer. Mode effect in self-correction behavior was more pronounced when assessed at the rule-specific level. Again, children with dyslexia did not differ from controls.

The sole difference between children with and without dyslexia was their spelling accuracy. However, while recognizing that writing mode affects spelling accuracy from a holistic perspective, we suggest that evaluation of children's spelling accuracy should focus on individual spelling rules, and thus on different writing processes.

Acknowledgements

Stefanie Jung was supported by the first Leibniz-WissenschaftsCampus Tübingen “Bildung und Informationsumwelten” as well as by the German Research Foundation (DFG) providing funding to Elise Klein (KL 2788/2-1). Elise Klein was supported by the Margarete von-Wrangell Fellowship (European Social Fund and the Ministry of Science, Research and the Arts Baden-Württemberg). We thank the participating children, their parents and institutions who made this work possible. We also thank all research assistants for their assistance with data collection.

References

Alamargot, D., Chesnet, D., Dansac, C., & Ros, C. (2006). Eye and pen: A new device for studying reading during writing. *Behavior Research Methods*, *38*(2), 287–299.

Alexander-Passe, N. (2006). How dyslexic teenagers cope: An investigation of self-esteem, coping and depression. *Dyslexia*, *12*(4), 256–275.

Anderson, J. R. (2005). *No TitleCognitive psychology and its implications*. Macmillan.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version*, *1*(7), 1–23.

Berninger, V. W., Abbott, R. D., Augsburger, A., & Garcia, N. (2009). Comparison of Pen and Keyboard Transcription Modes in Children with and without Learning Disabilities. *Learning Disability Quarterly*, *32*(3), 123–141.

Berninger, V. W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology*, *30*(4), 363–382.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008a). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology*, *46*(2), 151–172.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008b). Writing problems in developmental dyslexia: Under-recognized and under-treated. *Journal of School Psychology*, *46*(1), 1–21.

Berninger, V. W., Winn, W. D., Stock, P., Abbott, R. D., Eschen, K., Lin, S.-J. C., … others. (2008). Tier 3 specialized writing instruction for students with dyslexia. *Reading and Writing*, *21*(1–2), 95–129.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.

Chesnet, D., & Alamargot, D. (2005). Analyse en temps réel des activités oculaires et grapho-motrices du scripteur: intérêt du dispositif {\guillemotleft}Eye and Pen{\guillemotright}. *L'année Psychologique*, *105*(3), 477–520.

Cidrim, L., & Madeiro, F. (2017). Studies on spelling in the context of dyslexia: a literature review. *Revista CEFAC*, *19*(6), 842–854.

Coleman, C., Gregg, N., McLain, L., & Bellair, L. W. (2009). A Comparison of Spelling Performance Across Young Adults With and Without Dyslexia. *Assessment for Effective Intervention*, *34*(2), 94–105.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204.

Connelly, V., Gee, D., & Walsh, E. (2007). A comparison of keyboarded and handwritten compositions and the relationship with transcription speed. *British Journal of Educational Psychology*, *77*(2), 479–492.

Deimel, W. (2002). Testverfahren zur Diagnostik der Lese-Rechtschreibstörung - eine Übersicht. *Legasthenie: Zum Aktuellen Stand Der Ursachenforschung, Der Diagnostischen Methoden Und Der Förderkonzepte*, (Rodgers), 149–160.

Döhla, D., & Heim, S. (2016). Developmental dyslexia and dysgraphia: What can we learn from the one about the other? *Frontiers in Psychology*, *6*(JAN), 1–12.

Ellis, A. W. (1982). Spelling and writing (and reading and speaking). *Normality and Pathology in Cognitive Functions*, 113–146.

Gaskill, J. L., & Marshall, M. (2007). *Comparisons Between Paper-and Computer-based Tests: Foundation Skills Assessment-2001 to 2006 Data*. Society for the Advancement of Excellence in Education.

Goldberg, A., Russell, M., & Cook, A. (2003a). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, *2*(1).

Goldberg, A., Russell, M., & Cook, A. (2003b). The Effect of Computers on Student Writing: A Meta-Analysis of Studies from 1992 to 2002. *The Journal of Technology, Learning, and Assessment*, *2*(1), 2–51.

Günther, H., & Ludwig, O. (1994). *Schrift und Schriftlichkeit / Writing and Its Use* (Vol. 10.1–). Walter de Gruyter.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Horowitz-Kraus, T., & Breznitz, Z. (2011). Error Detection Mechanism for Words and Sentences: A Comparison between Readers with Dyslexia and Skilled Readers. *International Journal of Disability, Development and Education*, *58*(1), 33–45.

Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, *20*(9), 921–950.

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a sensory–motor brain network. *Neuropsychologia*, *44*(14), 2937–2949.

Jerrim, J. (2016). PISA 2012: how do results for the paper and computer tests compare? *Assessment in Education: Principles, Policy & Practice*, *23*(4), 495–518.

Jung, S., Huber, S., Heller, J., Grust, T., Müller, K., & Nuerk, H.-C. (2016). Die TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE). *Lernen Und Lernstörungen*, *5*(1), 7–15.

Kandel, S., Lassus-Sangosse, D., Grosjacques, G., & Perret, C. (2017). The impact of developmental dyslexia and dysgraphia on movement production during word writing. *Cognitive Neuropsychology*, *34*(3–4), 219–251.

Kröhne, U., & Martens, T. (2011). 11 Computer-based competence tests in the national educational panel study: The challenge of mode effects. *Zeitschrift Für Erziehungswissenschaft*, *14*(S2), 169–186.

Landerl, K. (1997). Word Recognition in English and German Dyslexics: A Direct Comparison. In C. K. Leong & R. M. Joshi (Eds.), *Cross-Language Studies of Learning to Read and Spell: Phonologic and Orthographic Processing* (pp. 121–137). Dordrecht: Springer Netherlands.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., … Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(6), 686–694.

Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *NeuroImage*, *19*(4), 1492–1500.

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., & Velay, J.-L. (2008). Learning through Hand- or Typewriting Influences Visual Recognition of New Graphic Shapes: Behavioral and Functional Imaging Evidence. *Journal of Cognitive Neuroscience*, *20*(5), 802–815.

Longcamp, M., Zerbato-Poudou, M. T., & Velay, J. L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, *119*(1), 67–79.

Lottridge, S., Nicewander, A., Schulz, M., & Mitzel, H. (2008). Comparability of Paper-based and Computer- based Tests : A Review of the Methodology. *CCSSO Technical Issues in Large Scale Assessment Comparability Research Group*.

Lottridge, S., Nicewander, A., Schulz, M., & Mitzel, H. (2008). Comparability of Paper-based and Computer-based Tests: A Review of the Methodology. *Pacific Metrics Corporation*, *585*.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading. *Annals of Dyslexia*, *53*, 1–15.

MacArthur, C. A. (2009). Reflections on Research on Writing and Technology for Struggling Writers. *Learning Disabilities Research & Practice*, *24*(2), 93–103.

Mangen, A., & Velay, J.-L. (2010). Digitizing Literacy: Reflections on the Haptics of Writing. *Advances in Haptics*, 385–403.

Martlewm, M. (1992). Handwriting and spelling: Dyslexic children's abilities compared with children of the same chronological age and younger children of the same spelling level. *British Journal of Educational Psychology*, *62*(3), 375–390.

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690.

May, P. (2002). *Hamburger Schreibprobe (HSP): zur Erfassung der grundlegenden Rechtschreibstrategien*. Verlag für Pädagogische Medien VPM.

Mejía, C., Diaz, A., Jiménez, J. E., & Fabregat, R. (2012). BEDA: A Computerized Assessment Battery for Dyslexia in Adults. *Procedia - Social and Behavioral Sciences*, *46*, 1795–1800.

Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading*, *13*(1), 1–25.

Moll, K., & Landerl, K. (2010). *SLRT-II: Lese-und Rechtschreibtest; Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)*. Huber.

Morken, F., & Helland, T. (2013). Writing in dyslexia: Product and process. *Dyslexia*, *19*(3), 131–148.

Namkung, J. M., & Fuchs, L. S. (2016). Cognitive Predictors of Calculations and Number Line Estimation With Whole Numbers and Fractions Among At-Risk Students, *108*(2), 214–228.

Newman, S. D., & Twieg, D. (2001). Differences in auditory processing of words and pseudowords: An fMRI study. *Human Brain Mapping*, *14*(1), 39–47.

Nicolson, R. I., Fawcett, A. J., & Sheffield, S. (n.d.). Computer-based spelling remediation for dyslexic children using the Selfspell environment.

Nippold, M. A. (2002). Lexical learning in school-age children, adolescents, and adults: a process where language and literacy converge. *Journal of Child Language*, *29*(2), 449–488.

Öngören, S., & Turcan, A. İ. (2009). The Effectiviness Of Montessori Education Method In The Acquisition Of Concept Of Geometrical Shapes. *Procedia - Social and Behavioral Sciences*, *1*(1), 1163–1166.

Pellegrino, J. W., & Quellmalz, E. S. (2010). Perspectives on the integration of technology and assessment. *Journal of Research on Technology in Education*, *43*(2), 119–134.

Petermann, F., Daseking, M., Linder, M., Grissemann, H., & Weid, T. (2013). *Zürcher Lesetest-II: ZLT-II; Weiterentwicklung des Zürcher Lesetests (ZLT) von Maria Linder und Hans Grissemann; Manual*. Huber, Hogrefe.

Petermann, F., & Petermann, U. (2010). *HAWIK-IV: Hamburg-Wechsler-Intelligenztest für Kinder-IV; Manual; Übersetzung und Adaption der WISC-IV von David Wechsler*. Huber.

Raven, J. C. (1962). Coloured Progressive Matrices, Sets A, A_B, B. *HK Lewis*.

Richter, T., Naumann, J., & Horz, H. (2010). Eine revidierte fassung des inventars zur computerbildung (INCOBI-R). *Zeitschrift Fur Padagogische Psychologie*, *24*(1), 23–37.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, *7*(20).

Russell, M., & Plati, T. (2002). Does it matter with what I write? Comparing performance on paper, computer and portable writing devices. *Current Issues in Education*, *5*(4), 1–29.

Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General*, *113*(3), 345–371.

Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading and Writing*, *26*(6), 991–1008.

Vaughn, S., Schumm, J. S., Gordon, J., Vaughn, S., & Schumm, J. S. (1992). Early Spelling Acquisition : Does Writing Really Beat the Computer ? *15*(3), 223–228.

Wang, S., Hong Jiao, Young, M. J., Brooks, T., & Olson, J. (2007). A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, *67*(2), 219–238.

Wang, S., Hong Jiao, Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of Computer-Based and Paper-and-Pencil Testing in K–12 Reading Assessments. *Educational and Psychological Measurement*, *68*(1), 5–24.

Warschauer, M. (1996). Motivational aspects of using computers for writing and communication. *Telecollaboration in Foreign Language Learning*, 29–46. Retrieved from

Wendt, H. W. (1972). Dealing with a common problem in social science: A simplified rank-biserial coefficient of correlation based on the U statistic. *European Journal of Social Psychology*, *2*(4), 463–465.

Westerman, S. J., & Davies, D. R. (2000). Acquisition and application of new technology skills: The influence of age. *Occupational Medicine*, *50*(7), 478–482.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83.

Wollscheid, S., Sjaastad, J., & Tømte, C. (2016). The impact of digital devices vs. Pen(cil) and paper on primary school students' writing skills - A research review. *Computers and Education*, *95*, 19–35.

Wollscheid, S., Sjaastad, J., Tømte, C., & Løver, N. (2016). The effect of pen and paper or tablet computer on early writing - A pilot study. *Computers and Education*, *98*, 77–80.

Appendix A

Appendix A gives an overview of the parallelized test words and sentences used in the two test sets for handwriting and typing.

Table A1: Real words: item set A.

N	Cat.	Test word	Sentence	English translation
P1	N	Zeh	Petra trat mir auf meinen großen ...	Petra stepped on my big ... (toe)
P2	V	rollt	Die Billardkugel ... ins Loch.	The billiard ball ... into the hole. (rolls)
1	N	Pralinen	Oma liebt Süßes, am meisten mag sie ...	Grandma loves sweets, she likes ... the most. (pralines)
2	N	Entlassung	Es ist traurig, wenn einem Arbeiter mit der ... gedroht wird.	It is sad when a worker is faced with a ... (dismissal)
3	N	Zuckerwatte	Auf dem Jahrmarkt essen viele Leute süße ...	At the fair many people eat sweet ...(candyfloss)
4	N	Fernrohr	Die Sterne kann man am besten durch ein ... sehen.	Stars are best seen through a ... (telescope)
5	N	Schimmelkäse	Ein Milchprodukt, das streng riecht, ist ...	A dairy product that smells strong is ... (mouldly cheese)
6	N	Dieb	Die Handtasche wurde von einem ... geklaut.	The purse was stolen by a ... (thief)
7	N	Geburtstag	Meist gibt es einen Kuchen zum ...	Usually there is a cake for your ... (birthday)
8	N	Abfall	Eine Bananenschale ist ...	A banana peel is ... (garbage)
9	N	Nachsitzen	Wenn man zu viel Quatsch macht, drohen manche Lehrer mit ...	Some teachers threaten ... if you do too much silliness. (detention)
10	N	Rechnen	Mit Bruchzahlen ist das ... schwieriger als mit ganzen Zahlen**.**	Fractions are more difficult to ... than integers. (calculate)
11	N	Baby	In der Wiege schläft ein ...	In the cradle sleeps a ... (baby)
12	N	Chips	Ich könnte pfundweise ... essen.	I could eat a pound of ...(chips)
13	N	Portion	Im Sommer isst man gerne eine große ... Eis.	In summer you like to eat a large ... of ice cream. (portion)
14	N	Garage	Zum Schutz steht das Auto in der ...	The car is in the ... for protection. (garage)
15	N	Mus	Die reifen Äpfel kochen wir zu ...	We cook the ripe apples to ... (applesauce)
16	N	Lok	Die Waggons werden von der ... gezogen	The wagons are pulled by the ... (locomotive)
17	V	klappern	Hör auf mit den Töpfen zu ...!	Stop ... with the pots! (clattering)
18	V	schleppt	Fin keucht, weil er die Kisten ...	Fin gasps because he's .... crates. (dragging)
19	V	beißen	Hunde, die bellen, ... nicht.	Dogs that bark don't ... (bite)
20	V	dehnen	Vor dem Sport sollte man die Muskeln	Before doing sports you should ... your muscles. (stretch)
21	V	schaffen	Wenn alle mitmachen können wir es zusammen ...	If all of us cooperate, we can ... it together (do).
22	V	verspritzen	Beim Patronenwechsel bin ich vorsichtig, ich will die Tinte nicht ...	I'm careful when changing cartridges; I don't want to ... the ink. (splash)
23	V	malt	Oma freut sich, wenn Lisa ein Bild für sie ...	Grandma is happy when Lisa ... a picture for her. (draws).
24	V	bieten	Will man eine Auktion gewinnen, muss man hoch ...	If you want to win an auction, you have to ... high. (bid)
25	V	erzählst	Die Geschichten sind immer spannend, wenn du sie	The stories are always exciting when you ... them. (tell)
26	A	grässlich	Das verbrannte Essen schmeckte ...	The burnt food tasted ... (awful)
27	A	gierig	Hungrige Tiere sind ... auf Essen.	Hungry animals are ... for food. (greedy)
28	A	verwirrend	Das Durcheinander ist	The mess is ... (confusing).
29	A	näher	Von Italien nach Frankreich ist es ... als vom Mond zur Erde.	From Italy to France it is ... than from the moon to the earth. (closer)
30	A	entsetzliche	In den Nachrichten kam die ... Meldung von einem Erdbeben.	The news reported the ...news of an earthquake. (terrible)
31	A	stiller	Es ist viel zu laut hier, könnt ihr nicht ... sein?	It's way too loud in here, can't you be ...? (quiet)
32	A	schick	Für die Hochzeit macht sich die Braut ...	For the wedding, the bride ...(dresses up)
33	A	voll	Schenke mir mein Glas bitte ganz ...	Give me a ... glass, please. (full)
34	A	cool	Seinen neuen Haarschnitt findet Hannes richtig ...	Hannes thinks his new haircut is really ... (cool)

Table A2: Real words: item set B.

N	Cat.	Test word	Sentence	English translation
P1	N	Biss	Der Vampir versetzte dem Mädchen einen … in den Hals.	The vampire put a … in the girl's throat. (bite)
P2	V	fiel	Lilo weinte heftig als sie vom Pferd …	Lilo cried violently as she was … from the horse. (fall)
1	N	Mandarinen	Zu Weihnachten gibt es viele Nüsse und …	At Christmas there are many nuts and … (mandarins)
2	N	Überschwemmung	Wenn ein Fluss über die Ufer tritt, gibt es eine …	When a river overflows, there is a …(flood)
3	N	Trockenfutter	Ein Hase bekommt Salat und …	A hare gets lettuce and …(dry food)
4	N	Eselsohr	Eine umgeknickte Seite eines Buches nennt man …	A folded page of a book is called … (dog`s ear)
5	N	Nummernschild	Jedes Auto muss ein …haben.	Every car must have a … (number plate)
6	N	Seitenhieb	Weil ich ihn ärgerte, versetzte mir mein Sitznachbar einen…	Because I annoyed him, my seat neighbor gave me a … (side blow)
7	N	Zwerg	Im Märchen taucht oft ein kleiner … auf.	A small … often appears in fairy tales. (dwarf)
8	N	Nähe	Timo wollte nicht alleine sein, darum war immer jemand in seiner …	Timo didn't want to be alone, so there was always someone … (around).
9	N	Üben	Ohne …wird keiner ein Meister.	Without … no one becomes a master. (practice)
10	N	Beschmutzen	Im Park ist das … der Bänke verboten.	It is forbidden to … the benches in the park. (dirty)
11	N	Pony	Auf der Weide steht ein kleines …	On the pasture stands a small … (pony)
12	N	Chili	Nimm für das Essen bitte wenig …	Please take a little … for the food (chili)
13	N	Mumie	Die Besucher im Museum betrachten eine schaurige …	The visitors in the museum look at a gruesome … (mummy)
14	N	Orange	In den Obstsalat gehört auch eine …	In the fruit salad belongs also an … (orange)
15	N	Bus	Beeil dich, wir kommen zu spät zum …	Hurry up, we'll be late for the … (bus)
16	N	Lot	Die Mauer muss nach dem … ausgerichtet werden.	The wall must be aligned according to the … (plumb line).
17	V	abmessen	Mit dem Lineal kann man Längen …	With the ruler you can … length (measure)
18	V	stoppen	Beim 100 m Lauf muss man die Zeit …	At the 100 m run you have to … the time (measure)
19	V	schnappt	Der Fisch … nach dem Haken.	The fish … at the hook. (snaps)
20	V	heißen	Die Brüder … Max und Moritz.	The brothers … Max and Moritz. (are called)
21	V	erfahren	Die Nachricht ist unglaublich, Sarah muss unbedingt davon …	The news is unbelievable; Sarah has to … about it. (hear)
22	V	treffen	Können wir uns in der Pause auf dem Hof … ?	Can we … in the yard during the break? (meet)
23	V	verschmutzen	Beim Essen … kleine Kinder oft die Kleider.	While eating small children often … their clothes (dirty).
24	V	mahlt	Die Mühlsteine drehen sich, wenn der Müller das Korn …	The millstones turn when the miller … the grain (grinds)
25	V	bitten	Wenn die Gäste klingeln, … wir sie herein.	When the guests ring the bell, we'll … them (invite).
26	A	verwöhnt	Meine Schwester ist ein Nesthäkchen, sie wird meistens …	My sister is a nestling, she is mostly … (spoiled)
27	A	dick	Ein Buch mit 800 Seiten ist …	A book with 800 pages is … (thick)
28	A	hässlich	Manche finden das Kunstwerk schön, andere finden es …	Some find the work of art beautiful, others find it …(ugly)
29	A	schmierig	Wenn man zu viel Gel nimmt, werden die Haare oft …	If you take too much gel, your hair often becomes … (greasy)
30	A	knurrend	Ich traute mich nicht ins Haus, weil ein Hund … vor der Türe saß.	I didn't dare go into the house because a … dog was at the door. (growling)
31	A	plötzlich	Gerade hat die Sonne geschienen, jetzt fängt es … an zu regnen.	The sun has just shone, now it … starts to rain (suddenly)
32	A	schneller	Mein neues Auto fährt … als das alte.	My new car drives … than the old one. (faster)
33	A	hell	Die Sonne scheint …	The sun is shining … (brightly)
34	A	okay	Den neuen Kinofilm findet Max ganz …	Max finds the new movie quite … (okay)

Table A3: Pseudowords: item set A and B.

	N	Cat.	Test word	Sentence	English translation
Set A	P1	N	Ponu	Auf dem Tisch stehen noch drei Tassen ...	There are three cups... left on the table. (Ponu)
	P2	V	ralt	Nach der Schule ... Paula gern.	After school Paula likes ... (ralt)
	1	N	Mieb	Auf der Wiese steht ein großes ...	On the meadow there is a large ... (Mieb)
	2	N	Folb	Wenn du einkaufen gehst, denke bitte an den ...	When you go shopping, please think of the ... (Folb)
	3	N	Nabo	Im Laden sah sie einen wunderschönen	In the shop she saw a beautiful ... (Nabo)
	4	N	Sprief	Karla machte sich auf die Suche nach dem ...	Karla went on a search for the ... (Sprief)
	5	N	Frabu	Gestern Abend sah ich auf dem Dach einen ...	Last night I saw a ... on the roof. (Frabu)
	6	N	Tokale	In Julias Vitrine stehen viele glänzende ...	In Julia's vitrine are many shiny ... (Tokale)
	7	N	Mippokur	Das ... ist ein riesiges Ungeheuer	The ... is a huge monster. (Mippokur)
	8	N	Abworken	Das ... bei der Klassenarbeit, ist nicht erlaubt.	The ... for the class test is not allowed. (Abworken)
	9	N	Nulfen	Ohne viel ... kannst du nicht besser werden.	You can't get any better without a lot of ... (Nulfen).
	10	V	flappern	Stör mich nicht, ich muss etwas mit ihm ...	Don't bother me, I need to ... to him a little. (flappern)
	11	V	nehlt	Die Jungen warten gespannt, dass der Ball ...	The boys are eagerly waiting for the ball to ... (nehlt)
	12	V	lampfen	Warte auf mich, ich kann nicht so schnell ... !	Wait for me, I can't ... that fast. (lampfen)
	13	V	krulst	Ich kann mich nicht konzentrieren, wenn du so ...	I can't concentrate when you're so ... (krulst)
	14	V	klompern	Es lohnt sich, bei dem neuen Spiel zu ...	It's worth ... on the new game. (klompern)
	15	V	geflochen	Um Wasser zu sammeln, ... wir ein Regenfass.	To collect water, we ... a rain barrel. (geflochen)
	16	V	stiezeln	Es nervt ihn, wenn die Uhren zu laut . ____.	It annoys him when the clocks ... too loud (stiezeln)
	17	A	boff	Der große Korb war schnell .. .	The big basket was fast ... (boff))
	18	A	batull	Ich will das nicht lesen, es ist viel zu ...!	I don't want to read this, it's too ... (batull)
	19	A	schwelz	Heute fühle ich mich zu ... zum Lernen.	Today I feel too ... to learn (schwelz)
	20	A	strimsig	Die Brücke ist morsch und ...	The bridge is rotten and... (strimsig)
Set B	P1	N	Laba	Zum Kochen braucht Mama zwei Löffel ...	Mama needs two spoons to cook ... (Laba)
	P2	V	fahlt	Markus sitzt seit Stunden am Tisch und ...	Markus has been sitting at the table and ... for hours. (fahlt)
	1	N	Birp	Lass bloß die Finger von dem ...!	Don`t touch the ... (Birp).
	2	N	Wirda	Wenn man oben auf dem Berg steht, kann man den ... sehen.	When you're standing on top of the mountain, you can see the ... (Wirda)
	3	N	Schworg	Ich kann nicht glauben, dass du auf den ... herein gefallen bist.	I can’t believe you were tricked by the ... (Schworg)
	4	N	Prone	Nachher holt sie beim Nachbarn ihre ... ab.	Afterwards she picks up her ... at the neighbour's. (Prone)
	5	N	Medale	In Leons Buch gibt es nur wenige ...	There are only a few ... in Leon's book. (Medale)
	6	N	Bammogen	Das ... hat sechs kurze Beine.	The ... has six short legs. (Mippokur)
	7	N	Serg	Im Wasser schwimmt ein schneller ...	In the water swims a fast ... (Serg)
	8	N	Abmerfen	Das ... der Tür, ist nicht gestattet.	The ... of the door is not permitted.(Abmerfen)
	9	N	Lonken	Mit fleißigem ... kannst du die Klassenarbeit schaffen.	With diligent ... you can finish the class test. (Nulfen)
	10	V	spoppen	Warte bitte, ich muss noch etwas mit dir... .	Please wait, I have something to ... to you about. (spoppen)
	11	V	luhmt	Keiner hätte geglaubt, dass er so lange	No one would've believed he'd be so long ... (luhmt)
	12	V	lompern	Man hörte die Reisegruppe laut johlen und	You could hear the travel group screaming loudly and ... (lompern)
	13	V	nalcht	Peter hat einen Freund, der sehr laut	Peter has a friend who is very loud ... (nalcht)
	14	V	krolpen	Es macht Spaß, mit Geschrei in die Pfützen zu	It's fun to ... into the puddles ... (krolpen)
	15	V	strompig	Das Auto ist alt und ...	The car is old and ... (strompig)
	16	V	geplauchen	Wovon du redest, ... wir wirklich nicht.	What you're talking about, we really don't... (geplauchen)
	17	A	humm	Die dunkle Höhle zu durchwandern, ist ...	Walking through the dark cave is ... (hum)
	18	A	namoll	Der Bischof wirkte in seinem Gewand sehr ...	The bishop was very ... in his robe. (namoll)
	19	A	schnolz	Herr Maiers Tag im Büro war heute sehr ...	Mr. Maier’s office day was very ... (schnolz)
	20	A	stiezen	Sie ist gut gelaunt, wenn die Kinder ...	She is in a good mood if the children ... (stiezen)

Note: Test words were preceded by two practice trials (P). The Category (CAT.) column specifies if the test word was used as a noun, verb, or adjective in the sentence

Appendix B

Appendix B provides test material for the Copying task.

Table B1: Parallelized texts for the *Copying Task* in German

Text A:	**Text B:**
Bello und die Eisenbahn	Wuffi und der Rasenmäher
Die Mutter bittet ihren Sohn Felix, jetzt mit Bello raus zu gehen.	Die Mutter bittet ihren Sohn Lukas, jetzt den Rasen zu mähen.
Aber Felix hat überhaupt keine Lust. Er will lieber mit seiner neuen Eisenbahn spielen. Und was macht Bello? Der arme Hund sitzt mit seiner Leine ungeduldig vor der Tür und wartet.	Aber Lukas hat überhaupt keine Lust. Er will lieber mit seinem neuen Fußball spielen. Da kommt sein treuer Hund Wuffi in den Garten. Er setzt sich und wedelt freudig mit seinem Schwanz.

Appendix C

Appendix C summarizes model selection procedure of study 1.

Table C1: Model selection procedure by means of likelihood ratio tests.

		DF	AIC	BIC	*Loglik*	*Deviance*	*Chi²(DF)*	*p*
A)	**Spelling accurcy**							
	null model	3	5491.4	5511.3	-2742.7	5485.4	NA	<.001
	Group	4	5459.56	5486.09	-2725.78	5451.56	33.84 (1)	<.001
	Group + Lex	5	5451.42	5484.59	-2720.71	5441.42	10.13 (1)	.03
	Group + Lex + Mode	**6**	**5448.64**	**5488.44**	**-2718.32**	**5436.64**	**4.79 (1)**	**.14**
	Group * Lex + Mode	7	5448.46	5494.89	-2717.23	5434.46	2.18 (1)	.31
	Group * Lex * Mode	10	5450.86	5517.19	-2715.43	5430.86	3.60 (3)	<.001
	Group + Lex + Mode	6	5448.6	5488.44	-2718	5437	NA	NA
	Sex + Group + Lex + Mode	7	5443.8	5490.23	-2715	5430	6.84 (1)	.01
	Sex + Age + Group + Lex + Mode	**8**	**5439.9**	**5492.99**	**-2712**	**5424**	**5.87 (1)**	**.02**
	Sex * Age + Group + Lex + Mode	9	5438.7	5498.37	-2710	5421	3.25 (1)	.07
	Sex * Age * Group + Lex + Mode	12	5442.4	5522.01	-2709	5418	2.26 (3)	.52
	Sex * Age * Group * Lex + Mode	19	5446.9	5572.98	-2704	5409	9.46 (7)	.22
	Sex * Age * Group * Lex * Mode	34	5456.7	5682.23	-2694	5389	20.20 (15)	.16
B)	**Self-Corrections**							
	null model	3	4032.04	4051.94	-2013.02	4026.04	NA	NA
	Mode	4	3876.70	3903.23	-1934.35	3868.70	157.35 (1)	<.001
	Mode + Group	5	3870.36	3903.53	-1930.18	3860.36	8.34 (1)	<.001
	Mode * Group	6	3850.82	3890.62	-1919.41	3838.82	21.54 (1)	<.001
	Mode * Group + Lex	**7**	**3842.84**	**3889.27**	**-1914.42**	**3828.84**	**9.98** (1)	**<.001**
	Mode * Lex * Group	10	3845.47	3911.80	-1912.73	3825.47	3.38(3)	.34
	Mode * Group + Lex	7	3842.84	3889.27	-1914.42	3828.84	NA	NA
	Age * Mode * Group + Lex	**11**	**3843.07**	**3916.04**	**-1910.53**	**3821.07**	**6.71 (1)**	**.01**
	Sex + Age * Mode * Group + Lex	10	3847.78	3914.11	-1913.89	3827.78	0.38 (1)	.54
	Sex + Age * Mode * Group * Lex	12	3844.37	3923.97	-1910.18	3820.37	0.70 (1)	.40
	Sex * Age * Mode * Group * Lex	34	3855.87	4081.40	-1893.93	3787.87	32.50 (22)	.07
C)	**Writing times**							
	CopyTime + Read	7	83828.30	83874.69	-41907.15	83814.30	NA	NA
	CopyTime + Read + Mode	8	83830.29	83883.31	-41907.14	83814.29	0.01 (1)	.92
	CopyTime + Read + Mode + Lex	9	83819.88	83879.52	-41900.94	83801.88	12.41 (1)	<.001
	CopyTime + Read + Mode * Lex	10	83821.39	83887.66	-41900.69	83801.39	0.49 (1)	.48
	CopyTime + Read + Mode * Lex + Group	11	83822.08	83894.98	-41900.04	83800.08	1.30 (1)	.25
	CopyTime + Read + Mode * Lex * Group	**13**	**83814.24**	**83900.39**	**-41894.12**	**83788.24**	**11.84 (2)**	**<.001**
	CopyTime + Read + Mode * Lex * Group	13	83814.24	83900.39	-41894.12	83788.24	NA	NA
	[...] Age * Mode * Lex * Group	**20**	**83675.66**	**83808.21**	**-41817.83**	**83635.66**	**152.58 (7)**	**<.001**
	[...] + Sex + Age * Mode * Lex * Group	22	83675.80	83821.60	-41815.90	83631.80	3.86 (2)	.14
	[...] + Sex * Age * Mode * Lex * Group	37	83621.60	83866.81	-41773.80	83547.60	84.20 (15)	<.001

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Appendix D

Appendix D summarizes model selection procedure for each spelling rule of study 2.

Table D1: Capitalization: model selection procedure.

		DF	AIC	BIC	*Loglik*	*Deviance*	*$Chi^2(DF)$*	*P*
A)	**Spelling accurcy**							
	null model	3	3056.01	3075.91	-1525.01	3050.01	NA	NA
	Group	4	3046.83	3073.36	-1519.41	3038.83	11.19 (1)	<.001
	Group + Lex	5	3041.7	3074.86	-1515.85	3031.7	7.13 (1)	.01
	Group + Lex + WoCl	6	2947.2	2987	-1467.6	2935.2	96.50 (1)	<.001
	Group + Lex + WoCl + Mode	7	2946.25	2992.68	-1466.12	2932.25	2.95 (1)	.09
	Group + Lex + WoCl * Mode	8	2933.91	2986.98	-1458.96	2917.91	14.34 (1)	<.001
	Group + Lex * WoCl * Mode	***11***	***2927.4***	***3000.37***	***-1452.7***	***2905.4***	***12.51 (3)***	***.01***
	Group * Lex * WoCl * Mode	18	2933.23	3052.63	-1448.62	2897.23	8.17 (7)	.32
	Age + Group + Lex * WoCl * Mode	11	2927.4	3000.37	-1452.7	2905.4	NA	NA
	Sex + Age + Group + Lex * WoCl * Mode	12	2928.25	3007.85	-1452.12	2904.25	1.15 (2)	.28
	*Sex + Age + Group * Lex * WoCl * Mode*	*13*	*2926.19*	*3012.43*	*-1450.1*	*2900.19*	*4.05* (1)	*.04*
	Sex + Age * Group * Lex * WoCl * Mode	20	2932.35	3065.01	-1446.17	2892.35	7.85 (1)	.35
	Sex * Age * Group * Lex * WoCl * Mode	**66**	**2941.11**	**3378.92**	**-1404.56**	**2809.11**	**45.57** (3)	**.04**
B)	**Self-Corrections**							
	null model	3	1121.44	1141.34	-557.72	1115.44	NA	NA
	WoCl	4	1074.18	1100.71	-533.09	1066.18	49.26 (1)	<.001
	WoCl + Mode	5	1052.17	1085.35	-521.09	1042.17	24.00 (1)	<.001
	WoCl + Mode + Lex	**6**	**1048.67**	**1088.47**	**-518.33**	**1036.67**	**5.51 (1)**	**.02**
	WoCl + Mode + Lex + Group	7	1047.21	1093.66	-516.61	1033.21	3.45 (1)	.06
	WoCl * Mode + Lex + Group	8	1046.85	1099.93	-515.42	1030.85	2.36 (1)	.12
	WoCl * Mode * Lex + Group	11	1050.48	1123.46	-514.24	1028.48	2.37 (3)	.50
	WoCl * Mode * Lex * Group	18	1052.72	1172.14	-508.36	1016.72	11.77 (7)	.11
	WoCl + Mode + Lex	6	1048.67	1088.47	-518.33	1036.67	NA	NA
	Age + WoCl + Mode + Lex	7	1048.13	1094.58	-517.07	1034.13	2.53 (1)	.11
	Age + Sex + WoCl + Mode + Lex	8	1050.04	1103.11	-517.02	1034.04	0.10 (1)	.75
	Age * Sex + WoCl + Mode + Lex	9	1051.68	1111.40	-516.84	1033.68	0.35 (1)	.55
	Age * Sex *WoCl + Mode + Lex	12	1054.96	1134.58	-515.48	1030.96	2.72 (3)	.44
	Age * Sex *WoCl * Mode + Lex	19	1060.68	1186.74	-511.34	1022.68	8.28 (7)	.31
	Age * Sex *WoCl * Mode * Lex	34	1078.52	1304.10	-505.26	1010.52	12.16 (15)	.67
C)	**Writing Times**							
	CopyTime + Read	6	88062.22	88102.00	-44025.11	88050.22	NA	NA
	CopyTime + Read + Mode	7	87558.98	87605.38	-43772.49	87544.98	505.25 (1)	<.001
	CopyTime + Read + Mode + WoCl	8	87560.25	87613.28	-43772.12	87544.25	0.73 (1)	.39
	CopyTime + Read + Mode * WoCl	9	87547.53	87607.18	-43764.76	87529.53	14.72 (1)	<.001
	CopyTime + Read + Mode * WoCl + Lex	10	87548.83	87615.12	-43764.41	87528.83	0.70 (1)	.40
	CopyTime + Read + Mode * WoCl + Mode * Lex	**11**	**87542.11**	**87615.03**	**-43760.06**	**87520.11**	**8.72 (2)**	**<.001**
	CopyTime + Read + Mode * WoCl * Lex	13	87541.38	87627.55	-43757.69	87515.38	4.73 (8)	.09
	CopyTime + Read + Mode * WoCl + Mode * Lex	11	87542.11	87615.03	-43760.06	87520.11	NA	NA
	[…] + Age * Mode * WoCl + Mode * Lex	**15**	**87450.08**	**87549.51**	**-43710.04**	**87420.08**	**100.03 (4)**	**<.001**
	[…] + Sex + Age + Mode * WoCl + Mode * Lex	23	87461.11	87613.57	-43707.56	87415.11	4.97 (8)	.76
	[…] + Sex * Age + Mode * WoCl + Mode * Lex	37	87472.34	87717.60	-43699.17	87398.34	16.77 (14)	.27

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Table D2: Consonant doubling: Model selection procedure.

A)	Spelling accuracy	DF	AIC	BIC	*Loglik*	*Deviance*	*Chi²(DF)*	*P*
	null model	3	1430.83	1447.67	-712.41	1424.83	NA	NA
	Group	4	1405.73	1428.19	-698.86	1397.73	27.10 (1)	<.001
	Group + Lex	**5**	**1387.00**	**1415.08**	**-688.50**	**1377.00**	**20.73 (1)**	**<.001**
	Group * Lex	6	1389.00	1422.69	-688.50	1377.00	0.00 (1)	.98
	Group + Lex	5	1387.00	1415.08	-688.50	1377.00	NA	NA
	Sex + Group + Lex	6	1381.83	1415.52	-684.92	1369.83	7.17 (1)	.01
	Sex + Age + Group + Lex	**7**	**1375.84**	**1415.14**	**-680.92**	**1361.84**	**8.00 (1)**	**<.001**
	Sex + Age * Group * Lex *	11	1382.47	1444.24	-680.24	1360.47	1.37 (4)	.85
	Sex * Age * Group * Lex *	18	1383.18	1484.26	-673.59	1347.18	13.29 (7)	.07
B)	**Self-Corrections**							
	null model	3	333.68	350.53	-163.84	327.68	NA	NA
	Mode	**4**	**330.78**	**353.25**	**-161.39**	**322.78**	**4.90 (1)**	**.03**
	Mode + Group	5	330.82	358.90	-160.41	320.82	1.97 (1)	.16
	Mode * Group	6	331.68	365.38	-159.84	319.68	1.14 (1)	.29
	Mode	4	330.78	353.25	-161.39	322.78	NA	NA
	Age + Mode	5	332.68	360.77	-161.34	322.68	0.10 (1)	.75
	Age + Sex + Mode	6	332.71	366.40	-160.35	320.71	1.98 (1)	.16
	Age * Sex + Mode	7	331.25	370.56	-158.62	317.25	3.46 (1)	.06
	Age * Sex * Mode	10	332.95	389.11	-156.48	312.95	4.30 (1)	.23
C)	**Writing times**							
	CopyTime + Read	6	29711.92	29745.57	-14849.96	29699.92	NA	NA
	CopyTime + Read + Mode	7	29712.70	29751.96	-14849.35	29698.70	1.22 (1)	.27
	CopyTime + Read + Mode + Lex	**8**	**29444.68**	**29489.54**	**-14714.34**	**29428.68**	**270.03 (1)**	**<.001**
	CopyTime + Read + Mode * Lex	9	29446.50	29496.96	-14714.25	29428.50	0.18 (1)	.67
	CopyTime + Read + Mode * Lex + Group	10	29445.80	29501.87	-14712.90	29425.80	2.70 (1)	.10
	CopyTime + Read + Mode * Lex + Group * Lex	11	29447.80	29509.48	-14712.90	29425.80	0.00 (1)	.98
	CopyTime + Read + Mode * Lex * Group	13	29445.93	29518.83	-14709.97	29419.93	5.87 (2)	.06
	CopyTime + Read + Mode + Lex	8	29444.68	29489.54	-14714.34	29428.68	NA	NA
	CopyTime + Read + Age + Mode + Lex	9	29432.99	29483.45	-14707.49	29414.99	13.69 (1)	<.001
	CopyTime + Read + Age * Mode + Lex	**10**	**29385.54**	**29441.61**	**-14682.77**	**29365.54**	**49.44** (1)	**<.001**
	CopyTime + Read + Age * Mode + Lex	12	29386.47	29453.76	-14681.23	29362.47	3.07 (2)	.22
	CopyTime + Read + Age * Mode + Age * Lex	14	29386.89	29465.39	-14679.44	29358.89	3.58 (2)	.17
	CopyTime + Read + Age * Mode * Lex	21	29395.65	29513.40	-14676.82	29353.65	5.24 (7)	.63

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Table D3: Lengthening: Model selection procedure

A)	Spelling accuracy	DF	AIC	BIC	*Loglik*	*Deviance*	*Chi²(DF)*	*p*
	null model	3	846.22	861.06	-420.11	840.22	NA	NA
	Group	4	807.79	827.57	-399.89	799.79	40.44 (1)	<.001
	Group + Lex	5	793.17	817.90	-391.58	783.17	16.62 (1)	<.001
	Group * Lex	**6**	**783.98**	**813.66**	**-385.99**	**771.98**	**11.19** (1)	**<.001**
	Group * Lex + Mode	7	785.47	820.10	-385.73	771.47	0.51 (1)	.48
	Group * Lex * Mode	10	790.12	839.59	-385.06	770.12	1.35 (3)	0.72
	Group * Lex	5	793.17	817.90	-391.58	783.17	NA	NA
	Sex + Group * Lex	7	780.63	815.26	-383.31	766.63	16.54 (2)	<.001
	Sex + Age + Group * Lex	**8**	**773.99**	**813.57**	**-379.00**	**757.99**	**8.63 (3)**	**<.001**
	Sex + Age * Group * Lex	11	776.40	830.82	-377.20	754.40	3.59 (1)	.31
	Sex * Age * Group * Lex	18	781.73	870.78	-372.87	745.73	8.67 (7)	.28
B)	**Self-Corrections**							
	null model	3	177.67	192.44	-85.83	171.67	NA	NA
	Mode	**4**	**163.80**	**183.50**	**-77.90**	**155.80**	**15.87 (1)**	**<.001**
	Mode + Group	5	164.70	189.33	-77.35	154.70	1.10 (1)	.30
	Mode * Group	6	166.04	195.60	-77.02	154.04	0.66 (1)	.42
	Mode	4	163.80	183.50	-77.90	155.80	NA	NA
	Age + Mode	5	165.34	189.97	-77.67	155.34	0.46 (1)	.50
	Age + Sex + Mode	6	167.23	196.79	-77.62	155.23	0.11 (1)	.75
	Age * Sex + Mode	7	168.83	203.31	-77.41	154.83	0.41 (1)	.52
	Age * Sex * Mode	10	172.74	222.00	-76.37	152.74	2.08 (3)	.56
C)	**Writing times**							
	CopyTime + Read	6	16135.26	16164.90	-8061.63	16123.26	NA	NA
	CopyTime + Read + Mode	**7**	**16023.88**	**16058.46**	**-8004.94**	**16009.88**	**113.38** (1)	**<.001**
	CopyTime + Read + Mode + Lex	8	16024.71	16064.23	-8004.36	16008.71	1.17 (1)	.28
	CopyTime + Read + Mode * Lex	9	16026.29	16070.76	-8004.15	16008.29	0.42 (1)	.52
	CopyTime + Read + Mode * Lex + Group	10	16027.37	16076.78	-8003.69	16007.37	0.92 (1)	.34
	CopyTime + Read + Mode * Lex * Group	13	16033.09	16097.31	-8003.55	16007.09	0.28 (2)	.96
	CopyTime + Read + Mode	7	16023.88	16058.46	-8004.94	16009.88	NA	NA
	CopyTime + Read + Age * Mode	**9**	**16002.12**	**16046.58**	**-7992.06**	**15984.12**	**25.76 (2)**	**<.001**
	CopyTime + Read + Sex + Age * Mode	10	16001.96	16051.36	-7990.98	15981.96	2.16 (1)	.14
	CopyTime + Read + Sex * Age * Mode	13	16007.52	16071.75	-7990.76	15981.52	0.43 (3)	.93

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Table D4: Devoicing: Model selection procedure.

		DF	AIC	BIC	*Loglik*	*Deviance*	*Chi²(DF)*	*p*
A)	**Spelling accuracy**							
	null model	3	387.02	401.00	-190.51	381.02	NA	NA
	Group	4	374.95	393.58	-183.47	366.95	14.07 (1)	<.001
	Group + Lex	5	376.89	400.19	-183.45	366.89	0.05 (1)	.82
	Group * Lex	**6**	**375.30**	**403.26**	**-181.65**	**363.30**	**3.59 (1)**	**.05**
	Group * Lex + Mode	7	375.93	408.55	-180.97	361.93	1.37 (1)	.24
	Group * Lex * Mode	10	377.36	423.95	-178.68	357.36	4.57 (1)	.21
	Group * Lex	5	376.89	400.19	-183.45	366.89	NA	NA
	Sex + Group * Lex	7	376.71	409.33	-181.36	362.71	4.18 (2)	.12
	Sex + Age + Group * Lex	8	378.67	415.94	-181.33	362.67	0.05 (1)	.83
	Sex + Age * Group * Lex	11	384.40	435.65	-181.20	362.40	0.27 (3)	.97
	Sex * Age * Group * Lex	18	393.65	477.52	-178.83	357.65	4.75 (7)	.69
B)	**Self-Corrections**							
	null model	3	70.42	84.35	-32.21	64.42	NA	NA
	Mode	4	72.42	90.99	-32.21	64.42	0.00 (1)	.96
	Mode + Group	5	74.43	97.65	-32.22	64.43	0.00 (1)	1.00
	Mode * Group	**6**	**71.44**	**99.30**	**-29.72**	**59.44**	**4.99** (1)	**.03**
	Mode * Group	6	71.44	99.30	-29.72	59.44	4.98 (1)	.17
	Sex + Mode * Group	7	74.04	106.54	-30.02	60.04	0.00 (1)	1.00
	Sex * Mode * Group	**10**	**67.93**	**114.36**	**-23.97**	**47.93**	**11.51** (4)	**.02**
	Age + Sex * Mode * Group	11	69.95	121.02	-23.98	47.95	0.00 (1)	1.00
	Age * Sex * Mode * Group	18	89.28	172.84	-26.64	53.28	0.00 (7)	1.00
C)	**Writing times**							
	CopyTime + Read	6	11569.47	11597.42	-5778.73	11557.47	NA	NA
	CopyTime + Read + Mode	**7**	**11440.48**	**11473.10**	**-5713.24**	**11426.48**	**130.98 (1)**	**<.001**
	CopyTime + Read + Mode + Lex	8	11441.61	11478.89	-5712.81	11425.61	0.87 (1)	.35
	CopyTime + Read + Mode + Lex + Group	9	11443.50	11485.43	-5712.75	11425.50	0.11 (1)	.73
	CopyTime + Read + Mode * Lex + Group	10	11444.43	11491.03	-5712.22	11424.43	1.06 (1)	.30
	CopyTime + Read + Mode * Lex * Group	13	11447.79	11508.36	-5710.90	11421.79	2.64 (3)	.45
	CopyTime + Read + Mode	7	11440.48	11473.10	-5713.24	11426.48	NA	NA
	CopyTime + Read + Age * Mode	**9**	**11410.29**	**11452.22**	**-5696.14**	**11392.29**	**34.20 (2)**	**<.001**
	CopyTime + Read + Sex + Age * Mode	10	11411.07	11457.67	-5695.54	11391.07	1.21 (1)	.27
	CopyTime + Read + Sex + Age * Mode	13	11415.90	11476.47	-5694.95	11389.90	1.18 (3)	.76

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Table D5: Rule words: Model selection procedure**.**

		DF	AIC	BIC	*Loglik*	*Deviance*	*Chi²(DF)*	*p*
A)	**Spelling accuracy**							
	null model	3	865.41	879.58	-429.70	859.41	NA	NA
	Group	**4**	**854.73**	**873.63**	**-423.37**	**846.73**	**12.68 (1)**	**<.001**
	Group + Mode	5	856.70	880.32	-423.35	846.70	0.03 (1)	.85
	Group * Mode	6	858.19	886.53	-423.09	846.19	0.51 (1)	.48
	Group * Mode	4	854.73	873.63	-423.37	846.73	NA	NA
	Sex + Group * Mode	**5**	**854.71**	**878.33**	**-422.36**	**844.71**	**2.02** (1)	**<.001**
	*Sex + Age + Group * Mode*	6	852.16	880.50	-420.08	840.16	4.55 (1)	.85
	*Sex + Age * Group * Mode*	10	853.85	901.09	-416.93	833.85	6.10 (3)	.11
B)	**Self-Corrections**							
	null model	3	329.23	343.36	-161.62	323.23	NA	NA
	Mode	**4**	**322.71**	**341.55**	**-157.36**	**314.71**	**8.52** (1)	**<.001**
	Mode + Group	5	324.57	348.11	-157.28	314.57	0.15 (1)	.70
	Mode * Group	6	326.40	354.65	-157.20	314.40	0.17 (1)	.68
	Mode	4	322.71	341.55	-157.36	314.71	NA	NA
	Age + Mode	5	324.14	347.68	-157.07	314.14	0.58 (1)	.45
	Age * Mode	6	325.17	353.42	-156.58	313.17	0.97 (1)	.33
	Sex + Age * Mode	**7**	**323.27**	**356.23**	**-154.64**	**309.27**	**3.90** (1)	**.05**
	Sex * Age * Mode	10	326.68	373.76	-153.34	306.68	2.59 (3)	.46
C)	**Writing times**							
	CopyTime + Read	6	13232.35	13260.66	-6610.18	13220.35	NA	NA
	CopyTime + Read + Mode	**7**	**13163.77**	**13196.80**	**-6574.89**	**13149.77**	**70.58 (1)**	**<.001**
	CopyTime + Read + Mode + Lex	9	13167.44	13209.90	-6574.72	13149.44	0.34 (2)	.84
	CopyTime + Read + Age + Mode	7	13163.77	13196.80	-6574.89	13149.77	NA	NA
	CopyTime + Read + Age * Mode	**9**	**13155.77**	**13198.23**	**-6568.88**	**13137.77**	**12.01 (2)**	**<.001**
	CopyTime + Read + Sex + Age * Mode	10	13155.10	13202.28	-6567.55	13135.10	2.66 (1)	.10
	CopyTime + Read + Sex * Age * Mode	13	13155.38	13216.72	-6564.69	13129.38	5.72 (3)	.13

Note: The hierarchical nested models were tested against the null model (with random effect for subjects and items only) using likelihood ratio tests. The models included age and lexicality (Lex) as covariates.

Study 4:

Closing the gap – Bounded and unbounded number line estimation in secondary school children

Stefanie Jung, Stephanie Roesch, Elise Klein, Tanja Dackermann,
Juergen Heller, Korbinian Moeller

Abstract

Changes in number line estimation (NLE) performance are frequently used as an indicator for the development of number magnitude representation. For this purpose, two different task versions have been applied: a traditional bounded and a relatively new unbounded NLE task. Previous studies mainly assessed primary school children or adults showing that these tasks differ in terms of i) difficulty and solution strategies employed as well as ii) with respect to their relation to other basic numerical/arithmetic skills. So far, data from secondary school children are scarce for bounded NLE, and even no data is available for unbounded NLE in this age group. Thus, we assessed bounded and unbounded NLE in grade levels 5 to 7 to evaluate a) developmental as well as strategic influences, and b) the relation of bounded and unbounded NLE performance with basic arithmetic skills. Results showed that children performed better in bounded as compared to unbounded NLE. Children seemed to employ the use of different solution strategies for bounded (i.e., proportion-judgment) and unbounded (i.e., magnitude-estimation based) NLE. Moreover, only for bounded NLE, estimation accuracy increased with age. Furthermore, estimation performance for bounded but not unbounded NLE was strongly associated with basic arithmetic (addition, subtraction, multiplication, and division). Our findings indicate that differential results for bounded and unbounded NLE obtained in primary school children seem to generalize to older secondary school children. Presented results substantially contribute to the knowledge about the (consecutive) development of skills pertaining to bounded and unbounded number line estimation.

1. Introduction

Number line estimation (NLE) refers to the ability to locate numbers on a physical line according to their magnitude. NLE performance is a widely used indicator of basic numerical skills in studies evaluating the development of numerical skills through childhood. Traditionally, NLE refers to estimation performance on a *bounded* number line with a numerical interval determined by a given start and end point (e.g., 0-100, 0-1,000, etc.) upon which participants are required to estimate the spatial position of a given number (e.g., Siegler and Opfer, 2003). Less frequently, *unbounded* NLE was assessed (e.g., Cohen & Blanc-Goldhammer, 2011; Link, Huber, Nuerk, & Moeller, 2014). In the unbounded NLE task, the number line has no labelled endpoint, but a single unit distance (e.g., 0-1) in addition to the start point. Starting from this unit distance, target numbers have to be located upon the given line.

Bounded and unbounded NLE tasks have been used as indicators to draw conclusions from task performance onto the underlying magnitude representation and its development (e.g., for bounded NLE: Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Booth & Siegler, 2006, 2008; Ebersbach, Luwel, & Verschaffel, 2015; Siegler & Booth, 2004; Siegler & Opfer, 2003; Thompson & Opfer, 2010; Young & Booth, 2015; e.g., for unbounded NLE: Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014). However, the two tasks were found to differ in terms of i) the solution strategies employed (e.g., considering the usage of reference points [cf. Cohen & Blanc-Goldhammer, 2011; Cohen & Sarnecka, 2014; Link, Huber, et al., 2014]) and ii) their association with other basic numerical as well as arithmetic skills (for a meta-analysis see Schneider et al., 2018). More importantly, despite numerous studies relating primarily bounded NLE to other basic numerical and arithmetic skills, the vast majority of these studies focused on young children from kindergarten to the end of primary school (e.g., Link, Nuerk, et al., 2014; Moeller et al., 2009; Schneider et al., 2017; Zhu et al., 2017) or adults (Reinert, Huber, Nuerk, & Moeller, 2017, 2015). So far, the age range in-between (i.e., secondary school children) has largely been neglected for bounded NLE. For unbounded NLE, there is still no empirical data at all from secondary school children. Crucially, consideration of relatively narrow age ranges strongly limits general statements on the development of number representation.

In the current study, we aimed at addressing this gap by investigating bounded and unbounded NLE in secondary school children[10] (i.e., 5^{th} to 7^{th} graders) to evaluate both a)

[10] In the German school system, majority of secondary schools begins with the 5th grade.

developmental changes in *estimation accuracy and strategy use* and b) the relation of bounded and unbounded *number line estimation with basic arithmetic*.

1.1. Estimation accuracy and strategy use

Regarding bounded NLE, there are different views on the developmental changes of NLE performance and its relationship to the underlying representation of number magnitude. Generally, all views are based on the observation that children of different age groups perform differently in the same number range: Younger children (i.e., preschoolers and first-graders) were observed to systematically overestimate the spatial position of small numbers (i.e., placing 9 at about the location where 40 would go on a 0-100 scale; Moeller et al., 2009). In turn, they localized larger numbers in a compressed way towards the upper bound of the scale (e.g., Siegler & Booth, 2004; Siegler & Opfer, 2003). Older children (e.g., second-graders), in contrast, did not show such a pattern and were observed to perform more accurately in the same number range (e.g., Siegler & Booth, 2004).

These estimation patterns have been discussed in various theoretical accounts (see Dackermann et al., 2015, for a recent review). Most importantly, two competing accounts can be differentiated: the log-to-linear shift account (e.g., Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Opfer, 2003) and the proportion-judgment account (e.g., Barth & Paladino, 2011; Slusser, Santiago, & Barth, 2013).

The log-to-linear shift account proposes a representational change from an initially logarithmic layout, as indicated by overestimation of smaller numbers, to a more linear layout of estimation patterns with increasing age and experience. Accordingly, a logarithmic function was found to fit the observed estimation pattern best in younger children whereas a linear function reflected the observed estimation patterns best in older children and adults (Booth & Siegler, 2006, 2008; Siegler & Booth, 2004; Siegler & Opfer, 2003). The account assumes estimation patterns to directly reflect the underlying magnitude representation (e.g., Kim & Opfer, 2017; Siegler & Opfer, 2003). However, the interpretation of a direct assessment of the magnitude representation is questioned in alternative accounts that emphasize the importance of other numerical processes in addition to the pure magnitude estimation. Ebersbach, Luwel, Frick, Onghena, and Verschaffel (2008), for instance, mentioned the increasing familiarity of children with numbers (but see Thompson & Opfer, 2010 for results contradicting the assumption that children's familiarity with numbers predicts individual’s mental representation of the number line), whereas Moeller et al., 2009 suggested considering the understanding of the place-value system, which also affects NLE performance

The most prominent alternative account, the proportion-judgment account, postulates that reference points (i.e., start and end point of the respective scale, but also quartiles and midpoint, also referred to as [subjective] landmarks) are considered for estimation (e.g., Barth & Paladino, 2011; Siegler & Opfer, 2003; Slusser et al., 2013). The use of reference points is supposed to increase with age and to enhance estimation accuracy (e.g., Slusser et al., 2013, see also Schneider et al., 2008, for eye-fixation data,). Evidence for the account comes from children's estimation patterns: while target estimates seem to follow a linear layout, estimation errors for target numbers were more accurate and varied less at and/or around reference points, which results in a typical *M*-shaped pattern in contour analyses (Ashcraft & Moore, 2012; Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014). Additionally, use of reference points and thus proportion-judgment strategies were found to be indexed by (one- or two-) cyclic power models (Barth & Paladino, 2011; Hollands & Dyre, 2000). Contrary, fit of cyclic power models was also associated with feedback on the mid-point of the number line (i.e., 50 by a number line length of 100) provided to participants (Opfer, Siegler, & Young, 2011; Opfer, Thompson, & Kim, 2016).

In sum, all accounts share the idea that several different development-related cognitive skills and solution strategies are involved in NLE (cf. Barth & Paladino, 2011; Cohen & Sarnecka, 2014; Ebersbach et al., 2008; Moeller et al., 2009; Siegler & Booth, 2004; Siegler & Opfer, 2003; Slusser et al., 2013). This is reflected in an ongoing debate on whether the traditional bounded NLE task is generally a valid measure of children's number magnitude representation (D. J. Cohen & Sarnecka, 2014; Slusser & Barth, 2017) as number magnitude estimation processes seem to be masked by other underlying cognitive processes and solution strategies.

In contrast, unbounded NLE is argued to reflect a purer measure of number magnitude representation (e.g., Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014; see also Reinert, Huber, Nuerk, & Moeller, 2015 for eye-tracking evidence supporting this argument) and seems to be less influenced by age as no development-related performance changes have been found (D. J. Cohen & Sarnecka, 2014; Link, Huber, et al., 2014). This assumption is based on empirical observations of children of different age levels within primary school and adults. Error variability in unbounded NLE was found to increase linearly with number magnitude (D. J. Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014). This finding was interpreted to reflect the strategic procedure of estimating a multiple of the given unit distance (i.e., dead-reckoning strategy as suggested by Cohen & Blanc-Goldhammer, 2011 or a counting-like strategy as recommended by Schneider et al., 2018), which leads to an accumulation of estimation errors (Reinert et al., 2015). As such, it was argued that unbounded NLE reflects a more direct type of number magnitude estimation (Link, Huber, et al., 2014; but see Kim & Opfer, 2017). Cohen & Blanc-Goldhammer (2011) showed first evidence that dead-reckoning strategy, for instance, is indicated by a repetitive scalloped estimation pattern.

1.2. Number line estimation and basic arithmetic

Numerous studies have examined associations of bounded and/or unbounded NLE and basic arithmetic. These studies have shown significant correlations, but only for the bounded version of the task (e.g., Link, Huber, et al., 2014; Schneider et al., 2018). Bounded NLE performance significantly predicts mathematical achievement concurrently but also longitudinally (e.g., Booth & Siegler, 2008; Schneider, Grabner, & Paetsch, 2009; but see Simms et al., 2016, for methodological aspects) and correlates with counting, basic arithmetic and algebra as well as with school achievement and grades (Schneider et al., 2018, for a meta-analysis) from kindergarten to primary school (Booth & Siegler, 2006). In this vein, Friso-van den Bos et al. (2015) argued that these associations may occur due to a mutual relation between both basic arithmetic and NLE skills in numerical development: With increasing practice in solving mathematical problems in childhood, the representation of number magnitude becomes more accurate, which is in turn reflected in better NLE performance.

Comparing data from both versions of the NLE task, Cohen and Sarnecka (D. J. Cohen & Sarnecka, 2014) argued that they are of different levels of difficulty and therefore likely to be associated with different basic arithmetic skills. The authors observed that young children (i.e., preschoolers) underperformed in bounded NLE, but were able to cope with the unbounded version approximately as well as older children (but see Kim & Opfer, 2017). Thus, Cohen and Sarnecka (2014) concluded that unbounded NLE may be less demanding, probably because task solution requires less advanced basic arithmetic skills, as compared to bounded NLE. Bounded NLE, instead, was assumed to be specifically associated with more demanding basic arithmetic skills such as subtraction and division.

Contrasting bounded and unbounded NLE performance in a sample of fourth-graders, the study of Link, Nuerk and Moeller (2014) partially confirmed these theoretical assumptions. The authors argued that bounded NLE may require a division of the number line to guide the application of proportion-based strategies by reference points. Subsequently, starting from a reference point, a leftward (reflecting subtraction) or rightward shift on the number line (reflecting addition) might be necessary to locate the respective target number correctly upon the line. This adjustment, however, is not appropriate in unbounded NLE. In this vein, Link and colleagues (2014) did not observe significant correlations between basic arithmetic and unbounded NLE in their sample of fourth-graders (but see Reinert et al., 2015 for different results regarding multiplication). Despite an increasing number of studies directly comparing bounded and unbounded NLE, it is not yet clear how both tasks differ in their associations with various basic arithmetic operations.

However, it is important that the results and assumptions described above are again mainly based on observations of children from kindergarten to the end of primary school and of adults, whereas the age group in-between, i.e., secondary school children, has so far largely been neglected. Crucially, the few studies investigating the relationship between NLE and basic arithmetic in typically developing secondary school children mainly assessed NLE of fractions and related estimation performance to either a standardized mathematical school achievement test (Siegler & Pyke, 2013) or to school grades (Schneider et al., 2009). Only few studies focused on whole numbers. Some of these studies included rather small sample sizes (32 sixth graders in Siegler & Opfer, 2003 and 24 sixth graders Thompson & Opfer, 2010) or did not use standard NLE tasks by evaluating NLE on both negative (-10,000-0) and positive numbers (-1,000-1,000) in sixth and seventh graders (Young & Booth, 2015).

On a neuro-functional level, Berteletti et al. (2015) provided first evidence that bounded NLE performance correlates with activation of brain areas associated with number magnitude processing (such as the intraparietal sulcus during single digit subtraction). Nevertheless, empirical data from secondary school children on unbounded NLE are still missing entirely. Moreover, so far, there is no comparison of bounded and unbounded NLE in secondary school children, which also compares the association of bounded and unbounded NLE performance and basic arithmetic in three consecutive age levels.

1.3. The current study

The current study focused on bounded and unbounded NLE in secondary school children from grade 5 to 7. In particular, we investigated whether bounded NLE performance changes over time within this age group and aimed at providing first empirical data on how secondary school children perform in unbounded NLE. For bounded NLE, previous studies assessing number line estimation beyond primary school used, number ranges between 0-1,000 and 0-100,000 (Link, Huber, et al., 2014; Siegler & Opfer, 2003; Thompson & Opfer, 2010; Young & Booth for negative numbers, 2015). For unbounded NLE, mainly smaller number ranges up to 20 (Cohen & Blanc-Goldhammer, 2011; Cohen & Sarnecka, 2014), 29 (Link, Nuerk, et al., 2014; Link, Huber, et al., 2014), and 50 (Reinert et al., 2015a; Reinert et al., 2015b) were used. Only Kim and Opfer (2017) used number ranges of 0-30/100/1,000 in unbounded NLE. In studies directly comparing bounded and unbounded NLE performance either equal number ranges in both tasks (Reinert, Huber, Nuerk, & Moeller, 2015, Cohen & Blanc-Goldhammer, 2011; Cohen & Sarnecka, 2014, Kim & Opfer, 2017) or markedly larger number ranges (from 0-100, 0-1,000, 0-10,000 for adults, Link, Huber, Nuerk, & Moeller, 2014) were applied in the bounded version. Importantly, and different from bounded number line estimation, there do not seem to be substantial influences of age (e.g., Link et al., 2014) and the respective number range (e.g., Kim & Opfer, 2017) on participants' estimation patterns. In fact, estimation patterns for all age groups and number ranges were characterized by estimation errors increasing in size and in variability with increasing magnitude of the respective target numbers. However, irrespective of the number ranges (equal or not) Kim and Opfer (2017) as well as Link and colleagues (2014) provided evidence that participants performed less accurate in the unbounded NLE task as compared to the bounded NLE task. Therefore, we used the same number ranges as in Link et al., (2014) study to tie on previous results.

Furthermore, to address the ongoing debate on which version of the NLE task – bounded or unbounded – is associated with which basic arithmetic operation, the relation of both bounded and unbounded NLE with the four basic arithmetic operations (i.e., addition, subtraction, multiplication, and division) was examined. Based on the findings from previous studies outlined above, our hypotheses were as follows:

First, we expected to find significant differences between bounded and unbounded NLE (cf. Cohen & Sarnecka, 2014). In particular, mean estimation errors were assumed to be higher in the unbounded task version (cf. Kim & Opfer, 2017; Link, Huber, et al., 2014). Second, regarding developmental changes, we expected a decline of estimation errors with age for bounded NLE, but no significant age-related differences in estimation errors for unbounded NLE (Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014). Third, we expected children to use proportion-judgment strategies in bounded but not in unbounded NLE and thus, only bounded NLE errors to follow a characteristic *M*-shaped distribution pattern (cf. Ashcraft & Moore, 2012). The use of reference points in bounded NLE should be predominantly reflected by cyclic models that best fit the estimation pattern of children. Finally, based on the theoretical considerations of Cohen and

Sarnecka (2014) and the findings of Link et al. (2014), we assumed both tasks to be associated with different basic arithmetic skills. In particular, we supposed NLE performance in the bounded task to be associated with subtraction, addition, and division. However, we did not formulate a well-directed hypothesis for the association of unbounded NLE performance and basic arithmetic. Results of the present study may help to further clarify these relations.

2. Methods

2.1 Participants

A total of 989 children (463 girls) from 12 German secondary schools participated in this study: children were recruited from 17 5th grade classes (n=342, mean age: 10.58, SD=8.24 months), 21 6th grade classes (n=383, mean age: 11.67, SD=8.48 months) and 15 7th grade classes (n=264, mean age: 12.57, SD=8.19 months). Cross-sectional data was collected at two different survey points (SP1 and SP2). Children were divided into two samples and completed a computerized assessment of either a bounded (Sample 1 assessed at SP1 and SP2: n=700 [SP1: n=389 and SP2: n=311] mean age: 11.53, SD=12.47 months) or unbounded (Sample 2 assessed only at SP2: n=289 mean age: 11.56, SD=12.93 months) NLE task. Each child was examined once. Furthermore, children's basic arithmetic as well as their spelling skills were assessed as control variables.

Data from 195 children who did not complete either the spelling or the arithmetic tasks (n=94) or had more than 33% missing data (n= 101) in the NLE tasks after outlier correction were not considered for further analyses. Finally, complete data of 573 children for the bounded NLE task and 221 children for the unbounded NLE task entered the analyses. Written informed consent was obtained from parents prior to the study besides children's verbal assent before actual testing. The study was approved by the local ethics committee (LEK 2014/19) as well as by the school authority.

2.2 Tasks and procedure:

Data collection took place during regular classes and lasted about 45 minutes. Arithmetic skills were assessed separately for each basic arithmetic operation (i.e., addition, subtraction, multiplication, division) together with either the bounded or unbounded NLE task. Spelling was examined with respect to most relevant German spelling rules (i.e., capitalization, consonant doubling, lengthening, rule words) and served as control variable. Children completed the arithmetic and spelling tasks in counterbalanced order across participants. All stimuli were presented on a 15.6-inch Lenovo ThinkPad T530 laptop display with a resolution of 1024 x 768 pixels at normal viewing distance and with target items in black against a white background.

Task relevant laptop functions (e.g., capitalization, deletion of incorrect entries) were introduced immediately before the assessment, followed by at least one practice trial for each task. Instructions were presented both visually and auditory. This procedure ensured that all children had the same prerequisites to cope with the tasks in the computerized assessment.

2.2.1. Basic arithmetic

Children's basic arithmetic skills in *addition, subtraction, multiplication,* and *division* were assessed following this particular order using a production paradigm. Children were asked to type in the correct result of a presented problem. Overall, 20 problems were presented for each operation. Stimuli were selected as follows: in a*ddition* (e.g., 36+47=____) and *subtraction* (e.g., 91-67=____), problems were controlled for problem size. Additionally, half of the problems required a carrying- or a borrowing-operation. In *multiplication* (e.g., 8*7=____), three problems each were picked from the tables of 3 to 9 as well as two problems from the table of 5. Neither ties nor problems in reversed order were presented. For *division* (e.g., 56/7=____), multiplication problems were reversed. Table A1.1 in Appendix A1 provides an overview of all arithmetic problems.

2.2.2. Number line estimation

Children completed either a bounded (range 0-10,000) or an unbounded (range 0-29) version of the NLE task. In the bounded and unbounded NLE task, target items were presented on the left end above the lower bound of the number line. The physical length of the number line was 716 pixels. Children were required to mark the correct position of a given number by means of the mouse cursor. Number line and target number remained on the screen until a response was given by a mouse click. Both NLE tasks were preceded by a practice trial (5,000 for the bounded NLE task and 1 for the unbounded NLE task) to ensure that the children understood the task and to familiarize the children with the use of the computer mouse. In the practice trial no feedback was provided on the children's estimation accuracy. In the bounded version (range 0-10,000*)*, a total of 24 target numbers were presented in randomized order (i.e., 74, 135, 1097, 1203, 2137, 2315, 3408, 3476, 4542, 4712, 4798, 4957, 5103, 5239, 5298, 5372, 6594, 6781, 7685, 7812, 8793, 8946, 9786, and 9851). In the unbounded version (range 0-29), participants had to position 24 numbers on a number line, of which only the start and a unit (i.e., the length corresponding to 1) but no end-point were indicated. The set of target numbers to be located ranged from 2 to 25, leaving sufficient distance between the end of the physical line (i.e., 29) and the largest target number (i.e., 25) as children tend to overestimate in this task.

2.2.3. Spelling competencies

Spelling assessment followed a writing-to-dictation task which required children to complete gapped sentences (e.g., “Eine Bananenschale ist …” [Abfall]; engl. “A banana peel is …” [garbage]). First, each sentence was presented auditorily, followed by a repetition of the target word (depicted in square brackets). Children were instructed to read along the sentences during dictation, wait until the target word was repeated and then type in the respective target word. Overall, 55 test words were administered with respect to the following orthographic rules: capitalization (55 words, e.g., ‘**A**bfall’, [engl. ‘garbage’]), consonant doubling (21 words, e.g., ‘Abfa**ll’**, [engl. ‘garbage’]), lengthening (14 words, e.g., ‘D**ie**b’, [engl., ‘thief’]) and rule words (i.e., words for which the spelling cannot be derived by the phoneme structure; 14 words, e.g., ‘**Lok**’, [engl. ‘loco’]. In this example, the short vowel is not followed by a double consonant, as one would assume from the German pronunciation). Table A1.2 in Appendix A1 presents the entire test material for the spelling assessment. Individual spelling was assessed automatically as either correct (scored 1) or incorrect (scored 0) with respect to each orthographic rule in each test word (e.g., ‘**A**bfa**ll’** [engl. ’garbage’] was assessed concerning (inner sentence) capitalization and consonant doubling; correct spelling of ´**A**bfa**ll**´ leads to a word score of max. 2) resulting in a maximum overall sum score of 104.

2.3. Data Analysis

2.3.1. Estimation accuracy and strategy use

To compare number line estimation performance in both bounded and unbounded NLE tasks, children’s general estimation accuracy was analyzed using the percent absolute error as dependent variable (PAE=|Estimate− Target number|/ Scale ∗100; cf. Siegler and Booth, 2004). PAE is the standard dependent variable in the literature when evaluating NLE performance across different number ranges, age groups, versions of number line estimation tasks, etc. (e.g., Booth & Siegler, 2008; Kim & Opfer, 2017; Link et., 2014; Siegler & Booth, 2004; Slusser et al., 2012; Slusser & Barth, 2017; Young & Booth, 2015). Furthermore, contour analyses (cf. Ashcraft & Moore, 2012) were conducted separately for each grade to contrast children’s estimation errors around and in between reference points (e.g., start, end, quartile, and mid points) as an index of the use of proportion-judgment strategies (cf. Ashcraft & Moore, 2012) in bounded NLE. Estimates of the following numbers were pooled to reflect the usage of reference points beginning from start (74 and 135) via midpoint (4957 and 5103) to the end point (9851 and 9786). Numbers around the first (2315 and 3408) and third quartile (6781 und 7685) in-between these reference points were combined. In unbounded NLE, use of reference points was not expected as the absence of an end point of the scale prevents the possibility of a reference. Nevertheless, for better comparability of the results, estimates of target numbers around the start (2 and 3), first quartile (7 und 8) mid (13 and 14), third quartile (19 und 20) and end point (24 and 25) of the number line scale were combined as well. Data points of the contour analysis were then compared using an analysis of variance (ANOVA).

Additionally, we conducted model fittings on children`s individual estimates to validate number line estimation strategies derived from contour analyses. Data were modeled with Matlab 9.4.0 (MATLAB and Statistics Toolbox Release R2018a, The MathWorks, Inc., Natick, Massachusetts, United States.) using the trust region algorithm for non-linear model fitting.

2.3.2. Number line estimation and basic arithmetic

Association of number line estimation performance and basic arithmetic operations was assessed by means of partial correlations. With respect to correlation analyses, the Shapiro–Wilk test (Shapiro & Wilk, 1965) was conducted in order to assess normal distribution. As distribution of basic arithmetic and spelling data deviated significantly from normal distribution, Spearman correlation coefficients were calculated. In particular, semi-partial rank correlation coefficients of non-parametric data ("Spearman's rho") using the R package ppcor (Kim, 2015) were calculated to examine the correlation of number line estimation and the four basic arithmetic operations (i.e., addition, subtraction, multiplication, division), controlling for age and spelling ability. In order to validate results, cross-validation methods were applied. For this purpose, existing samples were randomly divided into two data sets each. As cross-validation revealed reliable outcomes in line with results for the overall sample, only latter results are presented below. Cross-validation results are reported in Table A2 in Appendix A2. Bonferroni-Holm procedure (Holm, 1979) was applied to (adequately) correct for multiple testing.

3. Results

In total, complete data sets of 573 children for the bounded NLE task, in particular, 198 fifth graders (94 female, mean age: 10.48, *SD* = 7.32 months), 231 sixth-graders (105 female, mean age: 11.69, *SD* = 8.64 months) and 144 seventh-graders (74 female, mean age: 12.54, *SD* = 8.28) entered analyses. For unbounded NLE, 221 data sets were analyzed, this means, data of 86 fifth graders (41 female, mean age: 10.62, *SD* = 8.05 months), 81 sixth-graders (37 female, mean age: 11.72 *SD* = 8.28 months) and 54 seventh-graders (22 female, mean age: 12.75, *SD* = 9.30 months). Children of the two samples did not differ significantly according to their sum scores in terms of spelling (5^{th} grade: $t(158.82) = -0.018$, $p = 0.99$, 6^{th} grade: $t(144.31) = -1.13$, $p = 0.26$, 7^{th} grade: $t(90.12) = 0.59$, $p = 0.59$) and basic arithmetic (5^{th} grade: $t(160.86) = -0.26$, $p = 0.79$, 6^{th} grade: $t(197.5) = -1.71$, $p = 0.09$, 7^{th} grade: $t(100.4) = -0.72$, $p = 0.47$) as assessed with Welch two sample *t*-tests. Moreover, at SP_2 bounded and unbounded NLE performance was assessed within the same school classes. Thus, school- and classroom related influences (e.g., school location, class size, teaching style, average math performance) were equal for both groups. Table 1 provides descriptive information regarding all variables used in the study.

A 2 (condition: bounded vs. unbounded NLET) x 3 (5^{th} vs. 6^{th} vs. 7^{th} grade) analysis of variance (ANOVA) on PAEs revealed a significant main effect for condition [$F(1,788) = 412.21$, $p < .001$, $\eta^2_p = 0.35$], such that average PAE was significantly lower for bounded NLE (*M* = 5.57, *SD* = 1.49) than for unbounded NLE (*M* = 9.57, *SD* = 4.07). The main effect for grade [$F(2,788) = 7.95$, $p < .001$, $\eta^2_p = 0.02$] indicated a significant influence of grade levels on NLE performance. Furthermore, there was no interaction between condition and group [$F(2,788) = 0.605$, $p > 0.05$, $\eta^2_p < 0.01$]. Post hoc comparisons for grade using Tukey HSD test applying an adjusted *p*-level for multiple comparisons showed a marginally significant [$p = .056$] difference between mean PAE of children in the 5^{th} (*M*=5.95, *SD*=1.66) and 7^{th} grade (*M*=5.17, *SD*=1.42,) in the bounded condition. In the unbounded condition, no significant differences in mean PAEs were observed between grade levels [all $p > .284$; 5-graders (*M*=9.77, *SD*=4.16), 6-graders (*M*=9.76, *SD*=4.39), and 7-graders (*M*=8.97, *SD*=3.36)]. However, this result is also likely to be attributed to lower power generated by the smaller sample size in the latter sample.

Table 1: Demographics

A) Demographics bounded sample

	Fifth grade		Sixth grade		Seventh grade	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Spelling abilities	83.44	11.71	85.55	11.84	91.59	8.65
Addition	18.24	2.79	17.94	3.03	18.59	2.64
Subtraction	14.98	4.85	14.11	5.43	15.93	4.39
Multiplication	17.47	3.38	17.21	3.84	18.08	2.30
Division	17.90	2.97	17.75	3.16	18.52	1.82
NLE task mean PAE	5.94	1.66	5.49	1.28	5.17	1.42

B) Demographics unbounded sample

	Fifth grade		Sixth grade		Seventh grade	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Spelling abilities	83.46	11.95	87.25	11.44	90.74	9.23
Addition	18.49	2.28	18.14	2.73	19.24	1.70
Subtraction	15.22	4.73	14.35	4.90	16.43	4.18
Multiplication	17.37	3.34	18.20	2.18	18.18	2.37
Division	17.90	3.46	18.54	2.48	18.28	2.51
NLE task mean PAE	2.83	1.20	2.82	1.27	2.60	0.97

3.1. Number line estimation accuracy and strategy use

Regarding children's performance in the bounded NLE task Figure 1 revealed estimation errors to vary for different target numbers. In contrast, children's estimation errors increased linearly in the unbounded NLE task. Summarizing estimation errors at/around reference points, contour analyses revealed an *M*-shaped pattern of error distribution (see Figure 1A) which has been argued to be characteristic for proportion-judgment strategies (cf. Cohen and Blanc-Goldhammer, 2011). Correlating PAEs and size of target number revealed no significant correlations for all grade levels (all $r < 0.333$, all $p > 0.11$). Bonferroni-Holm corrected *t*-tests were conducted to compare PAEs at specific reference points (start-, mid- and endpoint) and in the intervals between them (first and third quartile). Results substantiated the descriptive pattern of the contour analyses by indicating significantly reduced PAEs at reference points as compared to the intervals in-between for all grades [5th grade: $t(192) > -7.23$, $p < .001$, Cohen's d = 0.23 – 1.14; 6th grade: $t(227) > -21.53$, $p < .001$, Cohen's d = 0.40 – 1.43, and 7th grade: $t(142) > -14.79$, $p < .001$, Cohen's d = 0.37 – 1.23]. In unbounded NLE, in contrast, linearly increasing PAEs across all age groups were observed: PAEs increased monotonously with target number (see Figure 1B) resulting in significant correlations between PAEs and size of target number (from $r = 0.931$ to $r = 0.952$, all $p < 0.05$). This pattern was also found in contour analyses (Figure 1C). Bonferroni-Holm corrected *t*-tests revealed significant differences between both the start- and midpoint and the quartile [5th

grade: $t(85) > -12.09$, $p < .001$, Cohen's d = 0.58 – 1.30; 6th grade: : $t(80) > -12.09$, $p < .001$, Cohen's d = 0.25 – 1.50, and 7th grade: $t(53) > - 10.14$, $p < .001$, Cohen's d = 0.31 – 1.38], but not for the third quartile and the endpoint [5th grade: $t(85) > -1.64$, $p > 0.05$, Cohen's d = 0.18; and 7th grade: $t(53) > -2.25$, p(adj.) > 0.05, Cohen's d = 0.01), except for the 6th grade ($t(80) > 2.5432$, $p < .001$, Cohen's d = 0.28].

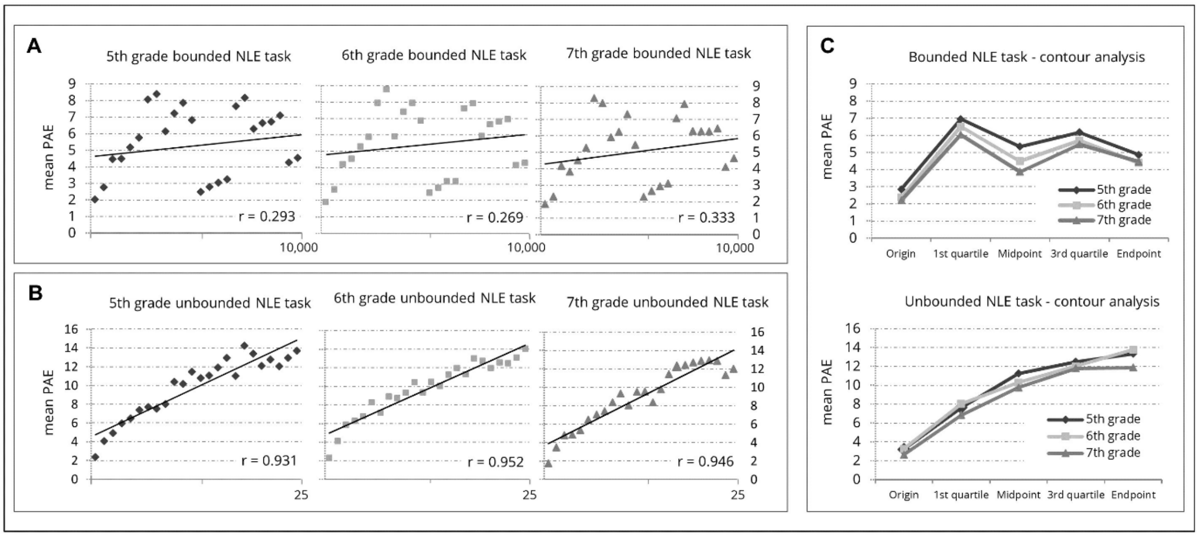


Fig. 1: The left panel depicts mean PAE and target numbers in the (A) bounded and (B) unbounded NLE task separately for all grade levels. The right panel (C) shows contour analyses summarizing percent absolute errors (PAE) at specific reference points (cf. Ashcraft and Moore, 2012) for all grade levels (i.e., grade 5-7) again for bounded (top) and unbounded (bottom) NLE.

Additionally, we fitted children's number line estimates with different mathematical models. Table 2 shows absolute and relative frequencies (in parenthesis) of best fitting models for bounded and unbounded NLE separated for the three different grade levels. In bounded NLE, linear models provided best fit for all grade levels (for 56% of fifth, 55% of sixth, and 56% of seventh graders). Only about 30% of all children's estimates were best determined by models indicating proportion-judgment strategies (i.e., one-cyclic or two-cyclic models), and thus the use of reference points. This finding may contradict evidence for proportion-judgment from the contour analyses. In unbounded NLE, over 90% of children's estimates were accounted best by linear (for 63% of the fifth, 68% of the sixth, and 65% of the seventh graders) or power models (for 29% of the fifth, 30% of the sixth, and 33% of the seventh graders) for all grade levels. Model fitting did not indicate model fits in favor of dead-reckoning in unbounded number line estimation. Crucially, and in particular for unbounded NLE, mean adjusted R2 resp. the AIC values were almost identical. As such, a unique derivation of the best fitting model seems not feasible.

Table 2: Model fittings for bounded and unbounded number line estimation

A		**Bounded NLE task**			**B**		**Unbounded NLE task**		
		5th	**6th**	**7th**			**5th**	**6th**	**7th**
	N	198	232	144		N	86	81	54
linear model	adj R²	0.945	0.954	0.962	**linear model**	adj R²	0.877	0.884	0.891
	AIC	281.453	283.290	287.868		AIC	35.573	37.565	37.684
	Δ_AIC	0	0	0		Δ_AIC	0	0	0
	n	**112 (56)**	**128 (55)**	**81 (56)**		n	**54 (63)**	**55 (68)**	**35 (65)**
Power model	adj R²	0.944	0.954	0.961	**power model**	adj R²	0.874	0.878	0.883
	AIC	281.655	283.569	288.116		AIC	36.599	38.917	39.385
	Δ_AIC	0.20	0.28	0.25		Δ_AIC	1.026	1.352	1.702
	n	32 (16)	27 (12)	20 (14)		n	25 (29)	24 (30)	18 (33)
one-cyclic model	adj R²	0.929	0.936	0.945	**dual scallop model**	adj R²	0.876	0.879	0.886
	AIC	287.959	290.357	295.015		AIC	36.734	39.512	39.785
	Δ_AIC	6.51	7.07	7.15		Δ_AIC	1.16	1.95	2.10
	N	31 (16)	35 (15)	23 (16)		n	6 (7)	2 (2)	1 (2)
two-cyclic model	adj R²	0.9273	0.9388	0.9472	**mutli scallop model**	adj R²	0.872	0.878	0.886
	AIC	288.749	290.450	295.241		AIC	37.300	39.520	39.697
	Δ_AIC	7.30	7.16	7.37		Δ_AIC	1.73	1.96	2.01
	N	23 (12)	41 (18)	20 (14)		n	1 (1)	0 (0)	0 (0)

Note: Absolute and relative frequency (percentages) of best fitting models for bounded (A) and unbounded (B) number line estimation separated for grade levels. The best fitting models are indicated in bold script.

Regarding the partially contradictory results of contour analyses and model fittings for bounded NLE, closer visual examination of children's individual estimation pattern was more revealing. Table 3 provides results of individual contour analysis for each child separated for the best fitting models as well as for grade level. The table differentiates between the M-shape pattern of number estimates and other patterns (i.e., N-shape, inverse N-shape, V-shape, etc.). At first glance, more than one third of fifths and sixth graders showed an M-shape estimation pattern on the individual level. This number increased in the seventh grade up to almost fifty percent indicating an increasing use of reference points with age. Yet, these findings cannot be clearly derived from the results of individual model fittings: successful use of reference points (i.e., proportion-judgment) and thus minimal estimation errors can be also accounted by linear models (e.g., Link, Huber, et al., 2014). Contrary, unsuccessful use of reference points (i.e., a V-shape or N-shape pattern) can be accounted by one- or two-cyclic models.

Table 3: Individual contour analyses and model fittings for bounded NLE

	5th grade		6th grade		7th grade	
	M-Shape	**Other**	**M-Shape**	**Other**	**M-Shape**	**Other**
linear model	41	71	50	78	40	41
power model	8	24	10	17	8	12
one-cylic model	8	23	14	21	13	10
two-cyclic model:	6	17	9	32	6	14
∑	63 (32)	135 (68)	83 (36)	148 (64)	67 (47)	77 (53)

Note: Absolute and relative frequency (percentages) for the use of proportion-judgment strategy (i.e., M-Shape estimation pattern) or other strategies derived from contour analysis in bounded number line estimation.

3.2. Number line estimation and basic arithmetic

First, we analyzed to what extent spelling and arithmetic skills were associated to control for influences of general cognitive abilities in subsequent analyses. Partial correlations controlling for age revealed relatively high correlations between basic arithmetic skills and overall spelling skills for both samples assessed with the bounded (n=573, r=.51, p<.001) and unbounded (n=221, r = .43, p<.001) NLE task. These correlations indicated a strong association between spelling and arithmetic performance. As spelling and arithmetic skills have been found to rely on common general cognitive abilities (e.g., working memory, verbal memory, and reasoning skills, e.g., Knievel, Daseking, & Petermann, 2010; Passolunghi, Mammarella, & Altoè, 2008), we used overall spelling skills as a proxy to control for influences of general cognitive ability. Subsequently, partial correlations were calculated controlling for age and spelling skills. Table 4 shows the results of the partial correlation between number line estimation performance and basic arithmetic separately for grade levels.

For bounded NLE, strength of correlations increased with age. In grade 5, NLE performance correlated significantly with subtraction (i.e., lower error rate in subtraction problems was associated with less estimation errors in NLE), but not with addition, multiplication, or division. In grade 6, significant correlations were found for addition, subtraction and multiplication, but not for division. In grade 7, bounded NLE performance was found to be significantly associated with all basic arithmetic operations, but highest with addition and subtraction. For unbounded NLE, no significant correlation was found between task performance and basic arithmetic.

Table 4: Partial correlations (one-tailed) between PAEs of bounded (left panel) and unbounded NLE and basic arithmetic

	Bounded NLE task				Unbounded NLE task			
Grade	***Add***	***Sub***	***Mul***	***Div***	***Add***	***Sub***	***Mul***	***Div***
5th grade	-0.09	**-0.20****	-0.08	-0.03	0.06	-0.14	0.03	-0.06
6th grade	**-0.29****	**-0.23****	**-0.21****	-0.06	-0.08	-0.05	0.02	0.06
7th grade	**-0.43****	**-0.44****	**-0.24****	**-0.25***	0.03	-0.17	0.14	0.05

Note: Overall partial correlations were calculated for all grade levels. Add=Addition, Sub=Subtraction, Mul=Multiplication, Div=Division. * p<.05, ** p<.001.

4. Discussion

In the following section, we discuss results on number line *estimation accuracy and strategy use* and the association of *NLE and basic arithmetic* separately with regard to the relevant literature.

4.1. Estimation accuracy and strategy use

The first objective of this study was to examine NLE performance in secondary school children. We expected results in estimation accuracy and solution strategies to differ for bounded (i.e., Siegler & Opfer, 2003) and unbounded (i.e., Cohen & Blanc-Goldhammer, 2011) NLE. More specifically, we assumed higher estimation accuracy in the bounded as compared to the unbounded task. Regarding developmental aspects, we expected a significant increment of estimation accuracy with age for the bounded but not for the unbounded task version (Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014).

Results substantiated that secondary school children performed differently in both tasks: estimation errors were almost twice as high in unbounded than in bounded NLE. These findings are consistent with previous results on primary school children (i.e., grade 3 and 4) and adults (e.g., Link, Huber, et al., 2014). In addition, comparably high PAEs for unbounded NLE have been reported by Kim and Opfer (2017). However, present findings also contradict results from Cohen and Sarnecka (2014) who found that children performed better in unbounded NLE (i.e., more accurate estimations) as compared to bounded NLE. The authors suggested that both tasks build on basic arithmetic operations of varying difficulty (i.e., easier addition/multiplication for unbounded NLE and more sophisticated subtraction/division for bounded NLE) and concluded the unbounded NLE task to be less difficult. Kim and Opfer (2017) questioned these results and attributed them to task sequence influences. Analyzing task sequence effects of bounded and unbounded NLE, they found higher estimation errors in the bounded version of the task (range 0-1,000) when the unbounded version was presented first. The authors considered this finding to be due to fatigue effects or confusion in the subsequent task and concluded from that unbounded NLE to be more demanding. As our study did not examine both NLE tasks within the same person, we cannot draw firm conclusions on possible sequence effects. However, a look at the number intervals used in our study might provide an argument against the conclusions of Cohen and Sarnecka (2014). For unbounded NLE, the number range covered was much smaller (i.e., 0-25) than that used for bounded NLE (i.e., 0-10,000). Although one may assume that all children were familiar with both intervals, the larger number interval should nevertheless be more error prone due to the problem size effect (Zbrodoff & Logan, 2005, for a review). However, mean estimation errors in bounded NLE were significantly lower. This result could arise due solution strategies applied: In bounded NLE, application of proportion-based strategies simplifies task solution; the presence of a start and an end point facilitated the derivation of the center of the number line and also decomposition into sections. This procedure reduces the number of mistakes made by the children. Thus, estimation errors were reduced. Unbounded NLE does not offer this solution strategy as the end point is missing.

Contour analysis in the present study showed that solution strategies differ between bounded and unbounded NLE. In line with Ashcraft and Moore (2012) as well as with Link et al. (2014) a characteristic *M*-shaped pattern of estimation errors (i.e., smaller and less variable estimation errors for target numbers at or around reference points) were observed in contour analyses for bounded NLE. This suggested that children used proportion-based strategies to solve the task. By contrast, we observed an almost linearly increasing error pattern for unbounded NLE, which has been previously assumed to be related to a magnitude estimation-based strategy in unbounded NLE (Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014).

However, these findings were only partially corroborated by model fitting results. Model fittings for bounded NLE indicated best fit of linear models for the majority of children's estimates. Crucially, on an individual level, almost half of the children whose estimation patterns was determined best by the linear model showed an M-shaped estimation pattern in contour analyses. As recently discussed by Link and colleagues (2014), a very accurate estimation pattern can hardly be differentiated by model fittings. Contrarily, model fitting results for unbounded NLE were unambiguous. They did not provide any evidence for specific strategy use (i.e., proportion-judgment) in unbounded number line estimation which was also suggested by visual inspection in contour analyses. Taken together, conclusions about solution strategies in NLE derived from model fittings seem to be less conclusive than those derived from contour analyses. However, contour analysis may not achieve the resolution level of model fittings due to higher level of data aggregation.

Furthermore, bounded number line estimation performance in primary school children was found to improve with age (e.g., Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014; Slusser & Barth, 2017). This assumption was also true for secondary school children assessed in the present study as 5th-graders' estimation accuracy tended to be lower than that of children in grade 7. This argument is further supported by the increasing number of children from grade 5 to grade 7 for whom an M-Shape pattern and thus a use of reference points was indicated by the contour analysis. In contrast, for unbounded NLE no significant differences in mean PAEs were observed across grade levels. Interestingly, however, we observed age-related increase from 5th to 7th grade for both basic arithmetic and spelling performance which were considered as control variables. Considering the observation that estimation performance does not change in unbounded NLE across the different age groups, the question arises to what extent unbounded NLE performance is influenced by general age-related developmental processes.

Taken together, our data provide convergent evidence with previous studies on the assumption that secondary school children seem to apply proportion-based estimation strategies in bounded but not in unbounded NLE (cf. Cohen & Blanc-Goldhammer, 2011; Link, Huber, et al., 2014; Slusser et al., 2013). Nevertheless, results do not yet provide an explanation of the extent to which basic arithmetic operations are needed to solve bounded and unbounded NLE tasks.

4.2. Number line estimation and basic arithmetic

The second aim of the present study was to shed light on the association of NLE performance and basic arithmetic. Our results are consistent with findings of Cohen and Sarnecka (2014) and Link et al. (2014) that basic arithmetic skills were associated to secondary school children's estimation performance in bounded NLE. Moreover, correlation coefficients were comparable to those reported in other studies (ranging from r = 0.29 – 0.86) relating various mathematical competencies (i.e., problem solving, calculation, etc.) to bounded NLE performance (see Schneider et al., 2018, for an overview). However, in contrast to the theoretical suggestions of Cohen and Sarnecka (2014) but in line with the data of Link et al. (2014) for primary school children, none of the basic arithmetic operations was associated with unbounded NLE performance. In particular, we observed subtraction skills to be associated with bounded NLE performance consistently for all grades. However, in grade 6 correlations with addition and multiplication were of equal strength as with subtraction. In grade 7, all four arithmetic operations were significantly associated with bounded NLE.

Regarding our results, children who are well versed in solving subtraction tasks may be more likely to successfully use proportion-judgment strategies and, thus, show higher NLE accuracy. Similar to solving subtraction problems, subtraction upon the number line seems to be more difficult in comparison to addition (Artemenko, Pixner, Moeller, & Nuerk, 2018), and thus, to be more error prone. A closer descriptive look at the results of the contour analysis (Figure 1) and mean PAEs at the quartiles supports this assumption: starting from the mid-point a leftward shift (from mid-point to the first quartile reflecting subtraction) resulted in higher estimation errors than a shift to the right (from mid-point to the third quartile reflecting addition).

Several factors may account for the significance of subtraction for the bounded NLE performance: First, addition and subtraction are taught as the first arithmetic operations in school (e.g., Bildungsplan Sekundarstufe I, Ministerium für Kultus, Jugend und Sport, Baden Württemberg, 2016). However, subtraction is regarded to be more demanding as compared to addition (e.g., because of borrowing, Artemenko et al., 2018), and children were found to use calculation (i.e., instead of fact retrieval in simple addition and multiplication problems) to solve subtraction problems even in 2-digit numbers (e.g., Ischebeck et al., 2006). Thus, subtraction might be more sensitive to reflect differences in magnitude estimation skills as it also relies on magnitude manipulations to a higher degree. Second, according to the proportion-judgment account, a leftward (reflecting subtraction) or rightward shift (reflecting addition) starting from a reference point might be necessary to locate the number correctly upon the line (cf. Cohen and Sarnecka, 2014). Thus, children's subtraction and addition skills may be more likely to predict NLE performance. Third, further evidence may be derived from neuro-functional level: Berteletti and colleagues (2015) demonstrated that activation of brain areas associated with number magnitude processing (as required in subtraction) were correlated with NLE performance. Furthermore, Ischebeck et al. (2006) provided neuro-functional evidence that a training of subtraction problems in young adults encouraged the application of efficient procedural solution strategies. This result suggests that children with high subtraction skills are more likely to apply sufficient solution strategies (i.e., proportion-judgment) at least in bounded NLE.

An explanation for significant associations between bounded NLE and multiplication as well as division occurring not before grade 6 and 7, respectively, might be that, according to German math curricula, multiplication and division are increasingly used with the introduction of fractions in grade 6 (e.g., Bildungsplan Sekundarstufe I, Ministerium für Kultus, Jugend und Sport, Baden Württemberg, 2016). Although all four basic arithmetic operations have already been acquired in primary school (i.e., before grade 5; Huber, Moeller, & Nuerk, 2012), the increasing significance of multiplication and division (i.e., fractions) from grade 6 onwards might enable children to apply proportion-judgment strategies more proficiently, leading to better performance in bounded NLE. Visual inspection of contours on individual levels confirmed the increasing use of reference points with age. This in turn can influence NLE performance as presented by Siegler and colleagues (2011) in another domain of mathematics (i.e., fractions) showing that segmentation strategies involving dividing up the number line with subjective landmarks leads to more accurate fraction number line estimates.

Moreover, correlation coefficients became stronger for subtraction, addition, multiplication and finally division from grade 5 to 7, suggesting an increasing importance of basic arithmetic for bounded NLE. These findings substantiate the assumption of Friso-van den Bos et al. (2015) who argued that associations between arithmetic skills and NLE are predominantly driven by developmental processes.

For the association of unbounded NLE performance and basic arithmetic, we did not find significant relations with any basic arithmetic operation in secondary school children. These results are in line with those of Link et al. (2014) who did not observe an association between unbounded NLE and addition or subtraction skills in their sample of fourth graders. Building on the results of Cohen and Blanc-Goldhammer (2011) and Reinert et al. (2015), we would have expected at least a weak correlation of unbounded NLE with addition and multiplication, respectively. Even though, and unlike Reinert et al. (2015), we did neither focus on analyzing different unit distances (0-1, 0-5, etc.) nor specifically selected multiples of these as target numbers. Irrespective of this issue, our data did not show any association between unbounded NLE and basic arithmetic. This finding seems plausible considering the assumption that unbounded NLE may be solved by iterative counting like strategies (Reinert et al., 2015; Schneider et al., 2018). Counting as a highly automated process may not depend on basic numerical and arithmetical skills as needed for proportion-judgment in bounded NLE. In this vein, unbounded NLE seems to reflect purer numerical estimation, which may be superimposed less by underlying arithmetic processes (i.e., addition, subtraction, division) of other solution strategies as is proportion-judgment (see also Cohen &

Blanc-Goldhammer, 2011; Link, Huber, et al., 2014; Reinert et al., 2015). In sum, the present data provide additional evidence from secondary school children that unbounded number line estimation may indeed be solved primarily based on processing and estimating number magnitude. When interpreting these results, there are some constraints that need to be considered. First, in the present study different number ranges were used in bounded (0-10,000) and unbounded (0-29) number line estimation to tie on results of previous studies (e.g., Link, Huber, et al., 2014; Siegler & Opfer, 2003; Thompson & Opfer, 2010). Results documented unbounded NLE to be more difficult (higher mean PAEs for all grade levels). Conversely, this result implied that the application of a specific solution strategy (i.e., proportion judgment) seemed to simplify task solution in bounded NLE although the number range was significantly higher. These findings fit nicely with those of other studies directly comparing bounded and unbounded NLE using the same (Kim & Opfer, 2017) or different number ranges (Link, Huber, et al., 2014; also for a methodological discussion on the application of different number ranges). For all ranges and age groups investigated so far, it has been observed that estimation errors in unbounded NLE increase linearly with the size of the target number. However, this is clearly different for the bounded NLE task, for which the error pattern was shown to differ with age for the same range and over ranges at the same age. Both results (no developmental changes for unbounded but for bounded NLE) were replicated in the present study. Nevertheless, further research would be desirable to investigate potential influences of different number ranges on estimation performance and patterns more systematically.

Second, the study was cross-sectional, so we did not monitor intra-individual development in bounded and unbounded NLE. However, conclusions on numerical development with age seem warranted as for both spelling and basic arithmetic tasks our data documented increasing performance with increasing age. Nevertheless, it would be desirable to pursue intra-individual developmental changes in estimation accuracy in secondary school children more systematically in future studies. Third, as we assessed performance in bounded and unbounded NLE inter-individually, direct comparisons of both NLE tasks within one child were not possible. Nevertheless, both samples seem well comparable as bounded and unbounded NLE was assessed within the same school classes (at least at SP_2), and thus, class-room related influences (e.g., class size, teaching style, average math performance of the school class) were equally distributed across both groups. Moreover, and this is particularly important, the two groups solving either bounded or unbounded NLE neither differed significantly in mean basic arithmetic skills nor in spelling performance. We therefore assume that both groups are well comparable and sufficient to draw comparable conclusions on bounded and unbounded NLE performance. However, individual comparisons would be desirable for future studies. Finally, it is not clear yet whether and how the presentation medium (i.e., computerized assessment in the current study) moderates NLE (Schneider et al., 2018), and potentially leads to inconsistent results (e.g., increased error variability) as compared to earlier paper-pencil studies. Future mode effect studies on NLE comparing computerized and paper-pencil testing could provide further insights.

5. Conclusion

In the current study, we investigated bounded and unbounded NLE in secondary school children (i.e., grade 5 to 7). We found that children's performance in bounded NLE was significantly better as compared to unbounded NLE for all grade levels. Moreover, estimation accuracy seems to improve with age only in bounded NLE. With respect to the association of NLE performance with basic arithmetic, we observed that estimation performance in bounded but not unbounded NLE was strongly associated with basic arithmetic operations (i.e., addition, subtraction, multiplication, division). Interestingly, these associations increased with age, and thus, are probably subject to developmental changes. With respect to solution strategies employed in NLE, current results corroborated the assumption of bounded NLE to rely on proportion-based estimation strategies. For unbounded NLE we did not find indications to support recent hypotheses on strategies and arithmetic competencies specifically used in task performance (cf. Cohen & Sarnecka, 2014). However, conclusions from model fittings about solution strategies in NLE seemed to be less conclusive as compared to those derived from contour analyses - at least for bounded number line estimation. In summary, these data indicate that differential results on the association between bounded and unbounded NLE with arithmetic operations obtained in primary school generalize to older secondary school children. Results substantially contribute to the knowledge about the (consecutive) development of skills pertaining to bounded and unbounded number line estimation. In particular, these findings provide additional evidence for the assumption that unbounded NLE may indeed reflect a more direct and pure measure of number magnitude estimation.

Acknowledgements

Stefanie Jung was supported by the first Leibniz-WissenschaftsCampus Tuebingen "Bildung und Informationsumwelten" as well as by the German Research Foundation (DFG) providing funding to Elise Klein (KL 2788/2-1). Elise Klein was supported by the Margarete von-Wrangell Fellowship (European Social Fund and the Ministry of Science, Research and the Arts Baden-Württemberg). We thank the participating children, their parents and institutions who made this work possible. We also thank all research assistants for their assistance with data collection.

References

Artemenko, C., Pixner, S., Moeller, K., & Nuerk, H. C. (2018). Longitudinal development of subtraction performance in elementary school. *British Journal of Developmental Psychology*, *36*(2), 188–205.

Ashcraft, M. H., & Moore, A. M. (2012). Journal of Experimental Child Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology*, *111*(2), 246–267. h

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*(1), 125–135.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, *46*(2), 545.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189–201.

Booth, J. L., & Siegler, R. S. (2008). Numerical Magnitude Presentations Infleunce Arithmetic Learning. *Child Development*, *79*(4), 1016–1031.

Bos, I. F. Den, Kroesbergen, E. H., Luit, J. E. H. Van, Xenidou-dervou, I., Jonkman, L. M., & Schoot, M. Van Der. (2015). Journal of Experimental Child Longitudinal development of number line estimation and mathematics performance in primary school children. *Journal of Experimental Child Psychology*, *134*, 12–29.

Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin and Review*, *18*(12), 331–338.

Cohen, D. J., & Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations). *Developmental Psychology*, *50*(6), 1640–1652.

Ebersbach, M., Luwel, K., & Verschaffel, L. (2015). The Relationship Between Children ' s Familiarity with Numbers and Their Performance in Bounded and Unbounded Number Line Estimations. *Mathematical Thinking and Learning*, *17*(2–3), 136–154.

Hollands, J. G., & Dyre, B. P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review*, *107*(3), 500–524.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Huber, S., Moeller, K., & Nuerk, H.-C. (2012). Differentielle Entwicklung arithmetischer Fähigkeiten nach der Grundschule: Manche Schere öffnet und schließt sich wieder. *Lernen Und Lernstörungen*, *1*, 119–134.

Kim, D., & Opfer, J. E. (2017). A unified framework for bounded and unbounded numerical estimation. *Developmental Psychology*, *53*(6), 1088–1097.

Kim, S. (2015). ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Communications for Statistical Applications and Methods*, *22*(6), 665–674.

Lenhard, W., & Lenhard, A. (2014). Hypothesis tests for comparing correlations. Bibergau, Germany: Psychometrica.

Link, T., Huber, S., Nuerk, H. C., & Moeller, K. (2014). Unbounding the mental number line-new evidence on children's spatial representation of numbers. *Frontiers in Psychology*, *4*, 1–12.

Link, T., Nuerk, H. C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *Quarterly Journal of Experimental Psychology*, *67*(8), 1597–1613.

Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H. (2009). Journal of Experimental Child Children ' s early mental number line : Logarithmic or decomposed linear? *Journal of Experimental Child Psychology*, *103*(4), 503–515.

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science*, *14*(5), 1194–1204.

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition*, *149*, 11–17.

Peeters, D., Degrande, T., Ebersbach, M., Verschaffel, L., & Luwel, K. (2016). Children's use of number line estimation strategies. *European Journal of Psychology of Education*, *31*(2), 117–134.

Reinert, R. M., Huber, S., Nuerk, H., & Moeller, K. (2015). Strategies in unbounded number line estimation ? Evidence from eye-tracking. *Cognitive Processing*, *16*(1), 359–363.

Reinert, R. M., Huber, S., Nuerk, H., & Moeller, K. (2017). Sex differences in number line estimation: The role of numerical estimation. *British Journal of Psychology*, *108*(2), 334–350.

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental Number Line, Number Line Estimation, and Mathematical Achievement: Their Interrelations in Grades 5 and 6. *Journal of Educational Psychology*, *101*(2), 359–372.

Schneider, M., Merz, J., Sticker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development*.

Schneider, M., Thompson, C. A., & Rittle-Johnson, B. (2017). Associations of magnitude comparison and number line estimation with mathematical competence: A comparative review. In *Cognitive Development from a Strategy Perspective: A Festschrift for Robert Siegler* (pp. 100–119). Taylor and Francis.

Shapiro, S., & Wilk, M. B. (1965). Biometrika Trust An Analysis of Variance Test for Normality (Complete Samples) Published by : Oxford University Press An analysis of variance test for normality (complete sa, *52*(3), 591–611.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The developent of numerical estimation: Evidence for Multiple Representations of Numerical Quantity. *Psychological Science*, *14*(3), 237–243.

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive psychology*, *62*(4), 273-296.

Simms, V., Clayton, S., Cragg, L., Gilmore, C., & Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical achievement: The role of visuomotor integration and visuospatial skills. *Journal of Experimental Child Psychology*, *145*, 22–33.

Slusser, E. B., & Barth, H. C. (2017). Intuitive proportion judgment in number-line estimation: Converging evidence from multiple tasks. *Journal of Experimental Child Psychology*, *162*, 181–198.

Slusser, E. B., Santiago, R. T., & Barth, H. (2012). Developmental Change in Numerical Estimation. *Journal of Experimental Psychology: General*, *142*(1), 193–208.

Thompson, C. A., & Opfer, J. E. (2010). How 15 Hundred Is Like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition. *Child Development*, *81*(6), 1768–1786.

Young, L. K., & Booth, J. L. (2015). Student Magnitude Knowledge of Negative Numbers. *Journal of Numerical Cognition*, *1*(1), 38–55.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem-size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–345). New York: Psychology Press.

Zhu, M., Cai, D., & Leung, A. W. S. (2017). Number line estimation predicts mathematical skills: Difference in grades 2 and 4. *Frontiers in Psychology*, *8*(SEP), 1–8.

Appendix A

Appendix A provides test material used to assess basic arithmetic (Table A1) and spelling competencies (Table A1) of this study.

Table A1: Basic arithmetic problems

N	Addition	Subtraction	Multi-plication	Division
P	12 + 3 = 15			
1	36+47 = 83	91-67 = 24	8x7 = 56	48÷6 = 8
2	29+67 = 96	93-67 = 26	8x6 = 48	56÷7 = 8
3	25+67 = 92	51-37 = 14	8x4 = 32	32÷4 = 8
4	21+27 = 48	94-69 = 25	7x6 = 42	54÷9 = 6
5	65+28 = 93	92-76 = 16	9x7 = 63	42÷6 = 7
6	29+38 = 67	81-65 = 16	6x9 = 54	63÷7 = 9
7	38+25 = 63	41-18 = 23	4x7 = 28	24÷8 = 3
8	27+29 = 56	45-29 = 16	4x6 = 24	72÷8 = 9
9	24+68 = 92	52-27 = 25	4x9 = 36	35÷5 = 7
10	46+36 = 82	33-18 = 15	3x8 = 24	21÷3 = 7
11	53+36 = 89	96-72 = 24	9x8 = 72	18÷3 = 6
12	42+34 = 76	97-71 = 26	3x9 = 27	36÷9 = 4
13	65+21 = 86	36-14 = 22	5x8 = 40	28÷7 = 4
14	28+51 = 79	76-62 = 14	3x4 = 12	40÷8 = 5
15	37+41 = 78	39-26 = 13	7x5 = 35	24÷6 = 4
16	24+45 = 69	29-16 = 13	6x3 = 18	27÷9 = 3
17	41+35 = 76	56-41 = 15	7x3 = 21	12÷4 = 3
18	15+13 = 28	38-21 = 17	6x5 = 30	30÷5 = 6
19	13+24 = 37	39-11 = 28	5x4 = 20	20÷4 = 5
20	41+13 = 54	27-13 = 14	5x3 = 15	15÷3 = 5

Note: *Basic arithmetic problems were* preceded by one practice trial (P) to familiarize the children with the computerized testing.

Table A2: Spelling competencies assessed by gapped sentences

N	Target	C	D	L	R	Gapped sentence	English translation
P	Fuß	x			x	Lisa trat mir auf meinen rechten ...	Lisa stepped on my right ... (foot)
1	Abfall	x	x			Eine Bananenschale ist...	A banana peel is ... (garbage)
2	abmessen	x	x			Mit dem Lineal kann man Längen..	Use the ruler to ... lengths. (measure)
3	Baby	x			x	In der Wiege schläft ein...	In the cradle sleeps a ... (baby)
4	beißen	x				Hunde, die bellen, ... nicht.	Dogs that bark don't ... (bite)
5	Beschmutzen	x	x			Im Park ist das... der Bänke verboten.	It is forbidden to ... the benches in the park. (stain)
6	bieten	x		x		Will man eine Auktion gewinnen, muss man hoch...	If you want to win an auction, you have to ... high. (bid)
7	bitten	x	x			Wenn die Gäste klingeln,... wir sie herein.	We' ll ... them in when the guests ring the bell. (invite)
8	Blumenwiese	x		x		Das Mädchen pflückte Blumen von der ...	The girl picked flowers from the ... (field)
9	Bus	x			x	Beeil dich, wir kommen zu spät zum...	Hurry up, we'll be late for the ... (bus)
10	Chili	x			x	Nimm für das Essen bitte wenig...	Please take a little ... for your meal. (chili)
11	cool	x			x	Seinen neuen Haarschnitt findet Hannes richtig...	Hannes thinks his new haircut is really ... (cool)
12	dehnen	x		x		Vor dem Sport sollte man die Muskeln...	Before doing sports you should ... your muscles. (stretch)
13	dick	x	x			Ein Buch mit 800 Seiten ist...	An 800-page book is ... (thick)
14	Dieb	x		x		Die Handtasche wurde von einem... geklaut.	The purse was stolen by a ... (thief)
15	Entlassung	x	x			Es ist traurig, wenn einem Arbeiter mit der... gedroht wird.	It is sad when a worker is faced with a ...
16	erfahren	x		x		Die Nachricht ist unglaublich, Sarah muss unbedingt davon...	The news is incredible, Sarah needs to ... about it. (know)
17	erzählst	x		x		Die Geschichten sind immer spannend, wenn du sie...	The stories are always exciting when you ... them. (tell)
18	Eselsohr	x		x		Eine umgeknickte Seite eines Buches nennt man...	A folded page of a book is called a... (dog`s ear)
19	fällt	x		x		Es ist schon kalt und die Temperatur ... ständig noch weiter.	It's already cold and the temperature keeps ... (falling)
20	Fenstergriff	x	x			Jedes Fenster muss einen ... haben.	Every window must have a window ... (handle)
21	Geburtstag	x				Meist gibt es einen Kuchen zum...	Usually there is a cake for your ... (birthday)
22	grässlich	x	x			Das verbrannte Essen schmeckte...	The burnt food tasted ... (awful)
23	heißen	x	x			Manche finden das Kunstwerk schön, andere finden es...	Some people like the work of art, others consider it to be ... (ugly)
24	hell	x	x			Die Sonne scheint ...	The sun is shining ... (bright)
25	knurrend	x	x			Ich traute mich nicht ins Haus, weil ein Hund... vor der Türe saß.	I didn't dare come in the house because a dog was ... at the door. (grumbling)
26	Lok	x			x	Die Waggons werden von der... gezogen	The wagons are pulled by the ... (loco)
27	Lot	x			x	Die Mauer muss nach dem... ausgerichtet werden.	The wall must be aligned to the ... (perpendicular)

28	Target	x				Gapped sentence	English translation
29	mahlt	x		x	x	Die Mühlsteine drehen sich, wenn der Müller das Korn...	The millstones turn when the miller ... the grain. (grinds)
30	malt	x				Oma freut sich, wenn Lisa ein Bild für sie...	Grandma is happy when Lisa ... a picture for her. (draws).
31	Mandarinen	x			x	Zu Weihnachten gibt es viele Nüsse und...	There are lots of nuts and ... for Christmas. (tangerines)
32	Maß	x			x	Zentimeter sind ein ... für Längen.	Centimeters are a ... of length. (measure)
33	Mus	x			x	Die reifen Äpfel kochen wir zu...	We cook the ripe apples to ... (puree)
34	Nachsitzen	x	x			Wenn man zu viel Quatsch macht, drohen manche Lehrer mit...	Some teachers threaten ... if you do too much rubbish. (detention)
35	näher	x		x		Von Italien nach Frankreich ist es... als vom Mond zur Erde.	From Italy to France it is ... than from the moon to the earth. (closer)
36	okay	x			x	Den neuen Kinofilm findet Max ganz...	Max thinks the new movie is .... (okay)
37	Orange	x			x	In den Obstsalat gehört auch eine...	An ... also belongs in the fruit salad. (orange)
38	Pony	x			x	Auf der Weide steht ein kleines...	There's a little ... in the pasture. (pony)
39	Portion	x			x	Im Sommer isst man gerne eine große...Eis.	In summer you like to eat a large ... of ice cream. (portion)
40	Pralinen	x			x	Oma liebt Süßes, am meisten mag sie...	Grandma loves sweets, she likes ... the most. (pralines)
41	Rechnen	x				Mit Bruchzahlen ist das... schwieriger als mit ganzen Zahlen.	Fractions are more difficult to ... than integers. (calculate)
42	schafft	x	x			Die Fans feuern den Fahrer an, damit er es auf den ersten Platz ...	All fans cheer on the racer to ... it to the first place. (make)
43	schmierig	x		x		Wenn man zu viel Gel nimmt, werden die Haare oft ...	If you take too much gel, the hair often becomes ... (greasy).
44	sehnst	x		x		Wenn dich die Arbeit anstrengt,... du dich nach Pausen	When you work hard, you ... for breaks. (long)
45	Straßenbahn	x		x		Ohne Ticket darf man nicht mit der ... fahren	You may not take the ... without a ticket. (tram)
46	stumm	x	x			Ich hatte einen Schock, konnte nichts sagen und blieb ganz ...	I was in shock, couldn't say anything and remained completely ... (mute).
47	stützt	x	x			Der Kranke ... sich auf die Krücke	The patient ... on the crutch. (leans)
48	treffen	x	x			Können wir uns in der Pause auf dem Hof ...	Can we ... on the yard during the break. (meet)
49	Üben	x				Ohne ... wird keiner ein Meister	Without ..., no one will be a master. (practice)
50	Überschwemmung	x	x			Wenn ein Fluss über die Ufer tritt, gibt es eine...	When a river bursts its banks, there is a ... (flood)
51	verspritzen	x	x			Beim Patronenwechsel bin ich vorsichtig, ich will die Tinte nicht...	I'm careful when changing cartridges; I don't want to ... the ink. (splash)
52	verwöhnt	x		x		Meine Schwester ist ein Nesthäkchen, sie wird meistens...	My sister is a nestling, she is usually ... (pampered)
53	voll	x	x			Schenke mir mein Glas bitte ganz...	Give me a ... glass, please. (full)
54	Zaubertrick	x	x			Wer ein Kaninchen aus einem Hut holt, kann einen ...	When you take a rabbit out of a hat, you can do a ... (magic trick)
55	Zwerg	x				Im Märchen taucht oft ein kleiner... auf.	A little ... often appears in fairy tales. (dwarf)

Note: Test words and corresponding sentences are presented in alphabetic order. The table also specifies which German spelling rules appear in which test word (C= capitalization, D= consonant doubling, L=lengthening, R=rule words). The English translation is intended to give the interested non-German-speaking reader an impression of the used test material. However, it is not suitable to reflect the particular characteristics of the German orthographic rules for the test words in the spelling assessment.

Appendix B

In order to validate results of partial correlation analyses, cross-validation methods were applied. For this purpose, existing samples were randomly divided into two data sets (n_1 and n_2) for the bounded and unbounded NLE task.

First, associations of spelling and basic arithmetic in the overall sample were compared to the two sub samples. Partial correlations controlling for age revealed relatively high correlations between basic arithmetic skills and spelling skills in both samples assessing bounded (n=573, r=.51, p<.001) and unbounded (n=221, r = .43, p<.001) NLE performance. Cross-validated correlations were found to be similarly highly correlated (bounded NLET: n_1=287, r_1=.43, p_1<.001, n_2=286, r_2=.49, p_2<.001; unbounded NLET: n_1=111, r_1=.41, p_1<.001, n_2=110, r_2=.42, p_2<.001) and confirmed us to use overall spelling ability as a proxy to control for influences of general cognitive ability.

Table B1: Cross-validated partial correlations (one-tailed) between bounded (upper panel) and unbounded NLE task and basic arithmetic abilities

		A) Bounded NLE task				B) Unbounded NLE task			
Grade	Sample	**Add**	**Sub**	**Mul**	**Div**	**Add**	**Sub**	**Mul**	**Div**
5th grade	Overall	-0.09	-0.20**	-0.08	-0.03	0.06	-0.14	0.03	-0.06
	Sub 1	-0.14	-0.31**	-0.11	-0.10	0.26	0.03	0.16	0.12
	Sub 2	-0.02	-0.16*	-0.06	0.04	0.18	0.12	0.33	0.26
6th grade	Overall	-0.29**	-0.23**	-0.21**	-0.06	-0.08	-0.05	0.02	0.06
	Sub 1	-0.29**	-0.23**	-0.21*	-0.06	0.03	0.14	0.08	-0.08
	Sub 2	-0.35**	-0.29**	-0.29*	-0.12	-0.08	-0.01	-0.02	0.09
7th grade	Overall	-0.43**	-0.44**	-0.24**	-0.25*	0.03	-0.17	0.14	0.05
	Sub 1	-0.52**	-0.39**	-0.26*	-0.33**	0.40	0.10	0.08	0.25
	Sub 2	-0.42**	-0.44**	-0.31*	-0.28**	0.19	-0.32	0.24	0.09

Note: Overall and cross-validated (Sub 1 and Sub 2) partial correlations were calculated for all grade levels. Add=Addition, Sub=Subtraction, Mul=Multiplication, Div=Division. * p<.05, ** p<.001.

Second, partial correlations controlling for age and spelling skills were performed separately for all grade levels. Table A2 in Appendix A2 shows the results of partial correlations between number line estimation and basic arithmetic for the overall and the sub samples. Subsequently, correlation coefficients in the two random sub-samples (n_1 and n_2) were compared to the correlation coefficient of the overall sample (n) applying Student's t-distribution (Lenhard & Lenhard, 2014). Comparisons revealed no significant differences in both bounded (for all comparisons, z> -.925, p> .10) and unbounded number line estimation (for all comparisons, z< 1.635, p> .05). This finding could prove the correlation coefficients in the overall sample as reliable.

Section 3:

Writing and beyond: Assessment and training of orthography and arithmetic

Study 5:

Die TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE): Individualisierte Förderung durch adaptive Lernspiele

Stefanie Jung, Stefan Huber, Juergen Heller, Torsten Grust,
Korbinian Moeller, Hans-Christoph Nuerk,

Abstract (English):

Digital media have conquered children's rooms in recent years. Currently, they have been integral part of modern school education. The use of online learning environments and games, both within and outside of educational contexts, enables traditional learning content to be conveyed in a playful way, independent of time or place.

The "TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)" (TUebinger Learning Platform for the Acquisition of Numerical and Orthographic Competencies) presented below offers individual learning of central cultural techniques outside of the classroom. Adaptive methods were implemented to adapt the arithmetic and orthography games of the TULPE to individual learning needs: First, an adaptive diagnosis is used to assess the educational demand. Second, according to this assessment, learning games might be individually or game partners with similar performance might be selected. In this way, effective learning will be enabled which meets the needs of our children in their digitized life reality.

Zusammenfassung:

Digitale Medien haben nicht nur die Kinderzimmer erobert, sondern sind mittlerweile fester Bestandteil einer modernen Schulbildung. Der Einsatz von Online-Lernumgebungen und -spielen, in und außerhalb von pädagogischen Kontexten, erlaubt es selbst traditionelle Lerninhalte spielerisch und unabhängig von Ort und Zeit zu vermitteln.

Die im Folgenden vorgestellte „TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)" bietet individuelles Lernen zentraler Kulturtechniken auch außerhalb des Klassenraums. Mithilfe adaptiver Verfahren lassen sich die Rechen- und Rechtschreibspiele der TULPE an individuelle Lernermerkmale anpassen: Zuerst wird in einer adaptiven Diagnostik der Lern- und Förderbedarf eingeschätzt. Entsprechend dieser Einschätzung können in einem zweiten Schritt Lernspiele adaptiert bzw. Lern- und Spielpartner mit ähnlichem Leistungsniveau ausgewählt werden. Auf diese Weise soll effektives Lernen ermöglicht werden, das den Bedürfnissen unserer Kinder in ihrer digitalisierten Lebenswirklichkeit entspricht.

1. Introduction (Einleitung)

1.1. Digitale Medien – Realität und Lebenswirklichkeit

Digitale Medien gehören heute fest zur Lebenswelt von Kindern, Jugendlichen und jungen Erwachsenen, sie prägen ihre tägliche Freizeitgestaltung und sind Bestandteil ihrer sozialen Interaktion und Kommunikation. Der kompetente Umgang mit diesen Medien ist eine wesentliche Schlüsselqualifikation für den Wissenserwerb, weit über die schulische Bildung hinaus. Damit avanciert Medienkompetenz in der aktuellen bildungspolitischen Diskussion - neben Rechtschreiben, Lesen und numerisch-mathematischen Fähigkeiten - zu einer zentralen Kulturtechnik in unserer Informationsgesellschaft (Bundesministerium für Bildung und Forschung, BMBF, 2014).

Doch kaum ein anderes Thema wird in der Bildungsforschung so intensiv und kontrovers diskutiert wie die Potentiale und Risiken der Nutzung digitaler Medien – d.h. sowohl kommerzieller Unterhaltungsspiele als auch digitaler Lernspiele, die oft mit den englischsprachigen Begriffen ‚Educational Games' oder ‚Serious Games' beschrieben werden - in und außerhalb von Schulen (und anderen Bildungseinrichtungen). Es werden mögliche positive Effekte (z.B. verbesserte Aufmerksamkeit, gesteigerte Motivation, Wissenszuwachs, etc.) negativen Auswirkungen (z.B. Aggressivität, körperlicher Erregtheit, Suchtpotential, soziale Isolation, etc.) gegenübergestellt (für eine weiterführende Übersicht siehe u.a. Appel & Schreiner, 2014; Connolly, Boyle, MacArthur, Hainey & Boyle, 2012). Darüber hinaus steht auch die Ergänzung formeller (schulischer) Lernkontexte durch informelle (außerschulischer) Bildungsangebote in Form von digitalen Lernarrangements im Fokus der Diskussion. Insgesamt sind die Einstellungen von Schüler- und LehrerInnen zur Nutzung digitaler Medien in der Schule vergleichsweise positiv (Herzig & Grafe, 2010). Dennoch ist laut PISA Studien die Diskrepanz zwischen häuslicher und schulischer Nutzung digitaler Medien in Deutschland im Vergleich zu den anderen OECD-Staaten sehr groß. Die Verknüpfung von formellen und informellen Lernkontexten stellt hierzulande somit für die Institution Schule im Speziellen aber auch für die Bildungsforschung allgemein eine besondere Herausforderung dar.

1.1.1 Lernplattformen

Lernen und Lehren erfordern vielseitige Aufgaben und Prozesse. Zur Unterstützung von Lern- und Lehrprozessen gewinnen computerunterstützte Lernplattformen (= web-basierte oder virtuelle Lernumgebungen) auch im schulischen Kontext zunehmend an Bedeutung.

Der Begriff Lernplattform umfasst eine Vielzahl unterschiedlicher Software-Systeme, die Lern- und Lehrprozesse organisieren. Je nach System und Umsetzung

i) stellen sie Lerninhalte zur Nutzung bereit;
ii) bieten administrative Werkzeuge (z.B. Benutzer- und Kursverwaltung, Methoden zur Evaluation) für Lehrende an und
iii) fördern die Kommunikation zwischen Lernenden untereinander und mit Lehrenden, z.B. im Chat oder in Foren (vgl. Bäumer, Malys & Wosko, 2004).

Lernplattformen fungieren damit als flexible, orts- und zeitunabhängige Schnittstelle zwischen Lernern und Lehrenden. Zur Anwendung von Lernplattformen in der Schule gibt es hierzulande bislang nur wenige empirische Untersuchungen. Aus Vergleichen mit anderen europäischen Ländern lässt sich jedoch ableiten, dass Lernplattformen vor allem im Sekundarbereich eingesetzt werden, während ihr Einsatz in der Primarstufe eher selten ist (Petko, 2010). Zudem werden in Abhängigkeit von Schulform und Unterrichtsfach unterschiedliche Funktionen angeboten. Für den flexiblen Einsatz digitaler Medien sind geeignete technische und organisatorische Rahmenbedingungen erforderlich (Friedrich, Hron & Töpper, 2011), um online individuelles, didaktisch aufbereitetes und pädagogisch begleitetes Lernen realisieren zu können.

1.1.2. Digitale Lernspiele

Spielend entdecken Kinder ihre Welt. Sie lernen spielend. Spielen ist essentiell mit Lernen verbunden und doch wird Lernen in der Schule vor allem mit voranschreitender Klassenstufe zunehmend formalisiert. Ein großer Teil der Lehrinhalte wird über Schulbücher vermittelt (Herzig und Grafe, 2007) und spielerische Elemente lassen sich kaum (noch) entdecken.

Digitale Lernspiele basieren auf der engen, scheinbar untrennbaren Beziehung von Spielen und Lernen. Als Schnittstelle zwischen beiden nutzen sie Elemente aus digitalen Spielangeboten (Computer- oder Videospiele) zur Vermittlung von Lerninhalten. Sie versuchen Lernen an die digitale Lebenswelt der Kinder und Jugendlichen anzupassen und den Bedürfnissen einer neuen Generation von Lernern gerecht zu werden. Ihre Eignung als Lernmedium steht stark im Fokus öffentlicher Diskussionen, in der verschiedene, mitunter synonym verwendete Begriffe für digitale Lernspiele bzw. den Einsatz digitaler Spielangebote im Lernkontext verwendet werden: u.a. „Serious Games", „Educational Games", „Game-Based-Learning" oder „Edutainment". Alle Konzepte haben gemein, dass sie positive Spielelemente (z.B. Motivation und Begeisterung) mit didaktischen und pädagogischen Konzepten verbinden und sie „ernsthaft" im Bildungskontext einsetzen. Die eindeutige Abgrenzung dieser Begriffe ist jedoch schwierig (vgl. Egenfeldt-Nielsen, 2007). Eine mögliche Kategorisierung orientiert sich am Verhältnis von didaktischen zu spielerischen Elementen, wobei „Serious Games" und „Edutainment" die jeweils entgegengesetzten Pole besetzen (Wechselsberger, 2009). Während der Anteil didaktischer Elemente in „Serious Games" stark ausgeprägt ist, fokussiert „Edutainment" eher die Unterhaltungsebene.

Eine für die weitere Darstellung geeignete Definition digitaler Lernspiele findet sich bei Hawlitschek (2013, p. 23). Sie versteht darunter „*Computerspiele,*

- *die explizit und systematisch in Hinblick auf ein bestimmtes Lernziel und für den Einsatz in einem pädagogischen Kontext konzipiert wurden,*
- *die ein positives Spielerleben beim Spieler auslösen*
- *deren Effektivität bei der Vermittlung der Lerninhalte nachgewiesen werden konnte."*

Dass digitale Lernspiele vielseitig zur Vermittlung unterschiedlicher Lerninhalte eingesetzt werden können, ist weitgehend unstrittig. Die Frage nach ihrer Effektivität kann jedoch nicht umfassend beantwortet werden: Eine Reihe von Studien konnten positive Lerneffekte belegen, u.a. bei beim Mathe- und Sprachenlernen (Li & Ma, 2010; Torgerson & Zhu, 2003) sowie beim Training kognitiver Fähigkeiten (Ke & Graboski, 2007; Warren, Dondlinger & Barab, 2008), wobei sich herausstellt, dass die Effektivität von zahlreichen Faktoren abhängt. Nach einer aktuellen Metaanalyse von Wouters, Nimwegen, van Oostendorf und van der Spek (2013) sind Lerneffekte stärker, wenn digitale Lernspiele in der Gruppe gespielt werden. Darüber hinaus spielen u.a. spielerische Gestaltung, Verständlichkeit der Instruktionen und die Einbettung in einen sorgfältig durchdachten didaktischen Rahmen ebenso eine entscheidende Rolle (vgl. Ke, 2009). Jedoch zeigt sich auch, dass digitale Lernspiele per se nicht motivierender sind als klassische Lernarrangements (Wouters et al., 2013). Als Gründe dafür führen die Autoren an, dass i) Vorgaben, wie wann und auf welche Weise digitale Lernspiele gespielt werden, dem Lerner die eigene Kontrolle und damit auch die Motivation zum Spielen nehmen können, ii) der *flow* des Spiels durch eine Fokussierung auf zu lernende Inhalte unterbrochen werden kann und die Identifikation mit dem Spiel erschwert ist; iii) möglicherweise auch die Verwendung von Fragebögen zur Erfassung von Spielmotivation nicht geeignet sei. Darüber hinaus kann eine Vielzahl von Lernspielen nicht mit den grafisch-technischen Möglichkeiten kommerzieller Computerspielentwicklungen konkurrieren. Die gestalterischen Mittel sind meist nicht völlig ausgeschöpft, Spielehandlungen und Aufgaben wenig vielseitig und komplex (Petko, 2008). Diese Faktoren können die Glaubwürdigkeit des Spiels und die Entstehung eines Präsenzgefühls (engl. *presence*) beeinflussen. Aus Spieler-(Lerner)-Perspektive werden digitale Lernspiele daher den Erwartungen, die an sie geknüpft sind, nicht immer gerecht. Aus Lehrerperspektive besteht das Risiko, dass Kosten und pädagogischer Nutzen digitaler Lernspiele in einem Missverhältnis stehen. Eine weitere, häufig genannte Einschränkung sind die fehlenden individuellen Differenzierungsmöglichkeiten digitaler Lernspiele: Während ein Spiel einige Schüler überfordert, kann dasselbe Spiel andere langweilen. Den Lernspielen fehlt es bisweilen an erkennbarer Adaptivität (Petko, 2008). Vor allem Schüler mit einem geringeren Leistungsniveau erleben das Spielen gegen bessere Spielpartner dann oft als frustrierend (Liu, Li, & Santhanam, 2013). Dieses Erleben geht mit einem Motivationsverlust einher, der sowohl Nutzungsdauer und -häufigkeit beeinflussen und sich zudem auf Aufmerksamkeitsprozesse und die Qualität der Verarbeitung kognitiver Lerninhalte auswirken kann (z.B. Pekrun, 1992). Eine Möglichkeit dem entgegenzuwirken und das motivierende Element digitaler Lernspiele zu erhalten, bieten adaptive Verfahren.

1.2. Adaptivität - Der Individualität gerecht werden

Zentrales Ziel und zugleich hoher Anspruch aller Lehr- und Lernsituationen ist es, Lernende entsprechend ihres aktuellen Leistungsstandes zu fördern, zu fordern und Lernangebote an die unterschiedlichen Gegebenheiten und Bedürfnisse zu orientieren. Gerade computerbasierte Lernumgebungen, wie Lernplattformen und Lernspiele, bieten die technischen Voraussetzungen, um mithilfe von Adaptivität das „Verhalten“ der Lernumgebung an die Merkmale des Lerners anzupassen. Adaptive Lernsysteme werden bereits jetzt als eine der nächsten Generationen web-basierter Lernumgebungen diskutiert (z.B. Verdú, Regueras, Verdú, De Castro & Pérez, 2008). Insbesondere da sie es u.a. erlauben, Lernmaterialien automatisiert je nach individuell benötigter Lern- und Verarbeitungszeit in unterschiedlicher Dauer, Komplexität und Struktur anzubieten und dabei individuelles Vorwissen und Lernstrategien zu berücksichtigen. Damit vereinen sie diagnostische und lernförderliche Elemente. Diese grobe Zweiteilung findet sich auch bei der Anwendung von Adaptivität in computerbasierten Lernumgebungen:

1.2.1. Diagnostik: Adaptive Erfassung der Lernerkompetenzen

In einer adaptiven Eingangsdiagnostik erfolgt eine erste Einstufung des Lerners, meist in Form eines Wissenstests. Zusätzlich zum getesteten (domänenspezifischen) Wissen, können auch weitere Merkmale, z.B. individuelles Vorwissen, Alter und Geschlecht sowie Lernermerkmale und -gewohnheiten (vgl. Kickmeier-Rust, Albert & Roth, 2007; Oblinger, 2004) erfasst werden. Im Gegensatz zu klassischen Tests, in denen eine bestimmte Anzahl an Aufgaben in immer derselben Reihenfolge vorgegeben wird, orientiert sich die Aufgabenauswahl und -reihenfolge beim adaptiven Testen an dem individuell gezeigten Antwortverhalten: Wird eine Aufgabe „falsch“ beantwortet, erhält der Lerner als nächstes eine einfachere Aufgabe. Ist die Lösung „richtig“, folgt eine schwierigere Aufgabe. Durch dieses sukzessive Vorgehen wird der Lerner nicht über- und nicht unterfordert. Darüber hinaus kann die Testdauer erheblich reduziert werden, ohne dass Messpräzision und -information verloren gehen (für eine detaillierte Einführung: van der Linden & Glas, 2000).

1.2.2. Lernförderung: Adaption des Lernangebots

Auf Grundlage der Einstufung des Lerners wird in einem zweiten Schritt das Lernangebot bzw. die Lernumgebung selbst angepasst. Dies kann sehr unterschiedlich umgesetzt werden: Die Anpassung des Lernangebots auf individuelle Lernverläufe wird im Folgenden als *intra-individuelle* Adaptivität bezeichnet. Sie umfasst vorwiegend die Individualisierung hinsichtlich Schwierigkeit und Komplexität, kann jedoch auch individuelles Vorwissen, Lern- und Lösungsstrategien berücksichtigen. *Inter-individuelle* Adaptivität hingegen meint die Auswahl besonders geeigneter, d.h. im Allgemeinen ähnlich leistungsstarker Lern- oder Spielpartner (bzw. Spielgegner). Neben der adaptiven Gestaltung des Lernangebots wird in dieser Phase auch das individuelle Lernverhalten durch den Computer dokumentiert. So können u.a. die Entwicklung von Lernprozessen und -strategien, Spielstände aber auch Fehleranzahl und -typen erfasst werden. (u.a. Kalyuga, 2006).

Adaptivität kann in computerbasierten Lernumgebungen sehr unterschiedlich realisiert werden. Eine Vielzahl aktueller computerisierter adaptiver Testverfahren beruht auf mathematischen Modellen der Item-Response-Theorien (IRT, auch probabilistische Testtheorie), denen eine hohe Testökonomie (Effizienz) und Messgenauigkeit zugeschrieben wird (u.a. Kubinger, 1993). In diesen Modellen wird versucht das wahrscheinliche Verhalten eines Lerners (bzw. einer Testperson), dem bestimmte (latente) Fähigkeiten zugrunde liegen, vorherzusagen, wenn ihm eine bestimmte Aufgabe, deren Eigenschaften (wie z.B. Schwierigkeit) man berechnen kann, gestellt wird. Somit wird eine, an der Fähigkeit des Lerners orientierte, selektive Vorgabe einzelner Aufgaben ermöglicht.

2. Methods (Methoden)

2.1. Digitale Medien und die Förderung zentraler Kulturtechniken

Rechnen, Lesen und Rechtschreiben gehören zu den Schlüsselqualifikationen unserer Informationsgesellschaft. Sie sind wesentlicher Bestandteil und zugleich Grundlage von Schul- und beruflicher Bildung. Störungen im Erwerb dieser Fähigkeiten resultieren in Defiziten, die ohne geeignete Intervention bis ins Erwachsenenalter bestehen können (z.B. Parsons & Bynner, 2005; Daniel, Walsh, Goldston, Arnold, Reboussin & Wood, 2006).

Klassische Interventionsprogramme zum Rechtschreiben und Rechnen, die häufig schulischem Lernen ähneln, können Kinder in ihrer digitalen Lebenswelt heute meist nur schwer erreichen. Daher wurde in den letzten Jahren verstärkt nach neuen spielerischen Ansätzen zur Förderung schriftsprachlicher und numerisch-mathematischer Fähigkeiten gesucht. Aktuell existieren verschiedene digitale Lernangebote und -spiele, die übergreifend (z.B. Lernwerkstatt) oder relativ spezifisch klassische Bildungsangebote zu Lesen/Rechtschreiben (z.B. Dybuster) und Rechnen (z.B. The Number Race) ergänzen. Während viele dieser Lernumgebungen einen hohen Lernerfolg versprechen und z.T. mit werbewirksamen Preisen ausgezeichnet wurden (z.B. Meister Cody – Talasia, Kaasa health, Version 1.0.6, 2014), sind nur einige wissenschaftlich evaluiert worden(für Rechtschreiben: z.B. Kast, Meyer, Voegeli, Gross & Jaencke, 2011; Rechnen: z.B. Wilson, Dehaene, Pinel, Revkin, Cohen & Cohen, 2006; Butterworth & Laurillard, 2010). Unbestritten ist jedoch das Potential web-basierter Lernumgebungen attraktive Lernangebote zu gestalten, die mithilfe des Internets unabhängig von Ort und Zeit genutzt werden und sich mithilfe adaptiver Verfahren an die individuellen Voraussetzungen und Bedürfnisse der Lerner anpassen können.

1.3.1. TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)

Die „TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen“ bietet individuelles Lernen unabhängig von formellen Bildungsangeboten. Um dies zu gewährleisten, wurde sie so entwickelt, dass ein Zugriff (auf die Lernplattform auf http://lernplattform.iwm-kmrc.de/) mit verschiedenen Endgeräten (Computer, Tablet-PC und/oder Smartphone) möglich und das Lernen somit zeit-, orts- und situationsunabhängig ist. Ein wesentliches Unterscheidungsmerkmal und zugleich großer Vorteil gegenüber anderen Lernplattformen, ist die Anwendung adaptiver Verfahren zur individuellen Anpassung des Lernangebots: Auf intra-individueller Ebene werden Lernspiele entsprechend des in einer ersten Diagnostik eingeschätzten Lern- und Förderbedarfs angepasst. Auf inter-individueller Ebene werden Lern- und Spielpartner entsprechend ihres Leistungsniveaus ausgewählt, so dass eine möglichst ausgeglichene Paarung entsteht.

Die „TUebinger LernPlattform“ vereint (fast) alle Merkmale klassischer Lernplattformen (vgl. Bäumer et al., 2004). Sie

i) stellt eine Auswahl an Lernspielen bereit;
ii) bietet eine Benutzerverwaltung zum Anlegen von Benutzerprofilen, auf denen persönliche Informationen (Alter, Geschlecht, Lieblingsfach, Hobbies, etc.) der Nutzer hinterlegt werden können aber nicht müssen. Darüber hinaus können hier Lernstatus und –verlauf dokumentiert und ausgewertet werden;
iii) ermöglicht die Kommunikation der Spieler untereinander (auch während der Spiele) sowie den Kontakt mit Entwicklern und Betreuern (Psychologen) der Lernplattform, wenn nötig.

und ergänzt diese Elemente mit intra- und interindividueller Adaptivität für ein hohes Maß an Individualisierung von Spielen und Lernen.

1.3.2. Lernspiele der TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen (TULPE)

Derzeit bietet die TULPE eine Auswahl von vier Rechtschreib- und Rechenspielen. Alle Lernspiele können im Einzel- oder Mehrspieler-Modus für bis zu fünf Mitspieler gespielt werden. Gerade das Spielen, Interagieren und Kommunizieren mit mehreren Mitspielern gleichzeitig ist eine Besonderheit dieser Lernspiele und Grundlage gemeinsamen, interaktiven Lernens.

Rechtschreibspiele

Der Entwicklung der Rechtschreibspiele ging eine detaillierte Analyse der linguistischen Eigenschaften, auf denen die verschiedenen Rechtschreibregeln beruhen, voraus. Dazu wurden in einem ersten diagnostischen Schritt relevante Parameter identifiziert, u.a. phonologische Bewusstheit (z.B. Längenunterscheidung bei Lauten, Erkennen von Stimmhaftig- und Stimmlosigkeit, etc.) sowie morphologische Aspekte der Wortbildung. Diese Parameter wurden mit Modellen der Item-Response-Theorie (IRT) analysiert. IRT Modelle ermöglichen es, die Schwierigkeiten von Testitems und die Rechtschreibkompetenz der Schüler und Schülerinnen getrennt voneinander zu betrachten. Somit kann der individuelle Leistungsstand berücksichtigt und bestehende Defizite mit einem adäquat an den Lernstand angepassten Schwierigkeitsniveau verbessert werden.

1. *Doppelungsspiel:* In der deutschen Sprache folgt - mit wenigen Ausnahmen (z.B. Bus) - auf einen kurzen Vokal ein Doppelkonsonant (z.B. Mann). Das *Doppelungsspiel* trainiert die Diskrimination von Vokallängen. Sie stellt eine Grundvoraussetzung für die korrekte Realisierung von Doppelkonsonanten dar (Landerl, 2005; Ise & Schulte-Körne, 2010). Das Spiel besteht aus zwei Phasen: In der ersten Phase - Identifikation - geht es um die Vokallängenidentifikation bei einem auditiv vorgegeben Wort (z.B. Tasse). An diese Phase schließt sich ein sogenanntes „Minispiel“ an, in dem herabfallende Münzen eingesammelt werden sollen, wofür Punkte vergeben werden. Auf den Münzen sind verschiedene Doppelkonsonanten abgebildet. Die Spieler dürfen nur die Münzen einsammeln, die den identifizierten Doppelkonsonanten des eingangs präsentierten Zielworts tragen. (z.B. ‚ss‘). Im Anschluss an das Minispiel folgt die zweite Phase des *Doppelungsspiels* - Produktion. Hier soll das Zielwort korrekt eingetippt werden. In beiden Phasen erhalten die Spieler kindgerechtes Feedback zu ihrem Antwortverhalten. Dazu werden fröhliche oder traurige Gesichter (Smileys) gezeigt, je nachdem, ob die Aufgabe richtig oder falsch gelöst wurde. Abbildung 1A zeigt beispielhaft den Ablauf des *Doppelungsspiels.*

2. *LeTris:* Das Spiel “LeTris” ähnelt dem Computerspieleklassiker “Tetris”. Indem Wörter nach auditiver Vorgabe (z.B.: KETTE) aus zufällig herunterfallenden Buchstaben (z.B.: L, T, E, K, etc.) zusammengesetzt werden, trainiert es sowohl Wortsynthesefähigkeiten als auch Aufmerksamkeitsprozesse. In Analogie zu Tretis wird eine Zeile dann gelöscht, wenn das Zielwort korrekt zusammengesetzt wurde; andernfalls bleibt die Zeile

bestehen. Je mehr Wörter fehlerhaft sind, desto mehr Zeilen verbleiben auf dem Bildschirm.

3. *Wortfamilien:* Wörter werden einer Wortfamilie (z.B. spielen, Spiel, verspielt, spielerisch, etc.) zugeordnet, wenn sie das gleiche Stammmorphem (z.B. -spiel) besitzen. Das Erkennen dieser morphologischen Beziehungen ist ein weiterer Grundbaustein für die Entwicklung ungestörter Lese- und Rechtschreibfähigkeiten (Carlisle & Katz, 2006). Das Spiel *Wortfamilien* nimmt sich dieser Grundfertigkeit an: Zu Beginn wird ein Wort visuell präsentiert (z.B. passen). Dann wird eine Auswahl weiterer Wörter gezeigt (Schwimmflosse, Pass, Eis, verpassen, etc.). Die Spieler sollen sich möglichst alle Wörter zu der präsentierten Wortfamilie merken und diese korrekt eingeben. Sie erhalten Rückmeldung darüber, wie viele und welche Wörter sie richtig erkannt und geschrieben haben.

4. *Wortbausteine:* Dieses Spiel basiert ebenfalls auf dem Konzept der Wortfamilien: Visuell präsentierte Wörter (z.B. sperrig, versperrt, Läufer, verlaufen, Lesebuch, lesbar, etc.) sollen verschiedenen Stammmorphemen zugeordnet werden (z.B. sperr-, lauf-, les-). Beide Spiele *Wortfamilien* und *Wortbausteine* trainieren das Bewusstsein für morphologische Konsistenz, ein entscheidendes Konzept in der Orthographie. In einigen Untersuchungen konnte gezeigt werden, dass ein derartiges Training die Entwicklung korrekter Rechtschreibfähigkeiten fördert (Ise & Schulte-Körne, 2010).

Rechenspiele

Die Spiele zur Verbesserung der Rechenleistungen trainieren die klassischen Grundrechenarten sowie die Zuordnung von numerisch-räumlichen Relationen. Fast alle Spiele sind als sogenannte Wahl-Reaktionsaufgaben implementiert. Bei diesen Aufgaben werden zwei oder mehr mögliche Lösungen zu einem Problem präsentiert. Der Lerner muss das richtige Ergebnis aus einer Auswahl falscher Ergebnisse herausfinden.

1. *Multiplikationsspiel:* Zur Übung von Multiplikationsfakten, dem sog. „Einmaleins", wird zuerst das Ergebnis einer Multiplikationsaufgabe präsentiert (z.B. 24). Dann werden Multiplikationsaufgaben, die zum richtigen Ergebnis führen (z.B. 3x8 oder 4x6), und Distraktoren gezeigt (z.B. 4x8 oder 6x6). Es sollen die Aufgaben mit der Maus markiert werden, die das gesuchte Produkt ergeben. Beim *inversen Multiplikationsspiel* wird die Aufgabe präsentiert und das korrekte Ergebnis soll ausgewählt werden

2. *Partnerzahl:* Dieses Additionsspiel übt das Platz x Wert oder Stellenwertsystem arabischer Zahlen. *Partnerzahl* trainiert die Addition im Zahlenraum bis 10. Den Spielern wird eine Zahl präsentiert (z.B. 4) und sie sollen aus einer Menge von Lösungsmöglichkeiten die Zahl auswählen, die addiert werden muss, damit sich die Summe 10 ergibt (z.B. 6).

3. *Über 10: Über 10* ist eine Erweiterung des *Partnerzahl*-Spiels. Hier wird die Addition mit Übertrag trainiert. Zuerst markieren die Spieler den Summanden, der sich mit der vorgegebenen Zahl (z.B. 8 + 2) zu 10 addiert. Im Anschluss soll die Zahl ausgewählt werden (z.B. 7) die zur richtigen Lösung der eingangs präsentierten Aufgabe (z.B. 8 + 9) führt.

4. *NumberLine:* Trainiert werden räumliche Repräsentationen von Zahlen. In dieser Aufgabe sollen zwei oder mehr Spieler die ungefähre Position einer vorgegebenen Zahl auf einem Zahlenstrahl markieren. Sobald ein Spieler seine Markierung gesetzt hat, kann kein weiterer Spieler diese Position als richtige Lösung angeben. Weder sehr genaue und langsame Schätzungen noch schnelle und inakkurate Strategien führen hier zum Ziel. Zahlenstrahl-Spiele wurden bereits erfolgreich zum Training von Zahl- und Raumrelationen und Vermittlung entsprechender Lösungsstrategien eingesetzt (Link, Schwarz, Huber, Fischer, Nuerk, Cress, & Moeller, 2014). Abbildung 1B zeigt ein Beispiel des Zahlenstrahlspiels.

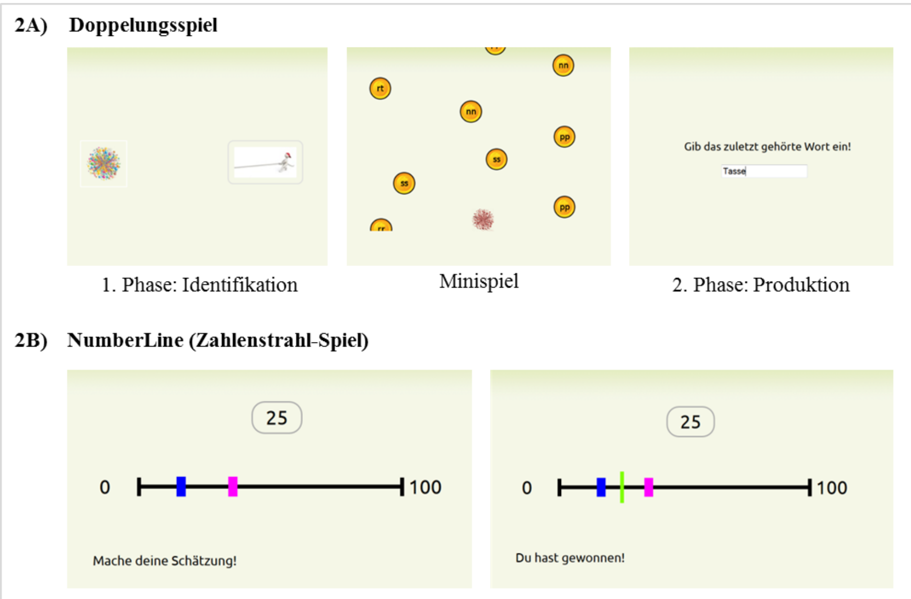


Fig. 1: (Abb.1): (A) Doppelungsspiel zum Training phonologischer Bewusstheit; **(B)** NumberLine (Zahlenstrahl-) Spiel zum Training numerischer und räumlicher Relationen

Die Rechtschreib- und Rechenspiele der TULPE entsprechen in vollen Umfang der obigen Definition digitaler Lernspiele (Hawlitschek, 2013, p.23):

i) Sie wurden explizit zur Förderung arithmetischer und schriftsprachlicher Fähigkeiten entwickelt;

ii) Die Spiele wurden von der entsprechenden Zielgruppe (fünfte und sechste Klasse) getestet, fanden positive Resonanz und wurden von einem großen Teil der Schülerinnen und Schüler mit Freude und Motivation gespielt.

iii) Ihre Effektivität bei der Vermittlung von Rechen- und Rechtschreibkompetenzen wurde wissenschaftlich evaluiert und nachgewiesen (Jung, Rösch, Huber, Heller, Grust, Nürk & Möller, 2015). Für die Rechtschreibspiele *Doppelungsspiel, Wortfamilien* und *Wortbausteine* fanden sich differentielle Interventionseffekte. Direkte Trainingseffekte konnten für das *Multiplikationsspiel, Partnerzahl* und *Über 10* nachgewiesen werden. Transfereffekte auf die Gegenoperationen blieben jedoch aus. Im Rahmen dieser Evaluation wurden Spielpartner entsprechend ihrer Rechen- und Rechtschreibfertigkeiten ausgewählt, so dass zwischen ihnen keine großen Leistungsunterschiede bestanden. Dieses Vorgehen erwies sich als sehr effektiv - auch wenn es in der Vorbereitung einige Zeit in Anspruch nahm und einer sehr genauen Einschätzung durch den Klassenlehrer voraussetzte – und bestärkt uns in der Entwicklung und Bereitstellung eines individuellen und adaptiven Lernangebots.

2.2 Individualisiertes, adaptives Lernen mit der TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen

Adaptive Lernspiele können unterschiedliche individuelle Fähigkeiten und Lernverhalten erkennen, darauf reagieren und das Spielerlebnis - stets mit Blick auf den Lernerfolg - selbständig differenziert gestalten. Auf der TULPE wird die Adaptivität mit Methoden der Item-Response-Theorie umgesetzt. Für numerische und schriftsprachliche Aufgaben wurden dazu spezifische Aufgabenschwierigkeiten geschätzt und zugeordnet. Diese Parameter werden für die automatisierte selektive Vorgabe von Aufgaben entsprechend des individuellen Kompetenzniveaus des Lerners benötigt. Die Umsetzung der Adaptivität erfolgt auf der TUebinger LernPlattform gemäß der oben beschriebenen Unterscheidung zwischen Diagnostik und Lernförderung.

2.2.1. Adaptive Diagnostik der Rechtschreib- und Rechenkompetenz

Zur ersten Einschätzung der Rechen- und Rechtschreibfähigkeiten und als Grundlage für individualisiertes Lernen wurde eine *Initiale Diagnostik* entwickelt. Sie ist den Lernspielen vorangeschaltet. Je nachdem, für welches Spiel sich ein Lerner entscheidet, wird mithilfe eines kurzen Tests der aktuelle Leistungsstand beim Rechtschreiben oder Rechnen erfasst. Im Bereich numerischer Fähigkeiten wird z.B. das Rechnen in den Grundrechenarten überprüft. Die Rechtschreibfähigkeiten werden hinsichtlich der verschiedenen Rechtschreibregeln (z.B. Groß-Klein-Schreibung, Doppelung, Dehnung, Auslautverhärtung, etc.) untersucht. Entscheidend dabei ist, dass bereits nach der Eingabe eines jeden Wortes automatisch bewertet wird, ob es entsprechend jeder einzelnen Rechtschreibregel richtig oder falsch geschrieben wurde. Dies ist notwendig, um für die nächste Aufgabe ein leichteres oder schwierigeres Wort auszuwählen und vorzugeben. Abbildung 2 zeigt den schematischen Ablauf der *Initialen Diagnostik* am Beispiel Rechtschreibung (A) und auch die Veränderung der geschätzten Fähigkeiten eines Lerners im Verlauf der Diagnostik für die drei Rechtschreibregeln: Groß-Klein-Schreibung, Doppelung und Dehnung (B).

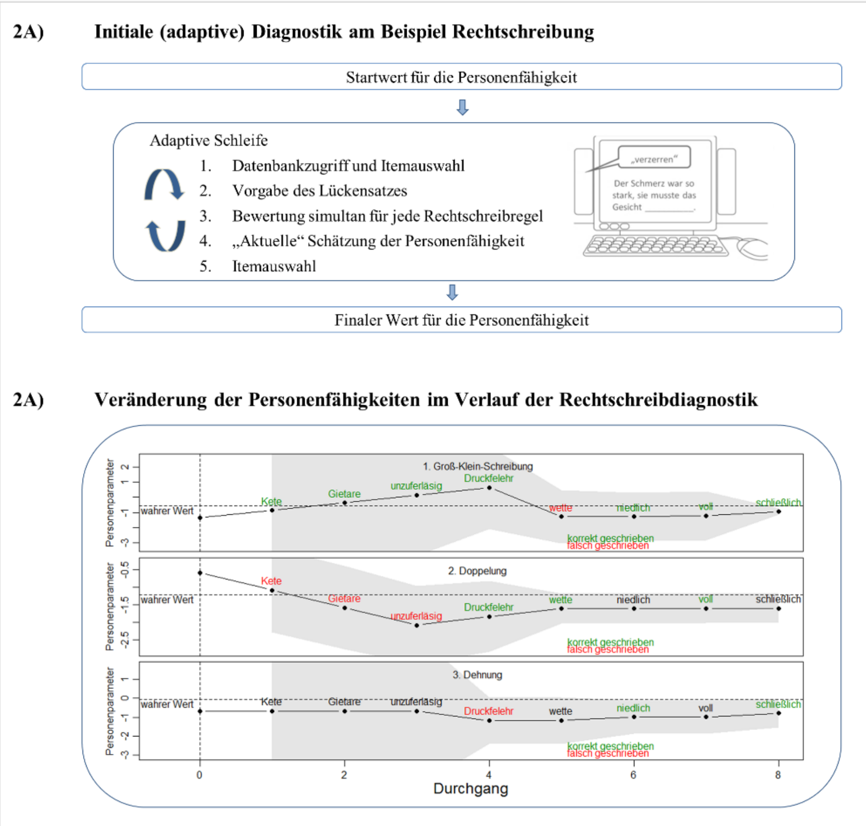


Fig. 2: (Abb. 2.): (A) A: Schematischer Ablauf der Initialen Diagnostik am Beispiel Rechtschreibung; **(B)**Veränderung der Personenfähigkeiten im Verlauf der Rechtschreibdiagnostik für die Groß-Klein-Schreibung, Dehnung und Doppelung: grün = korrekte Schreibung entsprechend der spezifischen Rechtschreibregel, rot = falsche Schreibweise, schwarz = Regel für dieses Wort nicht relevant

Auf Basis dieser *Initialen Diagnostik* können in einem nächsten Schritt die Lernspiele entsprechend des Leistungsniveaus des Lerners angepasst und Förderschwerpunkte abgeleitet werden, so dass, z.B. ein bestimmtes Lernspiel empfohlen werden kann.

2.2.2. Adaptivität der Lernspiele

Um das Lernen mit der „TULPE effektiv zu gestalten, werden die Lernspiele intra- und interindividuell adaptiv an das schriftsprachliche und numerische Kompetenzniveau der Lerner angepasst. Auf *intra-individueller* Ebene wird der individuelle Leistungstand eines Lerners berücksichtigt. Zum einen lässt sich daran Lernerfolg und -verlauf des Lerners bewerten und beurteilen. Zum anderen kann bei Spielen mit computergesteuerten Spielpartnern, die Spielweise des Computers so angepasst werden, dass die individuelle Erfolgsrate für den Lerner stets bei über 50% liegt. Die *inter-individuelle* Adaptivität findet in Lernspielen mit mehreren Lern- und Spielpartnern Anwendung. Das Lernspiel kann entsprechend des Leistungsniveaus des aktuellen Mit- oder Gegenspieler angepasst werden oder in Relation zu allen registrierten Spielern modifiziert werden. Mithilfe gespeicherter Spieldaten lässt sich für jeden Spieler und jedes Spiel eine Art Kompetenzindex bestimmen anhand dessen gleichstarke Mit- oder Gegenspieler ausgewählt werden können (ähnlich wie der ELO-Wert im Schach, vgl. Elo, 1978). Finden sich keine gleichstarken Spielpartner, so können z.B. leistungsstärkere Spieler ein Handikap erhalten. Solche Handicaps können etwa durch Verzögerungen bei der Aufgabenpräsentation oder bei der Registrierung des Antwortverhaltens realisiert werden. Dadurch ist es möglich, Erfolgsraten auszubalancieren und somit die Motivation der Lerner unabhängig vom individuellen Fähigkeitslevel aufrecht zu erhalten.

Während die adaptive Diagnostik bereits implementiert ist, stellt die Umsetzung und Bereitstellung der adaptiven Lernspiele auf der „TUebinger LernPlattform zum Erwerb numerischer und orthografischer Kompetenzen" den nächsten Meilenstein auf dem Weg zur maßgeschneiderten und motivierenden Förderung schriftsprachlicher und numerischer Kompetenzen dar.

3. Discussion (Diskussion)

3.1 Lernen Mit Digitalen Medien - Perspektiven und Ausblick

Web-basierte Lernumgebungen - Lernplattformen und digitale Lernspiele - haben sich in den letzten Jahren zu vielversprechenden Instrumenten für den Einsatz in und außerhalb von pädagogischen Kontexten entwickelt. Selbst beim Lernen der zentralen Kulturtechniken Rechtschreiben und Rechnen, das traditionell mühsam und zeitaufwendig ist und überwiegend in der Schule stattfindet, hat sich ihr Einsatz fraglos bewährt (siehe Butterworth & Laurillard, 2010, Kast et al., 2011). Mit computerunterstützten Lernumgebungen können diese Fähigkeiten *spielerisch* gefördert und Defizite verringert werden (siehe Li & Ma, 2010; Torgerson & Zhu, 2003). Lerneffekte sind dabei oft sogar höher, wenn digitale Lernspiele in der Gruppe gespielt werden (Wouters et al., 2013). Aus diesem Grund werden Lernspiele auf der TULPE auch im Mehrspielermodus angeboten.

Der oft als motivierend angesehene Charakter von Lernspielen ist jedoch anfällig. So deutet die aktuelle Metaanalyse von Wouters und Kollegen (2013) darauf hin, dass digitale Lernspiele nicht motivierender sind als klassische Lernmethoden. Dabei scheint es entscheidend, wie autonom ein Spieler über den Nutzen digitaler Lerninhalte entscheiden kann und wie stark involviert er sich fühlt. Dennoch scheinen sich motivationale Prozesse regulieren zu lassen: Die Evaluation der Rechtschreib- und Rechenspielen der TULPE ergab, dass ein Aufeinandertreffen gleichfähiger Lern- oder Spielpartner entscheidend dafür ist, dass insbesondere leistungsschwächere Lerner weder Motivation noch Freude am Spiel verlieren. Ein solcher Verlust kann sich negativ auf den Lernerfolg auswirken (Liu et al., 2013). Hier ansetzend bietet die Implementierung von Adaptivität eine vielversprechende Methode, um Lernspiele automatisiert entsprechend des Leistungsniveaus des aktuellen Mit- oder Gegenspieler anzupassen. Darüber hinaus ermöglicht sie eine hohe Individualisierung diagnostischer Verfahren und darauf aufbauend die Auswahl spezifischer Lerninhalte. Wie in der personalisierten Medizin werden unseres Erachtens auch Lernspiele und Lernangebote stärker personalisiert und an die individuellen Bedürfnisse adaptiert werden müssen, damit möglichst effektiv gelernt werden kann.

Damit steigen die Anforderungen an die Entwicklung von Lernplattformen und –spielen, für deren Bewältigung eine enge Zusammenarbeit von Informatik, Psychologie, Pädagogik sowie weiteren beteiligen Fachdisziplinen unerlässlich ist. Gelingt es adaptive Verfahren in web-basierte Lernumgebungen zu integrieren und somit ihr didaktisches und motivationales Potential nutzbar zu machen, bieten sie ein erfolgversprechendes Lernmedium, das der digitalisierten Lebenswirklichkeit unserer Kinder entspricht.

References (Literatur)

Appel, M. & Schreiner, C. (2014). Digitale Demenz? Mythen und wissenschaftliche Befundlage zur Auswirkung von Internetnutzung. *Psychologische Rundschau, 65*, 1-10.

Bäumer, M., Malys, B. & Wosko, M. (2004). Lernplattformen für den universitären Einsatz. In K. Fellbaum & M. Göcks (Hrsg.), *eLearning an der Hochschule* (pp. 121-140), Aachen: Shaker Verlag.

Bundesministerium für Bildung und Forschung (BMBF): Digitale Medien in der beruflichen Bildung. Zugriff am 13.11.2014 unter http://www.bmbf.de/de/16684.php/

Butterworth, B. & Laurillard, D. (2010). Low numeracy and dyscalculia: Identification and intervention. *ZDM Mathematics Education, 42*(6), 527-539.

Carlisle, J. F. & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, *19*(7), 669-693.

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T. & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A. & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities, 39*, 507-514.

Egenfeldt-Nielsen, S. (2007). *Educational potential of computer games*. Continuum.

Elo, A. E. (1978). *The rating of chessplayers, past and present* (Vol. 3). London: Batsford.

Friedrich, H. F., Hron, A. & Töpper, J. (2011). Lernplattformen in der Schule. In: *Schule in der digitalen Welt* (pp. 117-141). VS Verlag für Sozialwissenschaften.

Hawlitschek, A. (2013). *Spielend Lernen. Didaktisches Design Digitaler Lernspiele Zwischen Spielmotivation und Cognitive Load* (Vol. 20). Logos Verlag Berlin GmbH.

Herzig, B. & Grafe, S. (2007). *Digitale Medien in der Schule. Standortbestimmung und Handlungsempfehlungen für die Zukunft. Studie zur Nutzung digitaler Medien in allgemein bildenden Schulen.* Dt. Telekom, Zentralbereich Unternehmenskommunikation.

Ise, E. & Schulte-Körne, G. (2010). Spelling deficits in dyslexia: evaluation of an orthographic spelling training. *Annals of Dyslexia*, *60*(1), 18-39.

Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied cognitive psychology*, *20*(3), 333-342.

Kast, M., Baschera, G. M., Gross, M., Jäncke, L. & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of dyslexia*, *61*(2), 177-200.

Ke, F. & Grabowski, B. (2007). Gameplaying for maths learning: cooperative or not?. *British Journal of Educational Technology*, 38(2), 249-259.

Ke, F. (2009). A qualitative meta-analysis of computer games as learning tools. *Handbook of research on effective electronic gaming in education*, *1*, 1-32.

Kickmeier-Rust, M. D., Albert, D. & Roth, R. (2007). *A Methodological Approach to Address Individual Factors and Gender Differences in Adaptive E-Learning.* na.

Kubinger, K. D. (1993). Testtheoretische Probleme der Computerdiagnostik. *Zeitschrift für Arbeits- und Organisationspsychologie, 37*(3), 130-137.

Landerl, K. (2005). Reading acquisition in different orthographies: Evidence from direct comparisons (pp. 513-530). In M. Joshi, & P. G. Aaron (Eds.), *Handbook of orthography and literacy*. Hove: Erlbaum.

Li, Q. & Ma, X. (2010). A Meta-analysis of the Effects of Computer Technology on School Students' Mathematics Learning. *Educational Psychology Review*, *22*(3), 215-243.

Link, T., Schwarz, E. J., Huber, S., Fischer, U., Nuerk, H.-C., Cress, U. & Moeller, K. (2014). Mathe mit der Matte – Verkörperlichtes Training basisnumerischer Kompetenzen. Zeitschrift für Erziehungswissenschaft, *17*(2), 257-277.

Liu, D., Li, X. & Santhanam, R. (2013). Digital games and beyond: What happens when players compete? *Management Information Systems Quarterly, 37*(1), 111-124.

Oblinger, D. G. (2004). The next generation of educational engagement. *Journal of interactive media in education*, 2004(1), Art. - 10.

Parsons, S. & Bynner, J. (2005). *Does numeracy matter more?*. London: National Research and Development Centre for adult literacy and numeracy.

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology, 41*(4), 359-376.

Petko, D. (2010). *Lernplattformen in Schulen.* Springer Fachmedien.

Petko, D. (2008). Unterrichten mit Computerspielen. Didaktische Potenziale und Ansätze für den gezielten Einsatz in Schule und Ausbildung. *MedienPädagogik. Zeitschrift für Theorie und Praxis der Medienbildung*, *15*, 2008

Torgerson, C. & Zhu, D. (2003). A systematic review and meta-analysis of the effectiveness of ICT on literacy learning in English. In: *Research Evidence in Education Library*. (pp. 5-16). London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

van der Linden, W. J. & Glas, C. A. (Eds.). (2000). *Computerized adaptive testing: Theory and practice.* Dordrecht: Kluwer Academic.

Verdú, E., Regueras, L. M., Verdú, M. J., De Castro, J. P. & Pérez, M. A. (2008). An analysis of the research on adaptive learning: the next generation of e-learning. *WSEAS Transactions on Information Science and Applications*, *5*(6), 859-868.

Warren, S. J., Dondlinger, M. J. & Barab, S. A. (2008). A MUVE towards PBL writing: Effects of a digital learning environment designed to improve elementary student writing. *Journal of Research on Technology in Education*, *41*(1), 113-140.

Wechselberger, U. (2009). Lernspiele aus pädagogischer Sicht. In: Hofmann, Berward/ Ulbrich, Hans-Joachim (Hg.): *Geteilter Bildschirm – getrennte Welten? Konzepte für Pädagogik und Bildung*. (pp. 145-152). Konzepte für Pädagogik und Bildung. München.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of. *Behavioral and Brain Functions, 2*(1), 19.

Wouters, P., van Nimwegen, C., van Oostendorp, H. & van der Spek, Erik D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105* (2), 249-265.

Study 6:

Behavioral and neurocognitive evaluation of a web-platform for game-based learning of orthography and numeracy

Mojtaba Soltanlou, Stefanie Jung, Stephanie Roesch, Manuel Ninaus, Katharina Brandelik, Jürgen Heller, Torsten Grust, Hans-Christoph Nuerk & Korbinian Moeller

Abstract

Recent years have seen a considerable increase in informal educational environments complementing formal educational settings such as schools. In this chapter, we will report results on the efficacy of a web-platform for game-based learning of orthography and numeracy. Besides the behavioral assessment of the platform, we focused specifically on neurocognitive changes due to training on the platform. These neurocognitive data are particularly informative to understand how game-based learning leads to performance improvements, and also might help us to develop new instructional designs.

Our web-based platform hosts several learning games, aiming at fostering orthography and numeracy skills. Learning games enable individual learning independent from formal learning environments - anytime and anywhere. Behavioral results revealed promising learning effects, particularly for orthography. In the next step, neurocognitive changes during arithmetic learning were assessed. Results indicated that arithmetic learning in our informal environment led to strategy changes, previously reported for the development of arithmetic competences in formal learning settings for both adults and children. Altogether, the findings suggest that improvements in orthography and numeracy can be achieved in joyful and less stressful informational environments such as our web-platform for game-based learning. We suggest that the additional implementation of adaptivity in such learning games to better meet individual needs should further increase obtained training effects in the future. Instructional implications of these findings and the relevance of neurocognitive data for learning are discussed.

1. Introduction

Over the last decade, it has become increasingly evident that learning is not limited to traditional, formal, and institutional contexts such as schools, but is diffusing into informal contexts. As was seen in Chap. 5 of this book, informal game-based interventions might be helpful not only for academic achievement but also for public health issue (see Zurstiege et al., 2017). This is closely associated with the rise of modern digital media, above all the World Wide Web as an informational environment, allowing for accessing information anywhere and at any time. These developments seem to have the power and relevance to change education because learners can now specifically search for and find information according to their interests, needs, and abilities allowing for a personalization of their informational environment.

In the present chapter, we describe the developmental and behavioral as well as neurocognitive evaluation of a web-based platform for game-based learning of orthography and numeracy—reflecting the idea of providing informal, low-threshold, and adaptive learning opportunities for what have been called key competencies for our knowledge and information societies. To promote such key competencies, diagnostic assessment of individual abilities and needs is essential. Automatized adaptive methods as provided by some web-based learning platforms can provide such an assessment. First we give a brief introduction into the relevance of orthography and numeracy as well as the idea of using digital games for learning. Thereafter, we describe the web-based learning platform and its evaluation with a focus on neurocognitive effects of training. We conclude with a discussion of the results and their implications for considering neurocognitive correlates of learning.

1.1. The Relevance of Literacy and Numeracy

Literacy and numeracy (i.e., basic reading and writing as well as numerical and arithmetical abilities) are key competencies for navigating life in our post-industrial, knowledge-based societies. Consequently, it seems obvious that insufficient literacy and numeracy carry severe disadvantages for the affected individuals (Hanushek & Woessmann, 2010; Parsons & Bynner, 2005). Empirical data clearly indicate that poor literacy increases an individual's risk for school dropout, low educational achievement, and unemployment (Esser, Wyschkon, & Schmidt, 2002). Moreover, children with poor reading or writing abilities may be aware of their disadvantages and the social relevance of sufficient literacy. They are more inclined to suffer from behavioral and emotional problems (e.g., Daniel et al., 2006). Similarly, insufficient numerical and arithmetical abilities are known to be detrimental to an individual's educational achievement and career, as well as their health prospects (Parsons & Bynner, 2005). Despite these individual disadvantages associated with poor literacy and numeracy, it also needs to be noted that poor literacy and numeracy lead to immense socio-economic costs (e.g., Butterworth, Varma, & Laurillard, 2011). In the case of insufficient literacy, the estimated annual costs range between £45,000 and £53,000 per affected individual, accumulating to a total of £1.73 to £2.05 billion every year for the UK (Gross, 2006). A similar study addressing the costs of poor numeracy estimated those costs to sum up to £2.4 billion per annum, again for the UK (Gross, Hudson, & Price, 2009).

These results clearly indicate that consequences of insufficient literacy and numeracy are not only relevant for the affected individuals but also for society on a larger scale—particularly Western knowledge-based societies (Beddington et al., 2008). Against this background, it is obvious that ensuring successful numeracy and literacy education is of major societal importance. Additionally, it is necessary to develop effective training programs to benefit students with learning difficulties (i.e., dyslexia and/or dyscalculia). The potential returns from investing in such efforts are considerable. For instance, the OECD (Hanushek & Woessmann, 2010) estimated that "an improvement of one-half standard deviation in mathematics and science performance at the individual level implies, by historical experience, an increase in annual growth rates of GDP [gross domestic product] per capita of 0.87 %" (p. 17). Applied to Germany, this reflects a sum of about 25 billion EUR. However, as learning disorders were found to be largely resistant to conventional teaching methods, there is the need to look for new and innovative intervention approaches. Therefore, investigating new possibilities to foster literacy and numeracy, such as the application of information technology and new digital media, is a worthwhile endeavor for research and practice.

Over the last decades, digital media have become part of children's everyday life, with considerable influence on free-time activities, interpersonal interactions, and peer-group communication. Nevertheless, formal education and teaching of numeracy and literacy is still largely based on traditional approaches, such as textbooks and worksheets. As a medium for learning, digital games or gamebased environments offer interesting possibilities to motivate and engage students in learning (Chen, Liao, Cheng, Yeh, & Chan, 2012).

1.2. Digital Games for Learning

Playing digital games can be highly engaging and rewarding, and thus games have become a ubiquitous part of our children's daily lives. According to the Interactive Software Federation of Europe (ISFE), 25 % of Europeans, including both adults and children, play digital games at least once a week (ISFE, 2012). Although most people think of digital games as entertainment, there is increasing interest in using digital games to enhance education (for a review see Boyle et al., 2016), combining the compelling aspects of games with instruction. Recent studies indicate that the use of game-based tasks or the implementation of game elements in conventional cognitive tasks can not only increase users' motivation and engagement, but also improve their performance (e.g., Mekler, Brühlmann, Tuch, & Opwis, 2017; Ninaus et al., 2015; Prins, Dovis, Ponsioen, Ten Brink, &Van der Oord, 2011; for a review see Lumsden, Edwards, Lawrence, Coyle, & Munafò, 2016). Thus, digital games provide increasingly important strategies for learning, educational interventions, and cognitive training. The compelling nature of games keeps users motivated to play or interact with the learning application (Erhel & Jamet, 2013; Ninaus et al., 2013), and this can be attributed to certain mechanics within the game itself. For instance, Garris, Ahlers, and Driskell (2002) emphasized the importance of immediate feedback, reflection, and active involvement in games. Interestingly, according to the Self-Determination Theory (Deci & Ryan, 2000), the intrinsic appeal of games can be explained by their ability to satisfy basic psychological needs for competency, autonomy, and relatedness which—when experienced—increase users' motivation and engagement (Przybylski, Rigby, & Ryan, 2010).

Unfortunately, though, only a minority of games for learning were developed based on recent research findings, and even fewer are evaluated empirically. In the domains of literacy (e.g., Tintenklex: http://www.legasthenie-software.de/) and numeracy (e.g., Semideus, http://www.flowfactory.fi/semideus/; The Number Race: http://www.lacourseauxnombres.com; for an overview see also Moeller, Fischer, Nuerk, & Cress, 2015) game-based solutions seem to be particularly promising. Since research on game-based learning in the domains of literacy and numeracy is still in its infancy, only a few studies have examined positive effects on learning and their comparability to conventional learning methods (e.g., Kast, Baschera, Gross, Jäncke, & Meyer, 2011; Ninaus, Kiili, McMullen, & Moeller, 2016; Wilson, Dehaene et al., 2006; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). Thus, the current project is aimed at designing, implementing, and evaluating a web-based learning platform hosting various games for learning orthography and numeracy. In the following section, key features of the web-based platform and the embedded learning games are described.

2. First step: Behavioral Evaluation

2.1. Introduction

2.1.1 A Web-Platform for Game-Based Orthography and Numeracy Learning

The web-based learning platform for orthography and numeracy (http://lernplattform.iwm-kmrc.de) was developed in an interdisciplinary collaboration between psychologists, linguists, and computer scientists of the University of Tübingen and of the Leibniz-Institut für Wissensmedien in Tübingen. The learning platform hosts several learning games that aim at fostering orthography and numeracy skills in secondary school students, which starts from the fifth grade in Germany. The learning games are embedded in a web browser to facilitate cross platform playability, and, thus, to enable individual learning independent from formal learning environments- anytime and anywhere.

2.1.2. Features of the Learning Platform

The learning platform contains various features to personalize its use and encourages social interaction (a detailed description of the learning platform can be found in Jung et al., 2015).

On a front page, students are invited to create a profile and to register (Fig. 7.1, "User Login"). Setting up profiles on the learning platform offers the opportunity to monitor an individual's learning progress. Once the profile has been set up, an individual profile page is generated (Fig. 7.1, "Mein Konto"). This profile page allows students to add and to share personal information (e.g., age, gender, preference for school subjects, etc.) without being obliged to disclose personal data. The front page enables access to all hosted learning games (Fig. 7.1, "Spielen") and informs about currently logged in users and new users who recently joined the community (Fig. 7.1, "Logged in users" and "New users"). Moreover, a chat forum is provided that invites chat with other users or contact with psychologists and programmers when necessary (Fig. 7.1, "Forum"). These implemented features are provided by a client-server architecture based on the Google Web Toolkit (GWT). GWT is a development tool that enables implementation of complex browser-based applications. The included Java-to-Javascript compiler allows developers to design rich internet applications solely written in Java. Additionally, GWT provides a large library of widgets and panels and comprises several built-in methods to communicate with a server (e.g., remote procedure calls). A database server, which runs the object-relational database management system PostgreSQL, logs all user-generated data (e.g., textual input, movements, and current game states).

These logs are archived and provide crucial information about users' performance and development (Giorgidze, Grust, Schreiber, & Weijers, 2010).

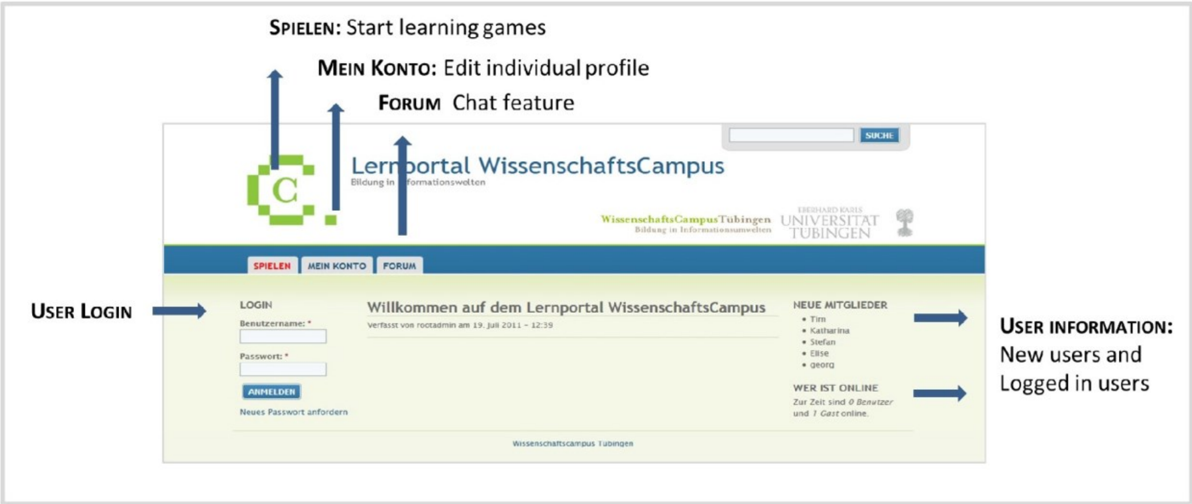


Fig. 1: Front page of the learning platform

2.1.3. Learning Games

The learning games rest on the identification of individual difficulties and developmental trajectories for orthography and arithmetic competencies assessed in more than 400 fifth and sixth graders (Huber, Fischer, Moeller, & Nuerk, 2013; Huber, Moeller, & Nuerk, 2012). Moreover, theoretical considerations including recent findings in the relevant domains should ensure optimally effective and individually tailored learning. Altogether, the learning platform hosts four numerical and four spelling games. A special feature of these games is their social interactivity: most of the games are developed as multiplayer games for up to four players and at least one computer-controlled opponent (e.g., number line game and multiplication game, see below for a description). To win a game, both fast and accurate responses are required.

Orthography games are based on in-depth linguistic analyses covering relevant characteristics of the German language, such as phonological (e.g., the perceptual sensitivity of vowel duration) and morphological (e.g., word-formation) aspects. These games are designed to foster an understanding of various German spelling rules. In German, double consonants (e.g., ff, mm, etc.) usually follow short vowels. As awareness of vowel duration (e.g., offen vs. Ofen) is vital for correct gemination (Ise & Schulte-Körne, 2010; Landerl, 2006), the "Gemination Game" aims at enhancing students' awareness of short and long vowels. The game procedure is threefold: First, vowel duration in an audibly presented word (e.g., Ball) needs to be identified by clicking on a respective symbol (i.e., an exploding ball for "short vowel" or a little man pulling a rope for "long vowel”). Second, in a mini-game, coins depicting double consonants fall from the top of the screen. Students should collect as many coins as possible showing the same double consonant as identified in the target word (e.g., ll). They gain a score point for each collected coin. Third, students are requested to spell the target word correctly by typing it in.

The game “LeTris” also addresses correct spelling. It was designed in the style of the well-known game “Tetris,” but instead of geometric figures, falling letters need to be rotated and arranged according to their order in a word, which is presented audibly. When the word is spelled correctly, the line of letters clears away and score points are gained. When the word is spelled incorrectly, the letters remain on the screen. Consequently, each spelling error yields one more line of letters piling up, finally resulting in “game” The notion of word families (i.e., words that share the same root or morpheme such as read, readable, etc.) is important in language education and associated with literacy performance (Carlisle & Katz, 2006). Word families might be used to derive the spelling of related words. The game “Word Families” fosters this strategy as follows: First, a target word is presented (e.g., child). Second, this target is followed by a range of other words of which some are part of the same word family. Students are requested to remember only those words that belong to the same word family as the target (e.g., children, childish, childhood, etc.). Finally, students are asked to type in and spell correctly all the words they can remember.

The game “Word Building Blocks” also focuses on the approach of word families and was designed as a choice-reaction task. At first, students are requested to select one out of four morphemes (e.g., “old”). Subsequently, they are asked to find and highlight as many words as possible belonging to the same word family, out of various other words (e.g., old timer, oldie, car, spoon, cold, etc.).

Numerical games were designed in line with recent results regarding number processing (e.g., Huber et al., 2013; Link, Huber, Nuerk, & Moeller, 2014; Link, Nuerk, & Moeller, 2014) and cover different numerical principles. The games are predominantly developed as choice-reaction tasks presenting two or more possible solutions to a given problem. The game “Multiplication” aims at enhancing students’ ability to solve multiplication problems. To this end, students are requested to select a multiplication problem (i.e., 3 _ 9) that corresponds to a previously given multiplication result (e.g., 27) out of several response options. The theoretical basis of this game is research indicating that multiplication facts are retrieved from long-term memory once they are learnt by heart (e.g., Delazer et al., 2003; Domahs, Delazer, & Nuerk, 2006).

Addition problems are trained in the game “Partner Number.” First, students are asked to pick one out of four given numbers (e.g., 6). Subsequently, they need to select the corresponding number that adds up to 10 (i.e., 4) out of several response options. Only the player who is the first to choose the correct solution gains a score point. The theoretical background of this game is research showing that mastery of the base-10-place-value structure of the Arabic number system is essential for multi-digit number processing (Nuerk, Moeller, Klein, Willmes, & Fischer, 2011; Nuerk, Moeller, & Willmes, 2015 for reviews).

The “Carry Game” extends the training of addition problems. This game specifically trains carry-over operations that can be solved by the following strategy: adding up to the next decade before adding the rest of the addend. The procedure in the “Carry Game” is threefold: First, students choose an addition problem (e.g., 4 C 9). Second, they are requested to select the number that needs to be added to the first summand (i.e., 4) to add up 10 (solution: 6). Third, they are required to indicate which number remains to be added for the correct solution (i.e., 13; solution: 3). This game is based on research indicating that the carry operation poses a particular difficulty in mental arithmetic in both children and adults (Moeller, Klein, & Nuerk, 2011a, 2011b).

In the “Number Line Game” students are trained in mapping numbers to space (Link, Huber et al., 2014). To pursue this aim, students are required to mark the correct position of a given number (e.g., 43) on a plain number line by means of the mouse cursor. Once a marker has been set, slower opponents cannot place their marker at or around the same location (i.e., 5 % deviation). Thus, players need to answer both quickly and accurately. The theoretical basis of the number line game are findings showing that (a) performance in the number line estimation task is associated reliably with arithmetic performance (e.g., Booth & Siegler, 2006) and (b) that training this task also improves arithmetic performance (Link, Moeller, Huber, Fischer, & Nuerk, 2013; Whyte & Bull, 2008).

2.2. Methods

In a first evaluation of the learning platform, intervention effects of three of the arithmetic (“Multiplication,” “Partner Number,” and “Carry Game”) and three of the spelling games (“Gemination Game,” “Word Families,” and “Word Building Blocks”) were appraised. Two fifth- and sixth-grade classes (n D 47, 19 female) of two public secondary schools participated in a study following a crossover design. One class (Group 1, see Fig. 7.2) received three sessions of the spelling training first, followed by three sessions of the arithmetic training (both about 45 min each). This procedure was reversed for the other class (Group 2).

Both arithmetic and spelling performance were assessed at three time points to evaluate learning progress: at the beginning, and then after three and six training sessions (see Fig. 7.2). At these three time points, arithmetic competencies were assessed by a speeded paper-pencil test covering the basic arithmetic operations (addition, subtraction, multiplication, division) with 36 problems each. In the spelling assessment, children had to perform a writing-to-dictation task and completed 28 fill-in-the-blank sentences by inserting a respective target word. Target words covered all central orthographic aspects of the German written language such as capitalization of initial letters, germination, and lengthening signs.

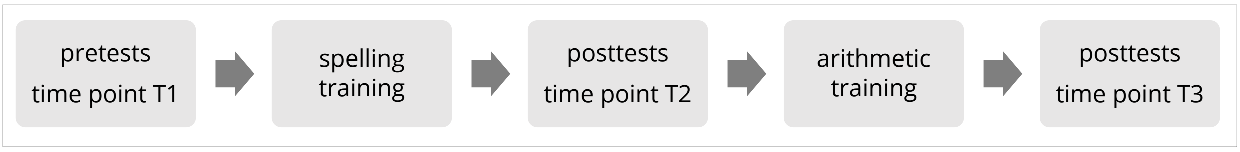


Fig. 2: Pretests, Posttests and training schedule, Group 1

2.3. Results

Results indicated that playing the learning games had a beneficial effect on children's arithmetic performance. However, this effect was not specific to training with the arithmetic games. This indicated that some aspects of the orthography training may have improved arithmetic performance as well (i.e., sequential, symbolic, and spatial processing of items). Finally, training effects were only observed for comparably easy problems across all basic arithmetic operations (see Roesch et al., 2016 for a more detailed description of the results). Thus, although arithmetic games successfully trained specific abilities, other training may be beneficial by training more domain-general aspects such as sequential problem processing, sustained attention, or working memory; this needs to be further studied in more detail.

Concerning orthography, training effects could be separated from other temporal effects by an analysis based on a Linear Logistic Test Model for measuring change (Fischer, 1995). The general temporal trend indicated an initial improvement, and a subsequent decrease back to the starting level (see Fig. 7.3a, no intervention). In addition to this trend, the analysis revealed a significant effect of the spelling training, which specifically improved children's writing performance (solid and dashed line in Fig. 7.3a). This model closely predicts the data at the group level. As shown in Fig. 7.3b, however, the training effect was not strong enough to counteract the general temporal trend (Group 1). This may be due to a lack of motivation in the post-test at T3, or the limited training time, making long-term effects unlikely. In sum, the observed learning effects were promising, especially for orthography. The behavioral results demonstrated the general applicability of the developed learning platform and its games, as well as their effectiveness in conveying orthography and arithmetic competencies. One major advantage of digital training on our web-based platform is the possibility to meet individual requirements by introducing adaptivity: implementations similar to those proposed, i.e., implementing adaptivity based on a Brain-Computer Interface (BCI) in an educational context, as it is shown in Chap. 8 of this book (Spüler et al. 2017), may further enhance the already obtained training effects in the future. In the following sections, we discuss neurocognitive data reflecting learning effects, including our initial results based on the online learning platform.

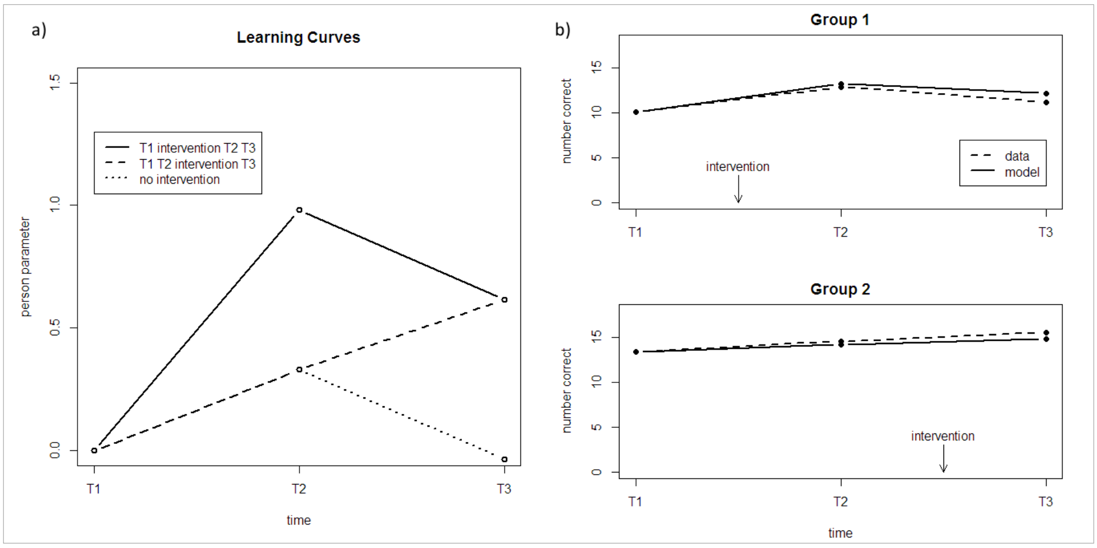


Fig. 3: Panel 3a provides an illustration of the general temporal trend (dotted line) and the intervention effects. Panel 3b shows that the observed number of correctly written items is closely predicted by the LLTM results at group level.

3. Second Step: Neurocognitive Evaluation of Arithmetic Learning

3.1. Introduction

Investigating the neurocognitive foundations of arithmetic development and learning is a complementary approach to uncover how children improve their behavioral math skills. For instance, Supekar et al. (2013) observed that while neural correlates predicted math learning in children, behavioral measures failed to do so. Therefore, it is essential to go beyond behavioral investigation of arithmetic development and learning in children in order to come up with appropriate educational and therapeutic interventions for each individual. As is seen in Chap. 8 of this book, Spüler et al. (2017) elaborate on an important application of neurocognitive data over and beyond behavioral assessment for adapting learning environments to individual needs.

Neurocognitive studies have already shown that individuals may rely on different brain networks to solve arithmetic problems (Grabner et al., 2007; Grabner, Ansari et al., 2009). Grabner, Ansari et al. (2009) observed higher activation of language-related parietal areas in the left hemisphere in individuals with higher math competency as compared to individuals with lower math competency. Activation of this area is usually interpreted to reflect retrieval strategies (e.g., knowing that 2 _ 3 gives 6 without effortful calculation) in arithmetic problem solving (Zamarian, Ischebeck, & Delazer, 2009), which is one of the dominant strategies used after sufficient training. Delazer et al. (2003) found that different training methods all led to a successful use of retrieval strategies in adults. However, these different training methods led to differing brain activation patterns, whereas behavioral performance improvement was similar.

3.1.1. Neurocognitive Foundation of Arithmetic Learning in Adults

According to Poldrack (2000), learning is a shift from general purpose processes to more task-specific processes. In line with this definition, arithmetic learning is characterized by a strategy shift from more effortful and algorithm-based calculations to more economic retrieval processes, which results in characteristic changes in brain activation patterns (Zamarian et al., 2009). The fronto-parietal network usually found to be active during number processing involves both areas associated with domain-general and domain-specific processing engaged in arithmetic problem solving. In this network, frontal areas are associated with supplementary domain general cognitive processes such as working memory (WM) and planning in mental calculation, while parietal areas are associated with magnitude processing of numerals and domain-specific processes (for a review see Arsalidou & Taylor, 2011). According to the triple-code model of number processing (Dehaene & Cohen, 1997; Dehaene, Piazza, Pinel, & Cohen, 2003), domain-specific processing of number magnitude information is subserved by neurons in and around the bilateral intra-parietal sulci, while visuo-spatial demands of number processing are associated with the superior parietal lobule, and language-related demands of number processing (e.g., retrieving arithmetic facts from long-term memory) are dedicated to the left angular gyrus.

In adults, Delazer et al. (2003) found that arithmetic learning goes along with a decrease of cognitive load in verbal and visuo-spatial WM, less engagement of attentional control, strategy planning and self-monitoring, and also less application of mathematical rules and algorithms in calculation associated with frontal activation (for a review see Zamarian et al., 2009). Furthermore, it is assumed to induce more specific processing of number magnitude information in parietal and language-related processes in left temporal areas of the brain. Such a shift in solution strategy was repeatedly observed to be accompanied by reduced activation in the fronto-parietal network of number processing and increased activation in language-related parieto-temporal areas in the left hemisphere in adults (Delazer et al., 2003, 2005; Grabner, Ischebeck et al., 2009; Ischebeck et al., 2006; Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007; Ischebeck, Zamarian, Schocke, & Delazer, 2009; Pauli et al., 1994; for a review see Zamarian et al., 2009). Generally, there seems to be a shift from frontal to parietal regions, and then within parietal regions with increasing proficiency reflecting reduced demands on domain-general cognitive processes and increased domain-specific numerical processes (Zamarian et al., 2009). Interestingly, it was found that this shift in brain activation can already happen after eight repetitions of arithmetic problems in adults (Ischebeck et al., 2007). However, there is no agreement on the specificity of this shift in brain activation with training. Importantly, the shift seems to depend on the learning method (Delazer et al., 2005), on the arithmetic operation (Ischebeck et al., 2006), and on the experimental design (Bloechle et al., 2016), and may not even be specific to arithmetic learning (Grabner & De Smedt, 2012; Grabner, Ischebeck et al., 2009). Moreover, recent studies suggested a pivotal role of hippocampal systems associated with long-term memory functioning in arithmetic learning in adults (for a review see Klein et al., 2016) as well as in children (e.g., Qin et al., 2014).

3.1.2. Neurocognitive Foundation of Arithmetic Development and Learning in Children

Neuroimaging studies of arithmetic learning in children are still scarce and most of our knowledge is drawn from studies comparing children with adults using cross sectional designs or involving math tutoring and assessing longitudinal age- and/or training-related effects. There is agreement on an increase of automated processes in arithmetic problem solving with age, which is reflected by enhanced activation of domain-specific parietal areas and reduced activation of complementary domain general frontal areas (e.g., Kaufmann, Wood, Rubinsten, & Henik, 2011; Kucian, von Aster, Loenneker, Dietrich, & Martin, 2008). This developmental increase of automated processing is reflected by a frontal-to-parietal shift of brain activation (i.e., reduced activation of frontal areas and increased activation of parietal areas (e.g., Ansari, 2008; Rivera, Reiss, Eckert, & Menon, 2005)—very similar to that observed in adults after training.

For instance, Kawashima et al. (2004) found greater activation of right parietal areas during subtraction and multiplication problem solving in adults compared to children, but no difference in frontal areas (see also Kucian et al., 2008). Moreover, they found reduced activation of areas associated with attentional processing in adults as compared to children (Kucian et al., 2008; see also Cantlon et al., 2009). Moreover, Rivera et al. (2005) suggested that adolescents recruit left parietal areas, and the left occipito-temporal area (see also Emerson & Cantlon, 2015), while children still recruit bilateral frontal areas and attention-related areas (see also Prado, Mutreja, &

Booth, 2014). This more pronounced frontal activation indicates that greater demands on WM and executive function in children are required to achieve similar performance to adolescents. Furthermore, stronger activation of the left hippocampus was observed in children, which was interpreted as revealing higher demands on declarative and procedural memory systems (Rivera et al., 2005). The transitional role of the hippocampal system in arithmetic development and learning was investigated by Menon and colleagues. Supekar et al. (2013) found that pre-training hippocampal volume, and the functional association of the hippocampus with frontal areas subserving domain-general processes, predicted training-related arithmetic improvement in children. Surprisingly, no behavioral measures, including IQ, WM, and general math abilities, did so (see also Evans et al., 2015). Moreover, in a longitudinal study on children, Qin et al. (2014) found a critical transient role of the medial temporal lobe, including the hippocampus, in arithmetic learning. They suggested that the hippocampal system is pivotal in the shift from procedural to retrieval strategies, as shown by the increasing involvement of the hippocampus and the decreased involvement of fronto-parietal networks in arithmetic problem solving with age (Qin et al., 2014). Furthermore, it was also suggested that the neurocognitive correlates of different arithmetic operations are not necessarily identical during development. In a cross-sectional study in children, Prado et al. (2014) found a grade-related activation increase in left language-related temporal areas for multiplication, but a grade-related increase of right quantity- and magnitude-related parietal activation for subtraction. Prado et al. (2014) concluded that fluency in arithmetic problem solving is achieved through different strategies depending on the arithmetic operation: by increasing retrieval, in the case of multiplication, but by increasing efficient procedural strategies in the case of subtraction. In sum, a developmental frontalto-parietal shift of activation along with increased engagement of hippocampal areas seems to accompany arithmetic development and learning in children. These results can be compared with the results of our own studies investigating children's numerical learning using the learning platform, which are described below.

3.2. Methods

3.2.1. Neurocognitive Foundation of Arithmetic Learning on the Learning Platform

Three studies were conducted to investigate changes in brain activation due to arithmetic learning by means of the learning platform: short-term effects of a 2-week training, immediate effects of one session training, and continuous changes during training. In order to evaluate changes in brain activation, functional near-infrared spectroscopy (fNIRS) and electroencephalography (EEG) were utilized. fNIRS comes with some advantages over fMRI involving the possibility of measurement in ecologically valid situations, like school settings (Dresler et al., 2009; Obersteiner et al., 2010), or with whole body movement (Bahnmueller, Dresler, Ehlis, Cress, & Nuerk, 2014). It is suitable for measuring brain activation in children in upright body postures, like sitting behind a desk and in front of a computer. fNIRS is a comparably cheap method in contrast to other brain-imaging methods such as fMRI; it is also easily applicable and highly versatile, which altogether allows frequent measurement repetitions (for more see Ehlis, Schneider, Dresler, & Fallgatter, 2014). With regard to its functioning, it can be said that activation of certain brain areas leads to increased cerebral blood flow in these areas, reflecting increased oxygen consumption (Scholkmann et al., 2014). This results in changes in oxyhemoglobins (O_2Hb) and deoxyhemoglobins (HHb). Non-invasively, fNIRS records these changes as an indirect measure of brain activation.

However, it needs to be mentioned that fNIRS has some limitations, including restricted penetration depth and spatial resolution (Wabnitz et al., 2010), confounding influences of extracranial signals (Haeussinger et al., 2014), and peripheral hemodynamic parameters such as skin perfusion (for a review see Scholkmann et al., 2014). Balancing advantages and limitations, fNIRS seems to be a promising tool to investigate cognitive development in children and adults (e.g., Dresler et al., 2009; Verner, Herrmann, Troche, Roebers, & Rammsayer, 2013). On the other hand, EEG offers very high temporal, but low spatial, resolution, and is relatively sensitive to motion. It is much cheaper than many brain imaging tools and because it is portable, it is easily applicable in very different situations such as in schools. An advantage of EEG is that brain oscillations (i.e., neural electric activation) can be recorded and analyzed in different ways. For instance, cognitive and motor processes lead to so-called event-related potentials (ERPs) and also to changes in continuous EEG in the form of event-related synchronization and desynchronization (ERS/ERD) (Pfurtscheller, 2001). ERS/ERD for specific frequency bands has been associated with particular cognitive functions (Pfurtscheller, 2001; Pfurtscheller & Da Silva, 1999) and thus reflect quantifiable measures of brain dynamics (Pfurtscheller & Aranibar, 1977). Previous studies indicated that theta and alpha frequency bands are sensitive to cognitive tasks such as arithmetic processing (e.g., Dolce &Waldeier, 1974). For instance, task complexity, attentional and domain-general cognitive demands, as well as memory load, lead to theta ERS (i.e., an increase in theta power) but alpha ERD (i.e., a decrease in alpha power) (Antonenko, Paas, Grabner, & van Gog, 2010; Gevins, Smith, McEvoy, & Yu, 1997; Klimesch, 1999; Moeller, Wood, Doppelmayr, & Nuerk, 2010; Pfurtscheller & Da Silva, 1999)3

3.3. Results

3.3.1. Short-Term Neurocognitive Changes

A group of typically developing children received seven sessions of training on simple (one-digit x one-digit, e.g., 3 x 7) and complex (one-digit x two-digit, e.g., 3 x 17) multiplication using the multiplication game implemented on the learning platform over 2 weeks. The game entailed a multiple-choice paradigm in which children had to click on the correct one out of 12 presented solution choices. The task was speeded and had to be performed in competition against a virtual computer player. The effect of a 2-week training period was evaluated using simultaneous fNIRS and EEG before and after the training. The 2-week training data indicated that children became more efficient in the trained compared to untrained multiplication problems, which was reflected in faster responses and fewer errors. With respect to the trained simple condition, no significant change was observed in the fNIRS data. However, decreased alpha ERD (i.e., increased alpha power) was found at the central frontal electrode. This decrease suggests reduced cognitive effort in general (Pfurtscheller, 2001), probably reflecting more retrieval-based solution strategies in trained simple multiplications after the training (see also Gevins et al., 1997).

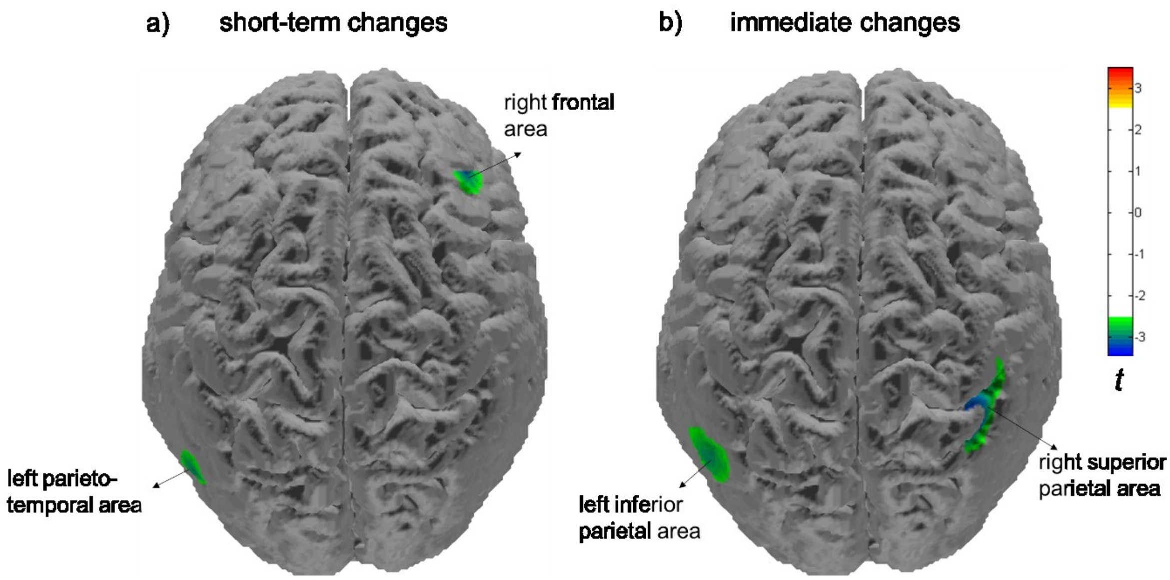


Fig. 4: Figure 4: Panel 4a illustrates changes in brain activation due to short-term arithmetic training in children. Panel 4b illustrates immediate brain activation changes due to arithmetic training in children. Colour coding: blue/green indicates a reduction of activation.

For the trained complex condition, fNIRS findings revealed reduced activation in left parieto-temporal as well as right frontal areas after training (cf. Fig. 7.4a). It has been shown that learning changes the relation of domain-general to more domain-specific processing demands, which is indicated by reduced activation in several brain regions (Poldrack, 2000). In agreement with this and also studies in adults (for a review see Zamarian et al., 2009), we observed reduced activation in right frontal areas, associated with executive control and WM, after the training (see also Soltanlou et al., 2017). This indicated an increase in retrieval-based solution strategies, which depend less on domain-general cognitive processes, following complex multiplication training (see also Prado et al., 2014).

Furthermore, the observation of decreased activation in left parieto-temporal areas is in line with longitudinal and training studies in children (Qin et al., 2014; Supekar et al., 2013), but is contradictory to multiplication training studies in adults, which have reported increased activation of left language-related parietal areas after training (for a review Zamarian et al., 2009; but see Bloechle et al., 2016). It seems that although a shift from effortful procedural to retrieval-based strategies is represented by a frontal-to-parietal shift of brain activation and increased activation of language-related parieto-temporal areas in adults, the same might not necessarily be true for children (see also Supekar et al., 2013). This difference might be due to more stable neural substrates of arithmetic processes in adults as opposed to children (Qin et al., 2014). Furthermore, the reduced activation of language-related left parietal areas is in line with an fMRI study by Menon, Rivera, White, Glover, and Reiss (2000), which also reported decreased activation of language-related parietal areas with increasing arithmetic expertise (see also Amalric & Dehaene, 2016). It should be noted, however, that even for adults, several brain areas, and not only left-hemispheric language-related parieto-temporal areas, have been associated with retrieval processes after multiplication training (Bloechle et al., 2016; Delazer et al., 2005).

3.2.2. Immediate Neurocognitive Changes

Immediate training effects were evaluated using simultaneous fNIRS and EEG efore and after a single session of the same training by means of a multiplication game implemented on the learning platform. Interestingly, in the absence of any significant behavioral improvement, a pre- and post-test comparison of fNIRS data showed reduced activation at left inferior parietal and right superior parietal areas after training, during the calculation of trained complex multiplication problems (cf. Fig. 7.4b). This finding is in line with the longitudinal fMRI data by Qin et al. (2014), indicating reduced activation of bilateral parietal regions in children after a year of schooling. A decrease in parietal activation, which is usually associated with processing numerical magnitude, may indicate that after the training, children seemed to rely less on manipulations of numerical magnitudes to solve the task. This finding is in line with an fMRI study by Ischebeck et al. (2007) that reported similar brain activation changes in adults after a training on complex multiplication Problems. In sum, this decrease seems to indicate reduced demands on exact calculation (e.g., Dehaene, Molko, Cohen, &Wilson, 2004). Moreover, we also observed increased alpha ERD at the central parietal electrode for complex multiplication, which might indicate increased visual attentional processes (Klimesch, Sauseng, & Hanslmayr, 2007). In simple multiplication, no significant difference was observed, probably because children of this age (i.e., fifth grade) were quite advanced in solving simple multiplication problems, and more training than just one session would have been needed to improve their performance.

3.2.3. Ongoing Neurocognitive Changes

In the next step, neurophysiological changes during arithmetic learning using the same multiplication game in the learning platform were investigated in children. Findings showed gradually increasing power in the theta (4–7 Hz) and lower alpha (8–10 Hz) bands, but not in the upper alpha band (10–13 Hz) over six repetitions of multiplication problems (cf. Fig. 7.5). Importantly, this fits nicely with the results of Klimesch, Vogt, and Doppelmayr (1999), who reported higher theta power in individuals with high as compared to individuals with low calculation skills—corresponding to the training effect observed in the present study. Increased theta power during multiplication training in children is probably associated with acquiring new information (Klimesch, 1999), which usually needs additional attention and mental effort, rather than retrieving existing knowledge (e.g., Gevins et al., 1997; Mizuhara & Yamaguchi, 2007; Sammer et al., 2007). In line with this argument, Gevins et al. (1997) found an increase in theta power resulting from a brief WM training in adults and attributed it to the extra effort required to focus attention over a long time. According to a model by Baroody (1983), mathematical training entails a shift from slow procedural processes towards compact procedural strategies, which can lead to a temporarily increased involvement of domain-specific and domain general cognitive processes (see also Núñez-Peña & Suárez-Pellicioni, 2012; Prado et al., 2014). In the present study, it seems that children primarily improved their procedural and algorithm-based strategies over six repetitions of the multiplication problems, which led to fewer errors but also increased theta power (see also Barrouillet & Thevenot, 2013). These findings point to the idea that procedural strategies improve as a function of learning in children in an early phase of building up new arithmetic fact knowledge (see also Lemaire, 2016; Zhou et al., 2011).

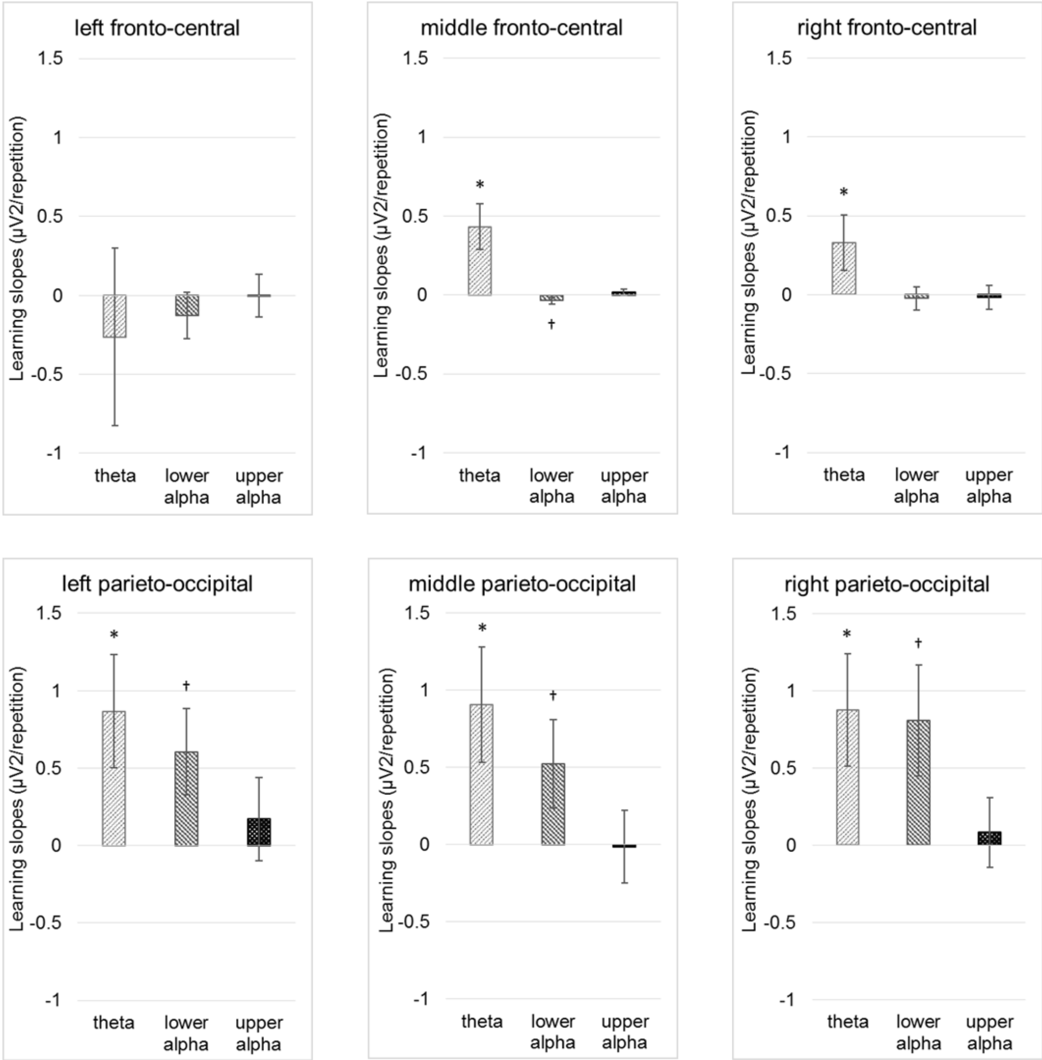


Fig. 5: Unstandardized coefficients (learning slopes) of EEG power density in theta, lower alpha, and upper alpha bands for different brain regions. Positive values display increased power density during six repetitions. Negative values display decreased power density during six repetitions. Error bars depict 1 SE. *: FDR corrected $p < .05$; †: FDR corrected $p < .1$.

Although it seems that these few repetitions led to a shift from more effortful to more efficient procedural processes, there might be a shift to retrieval strategies as well, or at least an increase in retrieval processes as a part of efficient procedural strategies. It has been shown that executive functioning is involved even in arithmetic fact retrieval (Hinault & Lemaire, 2016). Therefore, an increased use of retrieval processes during arithmetic learning might also lead to increased involvement of domain-general cognitive processes, resulting in theta increase. In sum, increased theta power seems to indicate more efficient performance, resulting from not only more efficient procedural but also more retrieval-based strategies. Furthermore, an increase in lower alpha power at bilateral occipito-parietal electrodes was observed in the present study. This is in line with previous arithmetic and cognitive training studies in adults (Gevins et al., 1997; Grabner & De Smedt, 2012). For instance, Gevins et al. (1997) suggested that increased automaticity through training is associated with increased power in the lower alpha band, reflecting a reduction in domain-general cognitive demands (Pfurtscheller, 2001), task difficulty (e.g., Gevins et al., 1997), and attentional demands (Ray & Cole, 1985). It has been found that procedural strategies demand more cognitive resources than retrieval strategies in multiplication problem solving in children (Lemaire, Barrett, Fayol, & Abdi, 1994). Therefore, in agreement with Grabner and De Smedt (2012), we conclude that decreased alpha power is associated with more procedural strategies. In contrast, increased power in lower alpha as observed in the current study may represent more automatic, presumably retrieval-based, strategies in arithmetic problem solving (e.g., Moeller et al., 2010). It seems that children first shift to more efficient procedural strategies before they shift to retrieval strategies.

3.4. Conclusion of Neurocognitive Evaluation

The findings of the above-described studies nicely demonstrated that arithmetic learning using the multiplication game of our learning platform in typically developing children led to measurable neurocognitive changes (even in the absence of reliable improvements in behavioral performance). In particular, the results of our studies indicated that the development of arithmetic competencies occurs in two steps: the first one from slow effortful procedural processes to faster more efficient procedural processes, and the second step to retrieval-based solution strategies. In the following section, the theoretical and practical implications of these results are discussed.

4. General Discussion, Limitations, and Perspectives of the Neurocognitive Approach

A neurocognitive approach to learning and education, which we have embraced in the above studies and which is specifically pursued in the emerging field of Educational Neuroscience, has been both hyped and debunked. For instance, Bowers (2016a) puts forth his skepticism about such an approach: *"Neuroscience cannot determine whether instruction should target impaired or nonimpaired skills. More importantly, regarding the assessment of instruction the only relevant issue is whether the child learns, as reflected in behavior. Evidence that the brain changed in response to instruction is irrelevant".* In the on-going debate, there have been fierce responses arguing that *"Bowers' assertions misrepresent the nature and aims of the work in this new field [of Educational Neurosceince]. We suggest that, by contrast, psychological and neural levels of explanation complement rather than compete with each other" (see also Bowers,* 2016b*; Gabrieli,* 2016*; Howard-Jones et al.,* 2016*).*Without reiterating the back and forth of this controversial debate in more detail, we wish to make two statements about what our approach as described above can and cannot contribute to (educational) science: one about the nature of neurocognitive data, and one on the localization of cognitive functions within the brain.

4.1. The Construct and Its Operationalization: Neurocognitive Data are Just Another Dependent Variable

At times researchers or the public may have gained the impression that neurocognitive data reflect the “real” underlying variables of human cognition and behavior, whereas behavioral variables may only be second-hand, indirect indices. It is important to mention that we do not want to promote such an impression nor do we share the underlying notion. Instead, our position on this is that neurocognitive data are at the beginning just another sort of dependent variables such as error rates, reaction times, eye movements, motion parameters, and other so-called *behavioral* measures. Most of these behavioral measures in learning and instruction research are indirect in the sense that they do not directly reflect the ultimate goal of instruction and education. For instance, in math education, our goal is that children can solve math problems correctly. However, since the beginning of the so-called cognitive era in psychological and educational research (starting in the late 1950s), it is almost undisputed that we do not only need to do behaviorist experiments, in which children simply need to commit fewer errors in math at the end, regardless of why and how. Instead, we are specifically interested in understanding the underlying cognitive processes and representations, the strategies and procedures employed, the contributions of domain (i.e., math)-specific and domain-general (e.g., WM) processes to human development; and, in turn, as we learn to understand these aspects, the best ways to teach such processes, along with developmental issues inherent in learning and instruction.

In fact, however, some of these processes, strategies, and so forth are quite abstract concepts. We cannot study them directly, but only indirectly through operationalization. A particular concept is broken down into graspable questions, and we derive particular hypotheses about specific operationalizations. We usually do not measure WM or mathematics competencies per se, but substitute performance in a task or test that we are confident provides an appropriate proxy measure for the respective construct. The dependent variables in such tasks or tests are not just accuracy (the primary concern in mathematics education), but much more often behavioral measurements like reaction times or eye movement or motion data. In this context, it may not be of interest for education per se whether a particular arithmetic problem is solved 30 ms faster or slower under a given condition or after a particular training—in the end it may have little practical relevance whether one solves a simple arithmetic problem in 800 or 830 ms. However, it may have important theoretical relevance, because it informs us about the way children and adults represent numerical knowledge. For instance, the observation that children represent multi-digit numbers in a decomposed fashion (i.e., separated into units, tens, etc., cf. Nuerk, Kaufmann, Zoppoth, & Willmes, 2004) from early on in their numerical development was derived from reaction time data. Furthermore, eye-tracking data indicated that even a simple addition task involving two-digit numbers seems to comprise at least three different underlying processes (Moeller et al., 2011b). Taken together, this illustrates that most behavioral data are not used to directly measure whether a child has learned something, but to understand what, how, and why a child has learned.

Importantly, the very same goal is pursued by neurocognitive data. We do not want to claim that neurocognitive data are superior to behavioral data. Like behavioral data, neurocognitive data only measure particular indices and not the underlying cognitive processes and representations in the brain - and most certainly not localizations of cognitive functions/representations in the brain (see below for a more detailed discussion of this). However, neurocognitive data are powerful dependent variables, and considering them will help us grasp the underlying processes in children's and adults' learning, by their temporal and spatial characteristics in different conditions and in contrast to other studies. The above-described studies provide a good example for this argument. We know that adults and children can learn multiplication facts by drill when the association between operands and result is repeated sufficiently often. In fact, they usually get faster and more accurate with training, at least in the trained problems. In case we are not interested in the processes and representations underlying this learning effect, we could simply leave it at that. However, when we are interested in the what, how, and why of the learning process, neurocognitive data may be informative. In previous adult studies, the increase in performance was associated with a shift in brain activation argued to reflect a change in strategy from effortful calculation to overlearned arithmetic fact retrieval (Delazer et al., 2003). No such shift was observed in our data. Rather the data suggested that children automatized and facilitated their calculation procedures first. When we only considered the behavioral data, we would only observe that children and adults improved very similarly after training; and thus, we would most likely assume that the same processes underlie their multiplication learning. However, complementary neurocognitive data suggested that brain activation patterns associated with the learning process may differ between children and adults. That is why we think that it is essential to understand the what, how, and why of the learning process to enhance and promote optimal learning procedures, instructions, and environments. Neurocognitive data can contribute to this understanding in addition to other (behavioral) data like eye-movement or motion data.

2.1. Function is Not Region: The Need to Go Beyond Localistic Modular Ideas of Neurocognitive Functioning

Early approaches in cognitive neuroscience have often been localistic and modular. This means that for a particular function like vision, hearing, motion, or somatosensory perception, the aim has been to identify brain regions that are selectively activated, which are then assumed to subserve the function. While this approach has worked to a certain degree for basic sensory and motor functions, as reflected by evidence for a visual cortex, an auditory cortex, a motor cortex, and a somato-sensory cortex, the idea of localized, the modular mental function is less realistic for higher cognitive functions, such as numerical cognition. The original triple code model of numerical cognition by Dehaene (Dehaene & Cohen, 1997) has focused—although white matter connections were implicitly suggested - heavily on such anatomo-functional associations. There was the proposition of a representation of number magnitude in the bilateral intra-parietal sulci, a verbal representation of number in the left-hemispheric superior temporal area and the angular gyrus, and a visual number form area in bilateral inferior occipito-temporal areas. In recent years, however, this model has been extended and networks with their white matter connections were included in network models of numerical cognition and arithmetic (Klein et al., 2016). Moreover, brain regions that were supposed to subserve a particular numerical representation are found to actually serve multiple functions when literature from different fields is considered. For instance, the angular gyrus thought to be involved in arithmetic fact retrieval may serve a more general role in a network with hippocampal structures jointly underlying fact retrieval (e.g., Bloechle et al., 2016; see also Harvey, Klein, Petridou, & Dumoulin, 2013 for the involvement of the intra-parietal sulcus).

The more a cognitive function is evolutionarily advanced, the less it is likely to be subserved by only a single brain region. Instead, whole networks seem to underlie complex cognitive functions like arithmetic, which makes neuroscientific research on these functions at the same time challenging but also more interesting and promising. After this discussion of the general benefits of neurocognitive data for educational science, the final section of the chapter focuses on challenges to and the implications of our results for the idea of numerical learning promoted by a web-based learning platform.

5. Conclusion: Challenges and Implications

A central goal of the development of the learning platform was to provide a personalized learning experience tailored to the specific needs of each individual learner. This is a highly desirable feature given the fact that there is considerable variance in orthography and arithmetic competencies. Moreover, motivation to use the platform needs to be maintained at a high level, in particular for children with specific deficits in orthography or numeracy. Presenting material that is too difficult may cause frustration and should thus be avoided. This may be achieved by adjusting the level of difficulty of the learning games to more directly target existing deficits. Such adaptivity in the platform can be realized at both an intra-individual and an inter-individual level. For intra-individual adaptivity, the system needs to have a valid model of strengths and weaknesses of the individual user. Therefore, these need to be assessed at a fine-grained level—at least for novices to the system—but also in between playing the learning games, to monitor learning progress. To keep these assessments short, an adaptive procedure seems most suitable.

A prototype of such an adaptive assessment approach was developed based on modeling relevant competencies by Item-Response Theory (IRT), a state-of-the-art methodology (Fleischer, Leutner, & Klieme, 2012; Hartig & Frey, 2012). Rich sets of items were analyzed and spelling/numerical errors were evaluated with respect to particular orthographic/arithmetic markers. Item difficulty parameters were determined separately according to the Rasch model, for each arithmetic operation (addition, subtraction, multiplication, etc.) and orthographic rule (capitalization, gemination, lengthening, etc.), and were used to assess proficiency with respect to all of these dimensions through the respective person parameters. The assessment proceeds in an adaptive manner by presenting the most informative item next. Notice that testing a German word like "Rennbahn" provides information on various competencies simultaneously: capitalization (initial capital letter), gemination (double consonant "nn"), and lengthening (long vowel "ah"). Thus, the corresponding information values were integrated to select the next word. Also, note that the spelling of a typed-in word needs to be checked automatically with respect to each of the orthographic markers. Efficient pattern matching algorithms were applied to provide such an online evaluation of the spelling. The adaptive assessment is stopped as soon as a certain level of confidence on the person parameters is achieved, at which point the learning games can be adapted based on these parameters. The difficulty of items in the learning games can be matched to the person parameters, so that the task is neither too demanding nor too easy. In sum, a proof of concept for computer-based adaptive assessment of spelling could be successfully developed and awaits implementation.

Inter-individual adaptivity refers to balancing the odds of two players in competitive learning games and is a functionality that is currently under development. The estimated person parameters of above-described assessments may be used to match players according to their competencies, or to simulate a computer-controlled opponent of the same ability profile. Moreover, game play behavior may be modified with respect to the student's performance relative to (a) her/his current opponent and (b) the set of all registered players. Individual log data may be exploited to provide an index of competency (like an ELO score in chess; Elo, 1978), which allows for specifying handicaps for the better performing student in any pair of players. Such handicaps can be realized via time-delays in item presentation or answer registration. This makes it possible to balance success rates and keep the motivation of the individual players high, quite independent of their ability level. Such inter-individual and intra-individual adaptivity remains a central goal of any learning game: Web-based and computer-based games provide - grounded on IRT based methods - a promising tool for incorporating such adaptivity. In our view, this is essential to prevent poor performers (most in need of learning) from quitting learning games due to lack of success and reward.

Finally, a recapitulating view on our project allows for some important statements about learning and educational practice:

- Informal learning settings (such as our learning platform) can successfully complement formal learning settings in schools as shown by the observed training effects. Newmedia like the internet can be particularly powerful, because the games we designed can be played anywhere where internet connections are available, with any device, inside or outside of school. The application of new digital media and game-based learning for interventions are not limited to academia and education, but also may also be used in a broader field of mental and physical healthcare, as it was shown in Chap. 5 of this book (Zurstiege et al. 2017).
- Game-based learning of cultural competencies like orthography or numeracy works. Math learning (and to a lesser degree, orthography) is usually considered highly aversive by many children, and, unfortunately, also by many teachers. Game-based approaches can provide a new approach to these important learning topics and nevertheless increase learning progress.
- Individualization is not just a matter of human judgment and expertise. While we strongly appreciate individualization and differentiation of the learning progress by teachers, we are also aware that it requires enormous effort. The time needed may not always be available, and in large classes it may sometimes not be easy or even possible to evaluate the particular strengths and weaknesses of each individual child on each new topic properly and correctly. Data-driven adaptivity as we have developed it for diagnostics of the spelling process may help to identify individual problems and help to tailor individual learning and instruction.

- Neurocognitive data can help to enhance our understanding of the learning and instruction process. In this project, we gained first insights that multiplication learning in children might be different from that in adults, although both improve in accuracy and reaction time during the learning process. Neurocognitive data are certainly not the only data allowing us to understand the what, how, and why of learning, but they are nevertheless a powerful complementary dependent variable to enhance this process.

Taken together, mastering literacy and numeracy remains one of the great challenges in education and instruction in our modern knowledge society. At the same time, respective deficits are a severe obstacle in professional and private life, impacting not only individuals but also societies at large. New, innovative methods for administrating and conducting orthography and numeracy training as well as evaluating and understanding the underlying learning processes revealed promising results in our web-based learning project. Because it is essential to achieve mastery of literacy and numeracy for as many as possible, we should use such computer supported, adaptive, and game-based methods to foster learning opportunities and thus enrich life and career prospects for as many children as possible.

References

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. Proceedings of the National Academy of Sciences, 201603205.

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. Nature Reviews Neuroscience, 9(4), 278-291.

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography to measure cognitive load. Educational Psychology Review, 22(4), 425-438.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+ 2= 4? Meta-analyses of brain areas needed for numbers and calculations. NeuroImage, 54(3), 2382-2393.

Bahnmueller, J., Dresler, T., Ehlis, A.-C., Cress, U., & Nuerk, H.-C. (2014). NIRS in motion—unraveling the neurocognitive underpinnings of embodied numerical cognition. Frontiers in psychology, 5, 743.

Baroody, A. J. (1983). The development of procedural knowledge: An alternative explanation for chronometric trends of mental arithmetic. Developmental Review, 3(2), 225-230.

Barrouillet, P., & Thevenot, C. (2013). On the problem-size effect in small additions: Can we really discard any counting-based account? Cognition, 128(1), 35-44.

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., . . . Thomas, S. M. (2008). The mental wealth of nations. Nature, 455(7216), 1057-1060.

Bloechle, J., Huber, S., Bahnmueller, J., Rennig, J., Willmes, K., Cavdaroglu, S., . . . Klein, E. (2016). Fact learning in complex arithmetic—the role of the angular gyrus revisited. Human Brain Mapping.

Bowers, J. S. (2016a). The Practical and Principled Problems With Educational Neuroscience.

Bowers, J. S. (2016b). Psychology, not educational neuroscience, is the way forward for improving educational outcomes for all children: Reply to Gabrieli (2016) and Howard-Jones et al.(2016).

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., . . . Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Computers & Education, 94, 178-192.

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: from brain to education. science, 332(6033), 1049-1053.

Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. Journal of Cognitive Neuroscience, 21(11), 2217-2229.

Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. Reading and Writing, 19(7), 669-693.

Chen, Z.-H., Liao, C. C., Cheng, H. N., Yeh, C. Y., & Chan, T.-W. (2012). Influence of Game Quests on Pupils' Enjoyment and Goal-pursuing in Math Learning. Educational Technology & Society, 15(2), 317-327.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. Journal of learning disabilities, 39(6), 507-514.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. Psychological inquiry, 11(4), 227-268.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic. Cortex, 33(2), 219-250.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. Current opinion in neurobiology, 14(2), 218-224.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. Cognitive neuropsychology, 20(3-6), 487-506.

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic—an fMRI study. Cognitive Brain Research, 18(1), 76-88.

Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C., . . . Felber, S. (2005). Learning by strategies and learning by drill—evidence from an fMRI study. NeuroImage, 25(3), 838-849.

Dolce, G., & Waldeier, H. (1974). Spectral and multivariate analysis of EEG changes during mental activity in man. Electroencephalography and Clinical Neurophysiology, 36, 577-584.

Dresler, T., Obersteiner, A., Schecklmann, M., Vogel, A. C. M., Ehlis, A.-C., Richter, M. M., . . . Fallgatter, A. J. (2009). Arithmetic tasks in different formats and their influence on behavior and brain oxygenation as assessed with near-infrared spectroscopy (NIRS): a study involving primary and secondary school children. Journal of neural transmission, 116(12), 1689-1700.

Ehlis, A.-C., Schneider, S., Dresler, T., & Fallgatter, A. J. (2014). Application of functional near-infrared spectroscopy in psychiatry. NeuroImage, 85, 478-488.

Elo, A. E. (1978). The rating of chessplayers, past and present: Arco Pub.

Emerson, R. W., & Cantlon, J. F. (2015). Continuity and change in children's longitudinal neural responses to numbers. Developmental science, 18(2), 314-326.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. Computers & Education, 67, 156-167.

Esser, G., Wyschkon, A., & Schmidt, M. H. (2002). Was wird aus Achtjährigen mit einer Lese-und Rechtschreibstörung. Zeitschrift für Klinische Psychologie und Psychotherapie; Forschung und Praxis, 31(4), 235-242.

Evans, T. M., Kochalka, J., Ngoon, T. J., Wu, S. S., Qin, S., Battista, C., & Menon, V. (2015). Brain Structural Integrity and Intrinsic Functional Connectivity Forecast 6 Year Longitudinal Growth in Children's Numerical Abilities. The Journal of Neuroscience, 35(33), 11743-11750.

Fischer, G. H. (1995). The linear logistic test model Rasch models (pp. 131–155). New York, NY: Springer.

Fleischer, J., Leutner, D., & Klieme, E. (2012). Modellierung von Kompetenzen im Bereich der Bildung: eine psychologische Perspektive. Psychologische Rundschau.

Gabrieli, J. D. (2016). The promise of educational neuroscience: Comment on Bowers (2016).

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. Simulation & gaming, 33(4), 441-467.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cerebral Cortex, 7(4), 374-385.

Giorgidze, G., Grust, T., Schreiber, T., & Weijers, J. (2010). Haskell boards the Ferry. Paper presented at the Symposium on Implementation and Application of Functional Languages.

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. Neuropsychologia, 47(2), 604-608.

Grabner, R. H., Ansari, D., Reishofer, G., Stern, E., Ebner, F., & Neuper, C. (2007). Individual differences in mathematical competence predict parietal brain activation during mental calculation. NeuroImage, 38(2), 346-356.

Grabner, R. H., & De Smedt, B. (2012). Oscillatory EEG correlates of arithmetic strategies: a training study. Frontiers in psychology, 3.

Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., & Neuper, C. (2009). Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. Human Brain Mapping, 30(9), 2936-2952.

Gross, J. (2006). The long term costs of literacy difficulties: KPMG Foundation.

Gross, J., Hudson, C., & Price, D. (2009). The long term costs of numeracy difficulties. Every Child a Chance Trust and KPMG, London.

Haeussinger, F. B., Dresler, T., Heinzel, S., Schecklmann, M., Fallgatter, A. J., & Ehlis, A.-C. (2014). Reconstructing functional near-infrared spectroscopy (fNIRS) signals impaired by extra-cranial confounds: an easy-to-use filter method. NeuroImage, 95, 69-79.

Hanushek, E. A., & Woessmann, L. (2010). The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes: ERIC.

Hartig, J., & Frey, A. (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. Psychologische Rundschau.

Harvey, B., Klein, B., Petridou, N., & Dumoulin, S. (2013). Topographic representation of numerosity in the human parietal cortex. Science, 341(6150), 1123-1126.

Hinault, T., & Lemaire, P. (2016). What does EEG tell us about arithmetic strategies? A review. International Journal of Psychophysiology.

Howard-Jones, P., Varma, S., Ansari, D., Butterworth, B., De Smedt, B., Goswami, U., . . . Thomas, M. (2016). The Principles and Practices of Educational Neuroscience: Commentary on Bowers (2016). Psychological Review.

Huber, S., Fischer, U., Moeller, K., & Nuerk, H.-C. (2013). On the interrelation of multiplication and division in secondary school children.

Huber, S., Moeller, K., & Nuerk, H.-C. (2012). Differentielle Entwicklung arithmetischer Fähigkeiten nach der Grundschule: Manche Schere öffnet und schließt sich wieder. Lernen und Lernstörungen.

Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., & Delazer, M. (2007). Imaging early practice effects in arithmetic. NeuroImage, 36(3), 993-1003.

Ischebeck, A., Zamarian, L., Schocke, M., & Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic—An fMRI study. NeuroImage, 44(3), 1103-1112.

Ischebeck, A., Zamarian, L., Siedentopf, C., Koppelstätter, F., Benke, T., Felber, S., & Delazer, M. (2006). How specifically do we learn? Imaging the learning of multiplication and subtraction. NeuroImage, 30(4), 1365-1375.

Ise, E., & Schulte-Körne, G. (2010). Spelling deficits in dyslexia: evaluation of an orthographic spelling training. Annals of dyslexia, 60(1), 18-39.

ISFE. (2012). Videogames in Europe: Consumer study. European summary report: Brussels: ISFE.

Jung, S., Roesch, S., Huber, S., Heller, J., Grust, T., Nuerk, H., & Moeller, K. (2015). An Interactive Web-Based Learning Platform for Arithmetic and Orthography. Paper presented at the Advances in Computers and Technology for Education–Proceedings of the 11th International Conference on Educational Technologies.

Kast, M., Baschera, G.-M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. Annals of dyslexia, 61(2), 177-200.

Kaufmann, L., Wood, G., Rubinsten, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. Developmental Neuropsychology, 36(6), 763-787.

Kawashima, R., Taira, M., Okita, K., Inoue, K., Tajima, N., Yoshida, H., . . . Fukuda, H. (2004). A functional MRI study of simple arithmetic—a comparison between children and adults. Cognitive Brain Research, 18(3), 227-233.

Klein, E., Suchan, J., Moeller, K., Karnath, H.-O., Knops, A., Wood, G., . . . Willmes, K. (2016). Considering structural connectivity in the triple code model of numerical cognition: differential connectivity for magnitude processing and arithmetic facts. Brain Structure and Function, 221, 979-995.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain research reviews, 29(2), 169-195.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. Brain research reviews, 53(1), 63-88.

Klimesch, W., Vogt, F., & Doppelmayr, M. (1999). Interindividual differences in alpha and theta power reflect memory performance. Intelligence, 27(4), 347-362.

Kucian, K., von Aster, M., Loenneker, T., Dietrich, T., & Martin, E. (2008). Development of neural networks for exact and approximate calculation: A FMRI study. Developmental Neuropsychology, 33(4), 447-473.

Landerl, K. (2006). Reading acquisition in different orthographies: Evidence from direct comparisons. Handbook of orthography and literacy, 513-530.

Lemaire, P. (2016). Cognitive aging: The role of strategies: Psychology Press.

Lemaire, P., Barrett, S. E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. Journal of Experimental Child Psychology, 57(2), 224-258.

Link, T., Huber, S., Nuerk, H.-C., & Moeller, K. (2014). Unbounding the mental number line- new evidence on children's spatial representation of numbers.

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H.-C. (2013). Walk the number line- An embodied training of numerical concepts. Trends in Neuroscience and Education, 2(2), 74–84.

Link, T., Nuerk, H.-C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. The quarterly journal of experimental psychology, 67(8), 1597-1613.

Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games, 4(2).

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2015). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. Computers in Human Behavior.

Menon, V., Rivera, S. M., White, C. D., Glover, G. H., & Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. NeuroImage, 12(4), 357-365.

Mizuhara, H., & Yamaguchi, Y. (2007). Human cortical circuits for central executive function emerge by theta phase synchronization. NeuroImage, 36(1), 232-244.

Moeller, K., Fischer, U., Nuerk, H.-C., & Cress, U. (2015). Computers in mathematics education–Training the mental number line. Computers in Human Behavior, 48, 597-607.

Moeller, K., Klein, E., & Nuerk, H. C. (2011a). (No) Small adults: Children's processing of carry addition problems. Developmental Neuropsychology, 36(6), 702–720.

Moeller, K., Klein, E., & Nuerk, H. C. (2011b). Three processes underlying the carry effect in addition—Evidence from eye tracking. British Journal of Psychology, 102(3), 623–645.

Moeller, K., Wood, G., Doppelmayr, M., & Nuerk, H.-C. (2010). Oscillatory EEG correlates of an implicit activation of multiplication facts in the number bisection task. Brain research, 1320, 85-94.

Ninaus, M., Kiili, K., McMullen, J., & Moeller, K. (2016). A Game-Based Approach to Examining Students' Conceptual Knowledge of Fractions Games and Learning Alliance (pp. 37-49): Springer.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Neuper, C., & Wood, G. (2015). Game elements improve performance in a working memory training task. Int. J. Ser. Games, 2, 3-16.

Ninaus, M., Witte, M., Kober, S. E., Friedrich, E. V., Kurzmann, J., Hartsuiker, E., . . . Wood, G. (2013). Neurofeedback and Serious Games. Psychology, Pedagogy, and Assessment in Serious Games, 82.

Nuerk, H.-C., Kaufmann, L., Zoppoth, S., & Willmes, K. (2004). On the development of the mental number line: More, less, or never holistic with increasing age? Developmental Psychology, 40(6), 1199.

Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., & Fischer, M. H. (2011). Extending the mental number line: A review of multi-digit number processing. Zeitschrift Für Psychologie-Journal of Psychology, 219(1), 3–22.

Nuerk, H.-C., Moeller, K., & Willmes, K. (2015). Multi-digits numerical understanding. In R. CohenKadosh & A. Dowker (Eds.), Oxford handbook of mathematical cognition (pp. 106– 139). Oxford: Oxford University Press.

Núñez-Peña, M. I., & Suárez-Pellicioni, M. (2012). Processing false solutions in additions: differences between high-and lower-skilled arithmetic problem-solvers. Experimental brain research, 218(4), 655-663.

Obersteiner, A., Dresler, T., Reiss, K., Vogel, A. C. M., Pekrun, R., & Fallgatter, A. J. (2010). Bringing brain imaging to the school to assess arithmetic problem solving: chances and limitations in combining educational and neuroscientific research. ZDM, 42(6), 541-554.

Parsons, S., & Bynner, J. (2005). Does numeracy matter more?

Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, N., Rickard, T. C., Yaroush, R. A., & Bourne, L. E. (1994). Brain potentials during mental arithmetic: effects of extensive practice and problem difficulty. Cognitive Brain Research, 2(1), 21-29.

Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. Vision research, 41(10), 1257-1260.

Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. Electroencephalography and Clinical Neurophysiology, 42(6), 817-826.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology, 110(11), 1842-1857.

Poldrack, R. A. (2000). Imaging brain plasticity: conceptual and methodological issues—a theoretical review. NeuroImage, 12(1), 1-13.

Prado, J., Mutreja, R., & Booth, J. R. (2014). Developmental dissociation in the neural responses to simple multiplication and subtraction problems. Developmental science, 17(4), 537-552.

Prins, P. J., Dovis, S., Ponsioen, A., Ten Brink, E., & Van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? Cyberpsychology, behavior, and social networking, 14(3), 115-122.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. Review of general psychology, 14(2), 154.

Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical functional reorganization underlies children's cognitive development. Nature neuroscience, 17(9), 1263-1269.

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes.

Rivera, S. M., Reiss, A., Eckert, M. A., & Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. Cerebral Cortex, 15(11), 1779-1790.

Roesch, S., Jung, S., Huber, S., Artemenko, C., Bahnmueller, J., Heller, J., . . . Moeller, K. (2016). Training arithmetic and orthography on a web-based and socially-interactive learning platform. International Journal of Education and Information Technologies, 10, 204-217.

Sammer, G., Blecker, C., Gebhardt, H., Bischoff, M., Stark, R., Morgen, K., & Vaitl, D. (2007). Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload. Human Brain Mapping, 28(8), 793-803.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Pavia, J. M., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. NeuroImage, 85, 6-27.

Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A.-C., & Nuerk, H.-C. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: A simultaneous fNIRS-EEG study. Cognitive, Affective, & Behavioral Neuroscience, 1-13.

Spüler, M., Krumpe, T., Walter, C., Scharinger, C., Rosenstiel, W., & Gerjets, P. (2017). Brain-computer interfaces for educational applications. In J. Buder & F. W. Hesse (Eds.), Informational environments: Effects of use, effective designs (pp. 177–201). New York, NY: Springer.

Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosenberg-Lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proceedings of the National Academy of Sciences, 110(20), 8230-8235.

Verner, M., Herrmann, M. J., Troche, S. J., Roebers, C. M., & Rammsayer, T. H. (2013). Cortical oxygen consumption in mental arithmetic as a function of task difficulty: a near-infrared spectroscopy approach. Frontiers in human neuroscience, 7, 217.

Wabnitz, H., Moeller, M., Liebert, A., Obrig, H., Steinbrink, J., & Macdonald, R. (2010). Time-resolved near-infrared spectroscopy and imaging of the adult human brain Oxygen Transport to Tissue XXXI (pp. 143-148): Springer.

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. Developmental Psychology, 44(2), 588

Wilson, A., Dehaene, S., Pinel, P., Revkin, S., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. Behavioral and brain functions, 2, 19.

Wilson, A., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of" The Number Race", an adaptive computer game for remediation of dyscalculia. Behavioral and brain functions, 2(1), 1.

Zamarian, L., Ischebeck, A., & Delazer, M. (2009). Neuroscience of learning arithmetic—evidence from brain imaging studies. Neuroscience & Biobehavioral Reviews, 33(6), 909-925.

Zhou, X., Booth, J. R., Lu, J., Zhao, H., Butterworth, B., Chen, C., & Dong, Q. (2011). Age-independent and age-dependent neural substrate for single-digit multiplication and addition arithmetic problems. Developmental Neuropsychology, 36(3), 338-352.

Zurstiege, G., Zipfel, S., Ort, A., Mack, I., Meitz, T. G. K., & Schäffeler, N. (2017). Managing obesity prevention using digital media—A double-sided approach. In J. Buder & F. W. Hesse (Eds.), Informational environments: Effects of use, effective designs (pp. 97–123). New York, NY: Springer.

PART III: GENERAL DISCUSSION

General discussion

The general discussion is organized into four parts: First, major results from the enclosed studies are summarized in terms of their contribution to the development of a theoretical model integrating writing of alphabetic and numerical scripts (Chapter 6). Second, the model framework is described: starting from writing letters and words (Chapter 7.1), writing digits and multi-digit numbers is integrated (Chapter 7.2). Subsequently, writing of both scripts is extended to various writing modes (i.e., handwriting and typing, Chapter 7.3). Third, the integrated and extended writing model is then evaluated by means of whether the model can explain findings from the studies enclosed in this thesis as well as findings from previous studies (Chapter 8). For this purpose, domain-specific (i.e., central) writing aspects are discussed separately in each domain (i.e., language and number domain, Chapter 8.1). Subsequently, domain-general (i.e., graphomotor execution) writing aspects are evaluated which are common to writing across domains (Chapter 8.2). The general discussion concludes with suggestions for further research that might help to evaluate cognitive mechanisms underlying writing, and hence, to adapt to the continuous evolution of writing (Chapter 9).

6 Executive summary

The aim of the present thesis is to identify the cognitive mechanisms of writing letters and numbers by jointly considering domain-specific cognitive functions of writing and domain-general graphomotor execution of writing with different writing technologies (i.e., handwriting and typing) as well as their anatomical support in the brain. This objective requires the development of a novel neuro-cognitive model toward writing which integrates writing of both scripts at different cognitive processing levels and writing modes. The model development is based predominantly on the study results presented in the following:

6.1 Section 1: Writing letters and digits – similar but not the same

In Studies 1 and 2 in Section 1 domain-specific writing differences and domain-general writing similarities in the cognitive processing of letters and digits were investigated on the behavioral and the neuro-cognitive level in a patient study (Study 1) and in healthy controls (Study 2). Study insights highlight the need to consider writing across domains.

Study 1 demonstrated, on a behavioral level, that it is possible to enhance handwriting performance in a chronic aphasic patient with severe peripheral agraphia. At the beginning of the writing training, handwriting (but not typing) of letters and words was severely impaired. Writing of Arabic digits and multi-digit numbers, however, was well preserved. However, number processing performance decreased with increasing numerical magnitude (i.e., problem size effects). During the training, the patient learned to associate semantic information depicting the letter's shape to single letters (e.g., a cocktail glass for Y), analogously to semantic magnitude information in Arabic digits (Dehaene & Cohen, 1995). After the training, the patient was able to write letters and words manually again by automating the retrieval of this associated semantic information. These training results argue for a successful intervention based on theoretical assumptions in different domains to promote handwriting. On a neural level, findings of Study 1 provided indications for the validity of these theoretical assumptions: First, in line with a large number of published studies (e.g., Beeson et al., 2003; Longcamp, Anton, Roth, & Velay, 2003; Planton et al., 2013; Roux, McKeeff, Grosjacques, Afonso, & Kandel, 2013), a left-hemispheric parieto-frontal network for writing letters in the brain was identified. Second, a bilateral fronto-parietal network of activation was found for writing Arabic digits (Arsalidou & Taylor, 2011; Dehaene et al., 2003), including the horizontal part of the intraparietal sulcus (hIPS). These different processing networks show differences in domain-specific writing processes at the central level.

Third, for peripheral writing processes, activation was observed in the vicinity of Exner's area (BA 6) for all writing tasks. Recently, Exner's area was found to be "the area of strongest and most reliable activity during handwriting" (Planton et al., 2013, p.10), involved in writing across domains (Eberhardt, 2010), and modes (Purcell, Turkeltaub et al., 2011). Thus, writing in both domains is based on a common brain region (i.e., Exner's area), which shows that this region is relevant for writing across domains.

Study 2 provided further empirical evidence suggesting that such commonly accessed cortex areas (i.e., Exner's area) receive domain-specific input from separate brain networks. This input is assumed to be conveyed via distinct processing pathways in the brain. Results of deterministic fiber-tracking in healthy adults of Study 2 indicated strong support for this assumption: in writing-to-dictation, writing of letters was found to be associated mostly with ventral connections encompassing the EC/EmC system, as well as the dorsal arcuate fascicle which enters into Exner's area more inferiorly (Weiller, Bormann, Saur, Musso, & Rijntjes, 2011). In writing of digits to dictation, semantic magnitude information was suggested to be transferred via the superior longitudinal fascicle (SLF II) into the superior part of Exner's area. Hence, both pathways to Exner's area were found at different locations (Catani et al., 2012; Weiller et al., 2011). In the patient, the ventral pathway was disrupted following a stroke. This disruption may have caused the writing impairment for letters. The dorsal pathway, however, was not affected and thus writing of digits may have been well preserved – explaining the observed dissociation between writing of letters and digits.

Taken together, study results argue for the consideration of domain-specific and domain-general cognitive mechanisms when developing a comprehensive theoretical model of typical writing. In this model, domain-specific mechanisms are reflected by central writing processes for which two distinct brain networks have been identified, one network for each domain. Domain-general mechanisms, however, are reflected by peripheral writing processes which share common brain areas required in the graphomotor planning and execution of writing in both domains.

6.2 Section 2: Handwriting and typing - modal influences on spelling and arithmetic

The studies in Section 2 investigated the influence of the actual writing mode on writing in typically developing children and children with dyslexia. The investigation of different writing modes, i.e., handwriting and typing, seems particularly important for this thesis: first, a theoretical explanation is still lacking as regards the difference in performance between handwriting and typing observed in the patient of Study 1. Second, a comprehensive writing model should comprise the evolution of writing toward digital media. Third, potential advantages of digital media might be identified which are assumed to be relevant for the assessment or promotion of language learning (difficulties).

Study 3 evaluated writing regarding two different aspects: first, from a holistic perspective, results did not indicate any group differences in terms of computer skills which is in line with previous studies. In particular, all children wrote more slowly on the computer keyboard than by hand (e.g., Berninger, Abbott, Augsburger, & Garcia, 2009; Connelly, Gee, & Walsh, 2007; Wollscheid, Sjaastad, & Tømte, 2016). With respect to spelling accuracy, typing increased spelling errors for typically developing children (but see Frahm, 2013) and children with dyslexia (e.g., Berninger et al., 2009) to the same extent. There was no difference between the groups regarding the mode effect. However, children with dyslexia underperformed in writing as compared to their peers in both writing modes. Apart from writing mode, lexicality (i.e., word and non-word writing) was found to affect spelling accuracy in both populations with writing of non-words specifically increasing the percentage of writing errors.

Second, from a rule-specific perspective, in which spelling was assessed separately for each spelling rule (i.e., capitalization, consonant doubling, lengthening, etc.), different results were observed: spelling accuracy differed between writing modes only for capitalization. Other spelling rules were not affected by writing mode. This finding emphasized the peculiarity of typing capital letters for which an additional 'Shift' key press is required. Moreover, additional processes related to writing (e.g., phonological awareness, rule and strategy knowledge as assessed by different spelling rules) did not differ between writing modes. Thus, writing mode seemed to predominantly affect peripheral but not central writing processes.

Regarding the comparability of test modes (i.e., computerized vs. paper-pencil approaches), Study 4 suggested comparable results from traditional and digital approaches in both the language and the number domain. These results contradicted findings from previous studies reporting mode effects for writing (e.g., Goldberg et al., 2003; Wollscheid, Sjaastad, & Tømte, 2016, for a meta-analysis), but not for mathematics (Wang, Hong Jiao, Young, Brooks, & Olson, 2007, for a meta-analysis). However, previous studies mainly investigated higher writing skills, such as text and essay writing, in students or adults. Typing skills in younger adults are very likely better developed as compared to those of children. For this reason, the mode effect might be more pronounced in adults. Moreover, higher writing skills seemed to be more sensitive to mode effects than writing at the single word level (Goldberg et al., 2003). Individually, children performed very differently. On the group level, children did about equally well in both spelling and arithmetic. Performance in these tasks correlated at a medium to high level (e.g., a higher spelling performance usually went along with a higher arithmetic performance). Similar correlations have been reported for traditional paper-pencil assessments (De Smedt et al., 2010, for further discussion; Landerl et al., 2009).

In sum, the study results of Section 2 show that writing mode does not impact linguistic and numerical processes to a different degree. Therefore, they do not substantiate differences in the application of digital media and traditional approaches in the assessment of written language. Furthermore, writing mode is found to predominantly affect peripheral but not central writing processes. Latter findings argue for differentiating between writing modes (i.e., handwriting and typing) at writing output stages in a theoretic writing model.

6.3 Section 3: Writing and beyond - assessment and training of orthography and arithmetic

The investigation of digital media applied in the diagnosis and intervention of orthography and arithmetic skills presented in Section 3 is particularly important for the present thesis because it provides cross-sectional and longitudinal data on the basis of which the integrated writing model can be evaluated.

Study 5 investigated potential advantages of digital media in the diagnostic procedure of spelling and arithmetic. In particular, an adaptive approach was presented in which the difficulty of the items presented was adjusted to individual performance level. Adaptive procedures were developed to reduce the number of test items, and thus, testing time. Accordingly, application of adaptivity for the diagnosis in both domains seemed to enable a high degree of individualization. Individualization, in turn, was assumed to prevent under- and overload during testing (Verdú, Regueras, Verdú, De Castro, & Pérez, 2008). The rule-specific approach presented in the studies of Section 2 provided the foundation to implement adaptivity in the spelling assessment. In the assessment of arithmetic skills, an operation-based approach (i.e., addition, subtraction, multiplication, and division) was implemented Building on these foundations, the adaptive diagnostic was proven to be promising also for the selection of specific learning contents.

Regarding learning and intervention, Study 6 provided empirical evidence for the successful application of digital learning games for spelling and arithmetic. These learning games were found to have different effects in both domains: for arithmetic, digital learning games were considered beneficial. But effects could not be specifically traced back to the training with the arithmetic games. In particular, training effects were observed for comparably easy tasks (see Roesch et al., 2016, for a more detailed description of the results), indicating that low-performing children achieved the greatest benefit from the training. For orthography, specific training effects were found. Orthography training improved children's spelling performance; improvements were found separately from other temporal effects. Overall, results demonstrated general applicability of digital learning games and their effectiveness in conveying orthography and arithmetic skills.

Taken together, results of the enclosed studies show that, on the one hand, writing errors in either domain (i.e., spelling or number transcoding errors) cannot be explained sufficiently if only central writing processes are considered exclusively. On the other hand, erroneous writing execution (i.e., distorted letters or digits) has to be investigated at the peripheral processing level. In the following, I will propose an integrated model of writing alphabetic and numerical scripts which is based on the study insights presented above. By proposing this model, I shed light on the domain-specific cognitive processes underlying writing of letters and numbers as well as on the universality of (hand)writing in terms of its graphomotor execution from a domain-general approach.

7 An integrated writing model for alphabetic and numerical scripts

The integrated writing model focuses on the production of graphemes and single digits, respectively, which are assembled to words and multi-digit numbers. I examine writing for two different writing modes, handwriting and keyboard typing, evoked within two different tasks, (a) in a writing-to-dictation and (b) in a copying task. Higher level writing skills as well as oral spelling output were not considered. In my model, serial modules are suggested to represent writing processes at particular processing levels. This representation allows to investigate the cognitive mechanisms of handwriting and typing by means of different research methods: on a behavioral level, for instance, analysis of spelling errors, latencies and the occurrence of linguistic effects may be informative. On a neural level, neuro-functional and neuro-anatomical data (i.e., lesion-symptom mapping) may provide further empirical evidence for the postulation of different processing routes and modules.

7.1 Theoretical foundations of the model: Writing letters and words

My integrated writing model is theoretically founded on the model of language functions by Ellis (1982, 1988; see Fig. 3.1 in the introduction part of this thesis). My model follows the classical architecture of the dual-route approach of language processing. Within this framework, basic elements from Ellis' model were adopted. However, several components were adapted to more recent research results and a new terminology was used to designate certain modules. These adaptations are indicated accordingly. Figure 7.1 depicts the first part of my model for writing of alphabetic scripts. For a concise presentation of the theoretical framework, I distinguish between three processing stages unlike Ellis in his model (1982): *input identification*, *central writing processes, and peripheral writing processes* (Bonin et al., 2015, for a similar approach)*. Input identification* (upper third of Fig. 7.1 on a light gray background) has to be differentiated from primary auditory and visual perception because it already involves early cognitive processing. *Central writing processes* (second third of Fig. 7.1 on a light orange background) are proposed to generate spelling for words and non-words; *peripheral writing processes* (last third of Fig. 7.1 on a light blue background) are proposed to convert abstract graphemes into neuro-muscular motor execution.

First, *Input identification* allows for (a) recognition and phonological encoding of the auditorily presented word in a writing-to-dictation task (i.e., verbal input given at the left part of Fig. 7.1 at the stage of input identification) and (b) for orthographic encoding of the visually presented verbal letter string in a copying task (i.e., written input given at the right part of Fig. 7.1 at the stage of input identification). During phonological and orthographic encoding, decisions are made as to whether the verbal string follows the phono-tactic rules of the mother tongue or whether it is a foreign word (Kay et al., 1996). This encoded information is kept active for processing in the phonological or orthographic input buffer.

Second, via *central writing (i.e., orthographic) processes* buffer information is conveyed either via i) the lexical-semantic (depicted with solid black arrows or solid-oranges arrows in Fig. 7.1) or ii) the sub-lexical routes (depicted with dashed arrows in Fig. 7.1).

i) In lexical-semantic processing, lexical entries in the (a) phonological or (b) orthographic input lexicon are activated. These lexical entries are then transferred to the graphemic output lexicon[11]: The graphemic output lexicon has been proposed to be particularly important during written language learning. "[L]earning the spelling of a word involves establishing an entry for it in a corner of one's memory [...] so that each time you wish to write the word in future the letter string which constitutes the word's accepted spelling is retrieved in its entirety from that lexicon" (Beattie & Ellis, 2017, p. 165).

[11] The idea of a lexicon, in which relevant entries are "looked up", emphasizes the functionality of the module in language processing. For convenience, this term is preferably used instead of the initial term 'graphemic output logogen system' (Ellis, 1982), or the more recent term of 'orthographic long-term memory' (McCloskey & Rapp, 2017; Purcell, Turkeltaub, et al., 2011). Existence of both an input and an output lexicon is assumed (but see Bonin et al., 2015, for a different approach). This assumption is based on findings in brain-injured patients with alexia and agraphia (Beeson & Rapcsak, 2015; Rapcsak et al., 2002), whose language processing was found to be selectively impaired at both input and output stages of the lexicon.

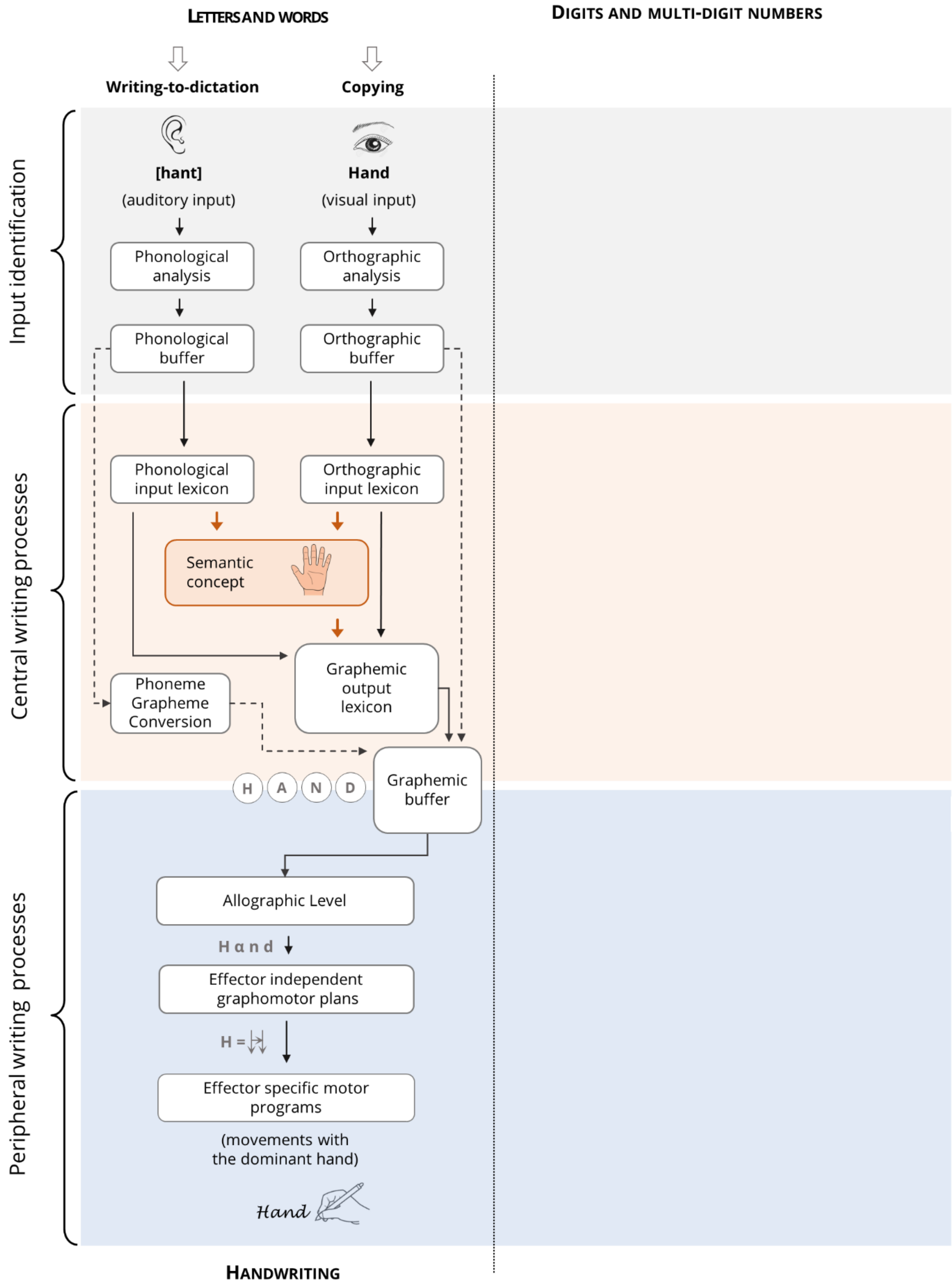


Fig. 7.1: Writing model for alphabetic scripts: Model components are given for handwriting the target word *'Hand'* within two tasks: a) writing-to-dictation following auditory presentation; b) copying the visually presented verbal string. Several processing routes are distinguished: lexical processing routes are given with solid arrows, semantic processing modules and routes are depicted in orange, and sub-lexical processing routes are indicated by dashed arrows. Curly braces indicate serial processing stages.

In the model, I assume two different ways for the retrieval of these lexical entries: first, lexical entries in the graphemic output lexicon are directly activated by receiving input from the (a) phonological or (b) orthographic input lexicon (indicated by solid arrows in Figure 7.1). Direct lexical activation is only proposed in case an external writing impulse is given (e.g., in a writing-to-dictation or in a copying task). For these tasks, direct lexical writing can also proceed incorrectly. Morton (1980) suggests that slip-of-the-pen phenomena might occur due to incorrect direct lexical processing. Second, lexical entries can also be indirectly activated by receiving input from the semantic system (orange colored module in Fig. 7.1). The semantic system comprises abstract conceptual knowledge of the world and the things in it (Saffran, 1982). Semantic processing (depicted with solid orange arrows) allows assigning the particular meaning of a word to the lexical entry. This lexical-semantic-assignment predominantly proceeds when writing is intrinsically motivated or when the retrieval of semantic information is explicitly required by a certain task (e.g., giving the meaning of a particular word). The graphemic output lexicon is then proposed to transfer the graphemic code to the graphemic buffer.

ii) In sub-lexical processing (indicated with dashed arrows in Fig. 7.1), the information stored in the phonological buffer is directly converted into the graphemic code by phoneme-grapheme assignments and then transferred to the graphemic buffer. Thereby, neither abstract lexical nor semantic information is activated.

Finally, *peripheral writing processes* (in the last third of Fig. 7.1 on a light blue background) are executed. According to Ellis (1988), I propose in my model that peripheral processes start when the graphemes are activated in the graphemic buffer. The activated graphemes are converted into specific letter shapes (i.e., allographs) in the sense of a grapheme-shape conversion at the allographic level. These allographs are then translated into effector-independent graphomotor plans. For the development of my writing model, the stage of effector-independent graphomotor plans was adopted from the work of McCloskey and Rapp (2017) and is suggested to be a step before graphomotor execution. “The graphic motor plans are effector independent in the sense that they are not tied to particular effectors (e.g., the right hand), and do not specify movements with respect to specific muscles or joints. Hence, the graphic motor plan for uppercase print could mediate writing of that letter with the right hand, left hand, left foot, or so forth” (McCloskey & Rapp, 2017, p. 69). Such an effector independent stage has not been proposed in previous models (Ellis, 1982; Van Galen, 1991) but it was found necessary to include different writing modes (i.e., handwriting and typing) in the writing model. Writing is finally executed by effector-specific motor programs which are converted into the neuro-muscular execution of the corresponding writing movements with the selected effector (i.e., writing of a certain letter on a sheet of paper with the dominant hand).

So far, the writing model covers writing of alphabetic scripts. In the following, writing of numerical scripts is integrated. Central components of writing digits and multi-digit numbers are outlined first.

7.2 Extension to the number domain: integration of numerical scripts

The extension of my writing model to the number domain is based on assumptions regarding numerical cognition postulated in both the TCM (Dehaene, 1992, 2009; Dehaene & Cohen, 1995; Dehaene et al., 2003) and the serial processing framework of the ADAPT model (Barrouillet et al., 2004). Furthermore, my model integrates multiple processing routes which have been suggested in number processing (Cipolotti & Butterworth, 1995). These models investigated either central processes of number transcoding or peripheral (hand)writing processes (Lochy et al., 2003). For this reason, they could not comprehensively describe the cognitive mechanisms underlying writing of numerical scripts. In particular, the TCM did not provide information about the serial processing stages in number transcoding. Such serial processes have been illustrated in the ADAPT model (Barrouillet et al., 2004). In the ADAPT model, however, number transcoding was only investigated in a writing-to-dictation task. Number transcoding according to visually presented numbers was assumed by the authors to proceed in a similar way. This assumption, however, has not yet been proven. By integrating central and peripheral aspects of writing numerical scripts evolved in different writing tasks in my writing model may provide the answer to the question: '[h]ow do people write numbers?' (Barrouillet et al., 2004, p. 368).

In my writing model, writing numbers is integrated within two steps: (a) in the first step, number transcoding is considered for writing-to-dictation (similar to the description of Barrouillet et al., 2004). For this purpose, central processing stages of the current writing model for alphabetic scripts are mirrored and transferred to the writing of numerical scripts. Figure 7.2 illustrates the mirroring from the language to the number domain. (b) In the second step, copying both the Arabic digit number format and the verbal number word format is integrated in the model. So far, neither the cognitive mechanism involved in copying of both number formats have yet been specified; nor has switching between these two formats been considered. For this reason, processing stages potentially required in copying Arabic multi-digit numbers and number words are adopted from the approach of copying letters and words. Figure 7.3 depicts the integration of copying of numbers (i.e., visual-Arabic digit number format) and Figure 7.4 depicts the integration of number words (i.e., verbal number word format), happening at the central processing level. Modules borrowed from Barrouillet's (2004) ADAPT model are given in boxes filled with a hatched pattern. The conceptual codes (i.e., modules) according to the TCM (Dehaene & Cohen, 1995; Dehaene et al., 2003) are marked with a double-row frame. The analogue semantic representation according to the TCM is additionally illustrated in orange. This illustration was chosen to show the origins of my integrated writing model and thus to emphasize its theoretical foundation. Furthermore, only the currently added or task relevant processing modules are given in black color, the currently not required processing stages are grayed out.

7.2.1. First part: Writing numbers to dictation (a)

Writing of digits and multi-digit numbers to dictation typically begins by auditory perception of the target number as writing alphabetic scripts. The target number is assumed to be phonologically encoded at the stage of *input identification*. Subsequent to the input identification, I propose that writing processes differ depending on the writing task: on the one hand, number transcoding is required when the auditorily presented number has to be converted into the Arabic number format (i.e., '*[ˈnɔɪ̯ntseːn]'* has to be written as '19'). Number transcoding is associated only with i) lexical-semantic processing. On the other hand, writing number words to dictation (i.e., *[ˈnɔɪ̯ntseːn]'* has to be written as *'neunzehn'*), is proposed to proceed i) lexically (i.e., for another example requiring lexical orthographic knowledge: *[ˈsi:ptseːn]* is written correctly as *'siebzehn'*) or ii) sub-lexically (i.e., *[ˈsi:ptseːn]* is written incorrectly as *'siepzeen'*). The processing routes for number transcoding and writing number words to dictation are described in more detail below.

<u>Number transcoding: i) direct lexical or lexical-semantic processing</u>

At the level of *input identification* (in the upper part of Fig. 7.2 on a light gray background), and in contrast to Barrouillet and colleagues (2004), I propose two distinct modules for phonological encoding (i.e., phonological analysis) and storage of the auditorily presented number. In line with Ellis' (1982) model, a phonological buffer is required to store phonological encodings for further processing. The phonological buffer is functionally associated with working memory capacity. The significance of working memory capacity in number transcoding is emphasized in the ADAPT model (Barrouillet et al., 2004) and led to further research particularly in children (Lopes-Silva et al., 2014; Moura et al., 2013; Zuber et al., 2009) and adults (Knops, Nuerk, Fimm, Vohn, & Willmes, 2006; Menon, 2016). In case of adults, clinical observations of number word repetition of the patient presented in Study 1 suggests the necessity of a phonological buffer at the level of input identification: when the patient was asked to repeat a multi-digit number, repetition performance decreased with increasing number of digits (i.e., problem size effect). This decrease in performance was hypothesized to be related to inappropriate working memory capacity because phonological analysis (i.e., identification of the number words as native language) was found unimpaired. In general, when unimpaired performance is observed in the phonological analysis versus poor performance in (number) word repetition, the latter deficit cannot be attributed to early receptive impairments. Therefore, different modules of phonological encoding and storage analogous to models of language processing, seemed to be also required in the number domain.

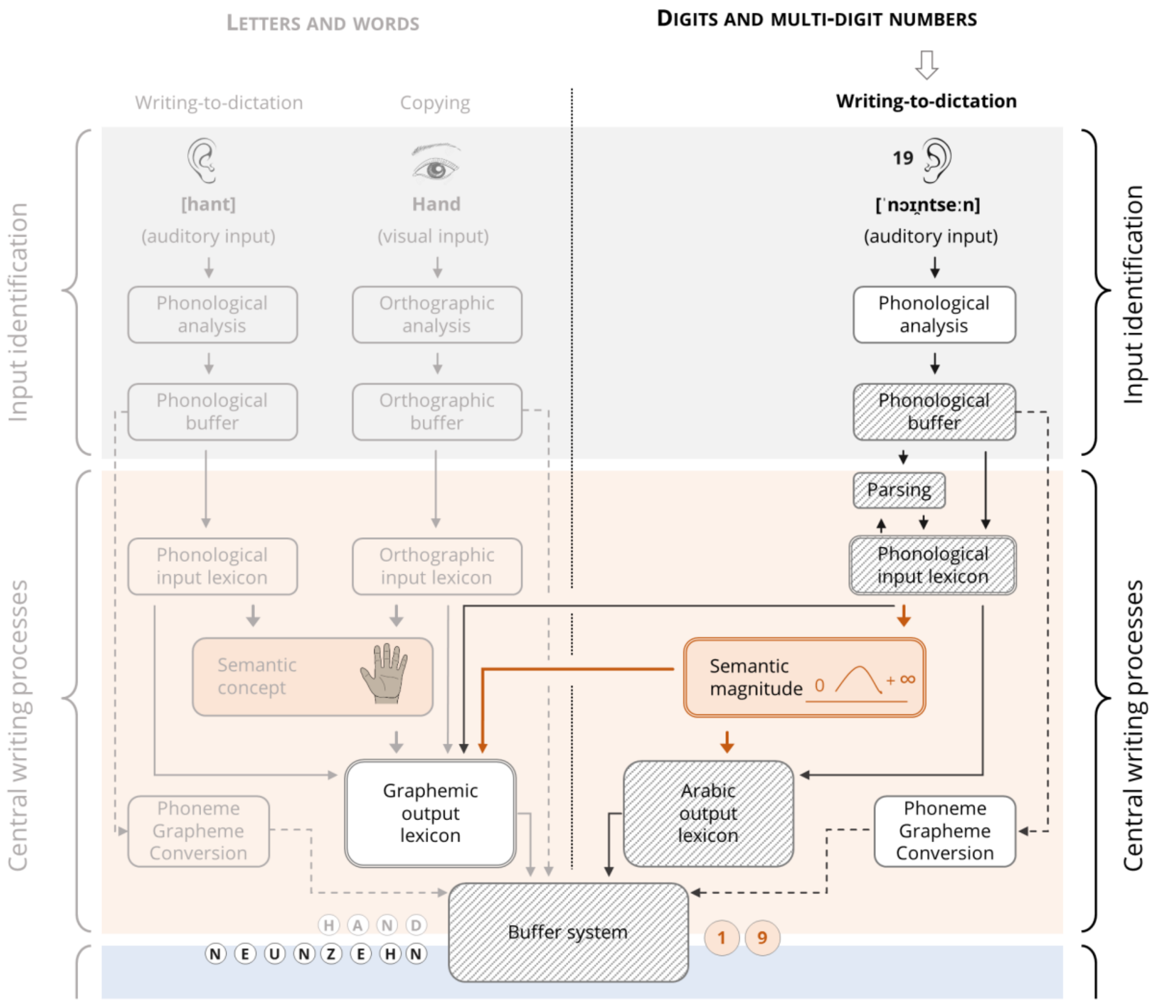


Fig. 7.2: Central writing components for writing numbers to dictation: Model components are given for number transcoding from auditory-verbal to the Arabic number format of the auditorily presented two-digit number '*19*'. Several processing routes are distinguished: number transcoding (i.e., lexical) routes are given by solid arrows, semantic processing module and routes are depicted in orange, and sub-lexical processing routes (i.e., number word writing without number transcoding) are indicated by dashed arrows. The modules corresponding to the auditory-verbal code and the analog number magnitude of Dehaene's TCM are marked with a double-row frame. The modules adopted from Barroulliet's ADAPT model (2004) are given in boxes filled with a hatched pattern. Curly braces indicate serial processing stages.

At the transition to the *central writing (i.e., number transcoding) processes* (lower part of Fig. 7.2 on a light orange background), the phonological information kept in the phonological buffer is transferred to the phonological lexicon. This transfer is proposed to be executed either directly or via a parsing module. The parsing process is adopted from the ADAPT model (Barrouillet et al., 2004, for a detailed description of the parsing process). It is required when larger multi-digit numbers are presented. During parsing, multi-digit numbers are suggested to be subdivided into smaller lexical and syntactical entries (i.e., separators or multiplicator words such as '*hundred, thousand*, *etc*.). These smaller units are then compared to the available entries in the phonological lexicon (cf. Dotan & Friedmann, 2018, for splitting multi-digit numbers into their elements at earlier processing stages in reading) in order to generate the correct sequence of digits according to the place-value structure of the Arabic number system (Zhang & Norman, 1995). Such parsing processes are domain-specific for number processing and not provided for writing alphabetic scripts. Finally, the parsed units are assumed to be reassembled, which is a prerequisite for the semantic retrieval of number magnitude information. For reassembling, I assume that the assembly is carried out also in the parsing module, in the sense of a reverse parsing process. For this purpose, I suggest bi-directional processing between the parsing module and the phonological lexicon indicated by a descending and an ascending solid black arrow in Fig. 7.2. Activation of the whole verbal number word form in the phonological lexicon is a pre-requisite for the semantic retrieval of number magnitude information.

Number magnitude information is retrieved from the semantic system. In line with the more recent versions of the TCM (Dehaene, 2009; Dehaene et al., 2003), the internal representations held in the semantic system are assumed to be spatially organized. Spatial organization is reflected by the assumption of a mental number line. According to Dehaene (1995) multi-digit numbers are assumed to be represented holistically (i.e., as in integrated entity) in order to place them correctly upon the mental number line according to their quantity. However, mental representations have also been suggested to be decomposed in nature maintaining the place-value structure of the Arabic system (Moeller, Nuerk, & Willmes, 2009). In order to meet these assumptions, I assume in my integrated writing model isolation and re-assembly of the lexical and syntactic units to be obligatory to activate the correct abstract semantic number magnitude. Subsequently, the meaningful abstract information is transferred to the Arabic digit output lexicon (i.e., lexical-semantic processing).

However, in the ADAPT model, asemantic processing (i.e., bypassing semantic) is proposed, in which lexical units are directly transcoded into Arabic digits. For this purpose, one module was postulated to be responsible for lexical access of long-term memory (Barroulliet et al., 2004), another module, namely a production system, was proposed following specific rules (P1 – P4) in number transcoding. My model also suggests two lexical modules in number transcoding: the phonological input lexicon is assumed to process the lexical-syntactic information of the phonologically encoded number (e.g. inversions). The processed language-dependent phonological information is then directly transferred to the Arabic output lexicon (i.e., direct-lexical processing). In the Arabic output lexicon, the respective lexical-syntactic frame of the Arabic digit is assigned (see Barrouillet et al., 2004, for a detailed description of the assignment from the verbal to the Arabic number system). The transcoded information is forwarded to a buffer system.

Writing number words to dictation: i) direct-lecxical, lexical-semantic, or ii) sub-lexical processing

Writing number words to dictation (i.e., without any transcoding process) follows either i) lexical or ii) sub-lexical processing routes. At the level of *input identification*, perceptual processes do not differ from those of number transcoding.

During *central writing (i.e., orthographic) processes*, lexical entries in the phonological input lexicon are transferred to the graphemic output lexicon. At this point it becomes apparent for the first time that graphemic representations of both words and verbal number word format are assumed to be stored in the same output lexicon. Thus, my model assumes cross-domain processing. In Figure 7.2 this is indicated by the crossing over the middle line which separated both domains. This crossing can be carried out either by access to (i.e., lexical-semantic processing, indicated in orange in Figure 7.2) or bypassing the semantic number system (i.e., direct-lexical processing, indicated in black solid arrows in Figure 7.2).

i) In lexical processing, the graphemic output lexicon is proposed to accomplish for irregular letter to sound assembly (i.e., *[si:ptsen]* – '*siebzehn'*) of number words similar to writing of alphabetic script. This lexical information is a prerequisite for writing irregular number words correctly. The connection between the phonological input lexicon and the graphemic output lexicon resembles the auditory-verbal code in the TCM by Dehaene (1992) and co-workers (1995, 2003, 2009, 2018). Accordingly, both modules are indicated with a double-row frame in Figure 7.2.

ii) Sub-lexical processing is also possible. In sub-lexical processing, lexical entries activated in the phonological input lexicon are converted into the graphemic code by phoneme-grapheme assignment (Damian, 2004). In this vein, conversion errors can occur in which the phonetic structure of the number word is exactly assigned to the associated graphemes (i.e., for German: '*siepzeen'* instead of *'siebzehn'*). Sub-lexical writing was not considered in the ADAPT model, probably, because it does not require number transcoding. Nevertheless, it is conceivable that a number word is written purely on the basis of its phonetic structure (e.g., when learning numbers in a foreign language, for instance, '*soixante-neuf'* in French).

For both scripts, central processing is suggested to terminate when the information converges in the (graphemic) buffer system. However, the question arises as to the structure of the buffer system. Should my model assume two separate working memory modules, this is, a buffer system for each domain, or a common working memory module for both domains? So far, research has not provided an unequivocal answer to this question: neuroimaging studies have identified a common network of brain regions involved in various working memory tasks (Cabeza & Nyberg, 2000, for a review). However, also domain-specific activation has been reported within this working memory network for letters and digits (e.g., Knops, Nuerk, Fimm, Vohn, & Willmes, 2006; Libertus, Brannon, & Pelphrey, 2009). Due to these inconsistent findings, I propose a shared multimodal buffer system in line with Baddeley's model of working memory (Baddeley, 2000). Hence, the term graphemic buffer is avoided, instead I will use the term buffer system in the following. To emphasize the domain-general functionality of the buffer system for writing letters and numbers, I placed the module directly at the center on the bottom of Figure 7.2.

7.2.2. Second part: Copying of both number formats (b)

As visual stimuli, letters and digits are quite similar, however, similarities and differences between letter and digit identification have been described (Schubert, 2017). One decisive difference is the occurrence of numerical scripts in different number formats (i.e., the visual Arabic number format as well as the [visual-]verbal number format). Copying of both formats is proposed to require different processing routes. Figures 7.3 and 7.4 depict the relevant central processing stages of copying both number formats. It is worth mentioning that, so far, copying of multi-digit numbers and number words has not yet been addressed in a serial model framework. Although the TCM (Dehaene & Cohen, 1995; Dehaene et al., 2003) has provided a visual-Arabic code, the internal processes of number transcoding within this code have not yet been specified. The serial processing ADAPT model (Barrouillet et al., 2004) could not provide further insights on this issue. I will integrate copying of Arabic multi-digit numbers in the writing model as follows:

<u>Copying of Arabic multi-digit numbers: i) direct-lexical or lexical-semantic processing</u>

Copying of Arabic multi-digit numbers typically begins with *input identification* (upper part in Fig. 7.3 on a light gray background), in which the number presented is visually encoded (to be a number). For visual encoding, I propose a domain-specific module for visual analysis of Arabic multi-digit numbers. Domain-specific separation for orthographic and visual encoding of alphabetic and numerical scripts is assumed because of the different structural properties between the scripts: “the decimal structure of digits is completely different from the morphological structure of words. Consequently, a dedicated visual analysis process that extracts the morphological structure of words could be very different from a dedicated visual analysis process that extracts the decimal structure of numbers" (Dotan & Friedmann, 2018b, p. 22). Subsequently, the encoded visual string is transferred to and stored in the visual buffer. Task- and domain-specific buffer systems at the level of input identification for words and numbers might explain the occurrence of dissociations in reading performance between different numerical scripts (i.e., preserved reading of Arabic digits but impaired reading of the verbal number format, Schubert, 2017) and domains.

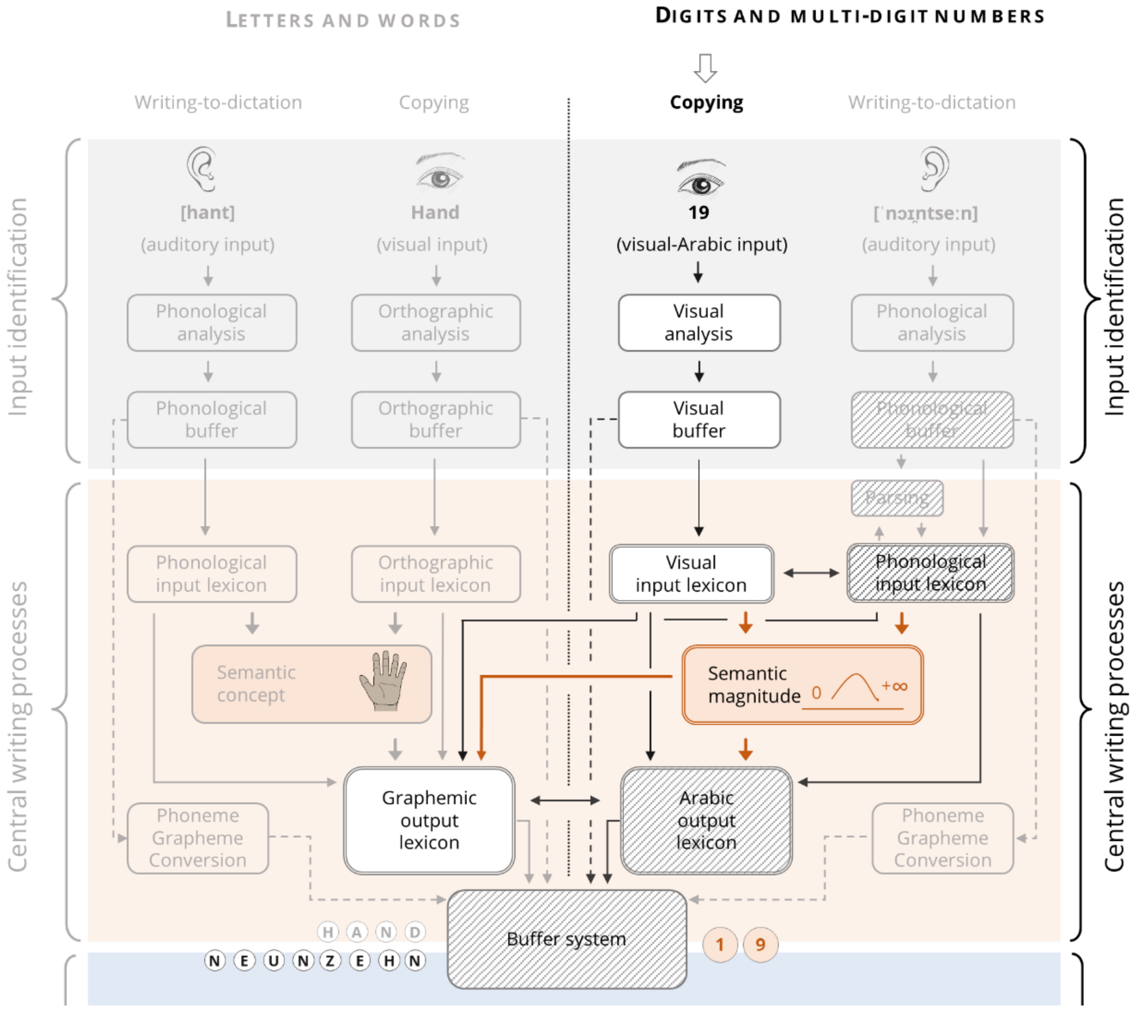


Fig. 7.3: Central writing components for copying and transcoding Arabic multi-digit numbers: Model components are given for copying Arabic digits and number transcoding following the visually presented two-digit number '19'. Several processing routes are distinguished: copying via direct-lexical processing and number transcoding are given by arrows, number transcoding by retrieval of semantic information is depicted in orange, and sub-lexical processing is indicated by dashed arrows. The modules corresponding to the auditory-verbal code and the analog number magnitude of Dehaene's TCM are marked with a double-row frame. The modules adopted from Barroulliet's ADAPT model (2004) are given in boxes filled with a hatched pattern. Curly braces indicate serial processing stages.

At the central processing level (lower part of Fig. 7.3 on a light orange background), copying of Arabic multi-digit numbers involves either i) lexical or ii) sub-lexical processing, analogous to copying letters or words. i) In lexical processing, I assume that the visual buffer information is transferred by the visual input lexicon to the Arabic output lexicon. This transfer is proposed to resemble the visual-Arabic code in the TCM. Hence, the modules for the visual input lexicon and the Arabic output lexicon are marked with a double-row frame in Figure 7.3. For this transfer, the visually encoded digit string is considered as unique and identifiable because of the place-value structure of the Arabic number system (Zhang & Norman, 1995). The activated visual digit string in the visual input lexicon might trigger retrieval of semantic numerical quantity (Dehaene & Cohen, 1995; Dehaene et al., 2003). However, the visually presented Arabic multi-digit number can also be transcoded into the respective number word. Transcoding necessarily involves activation of the visual input lexicon and the graphemic output lexicon; it can potentially involve retrieval of semantic number magnitude. For number transcoding, retrieval of phonological information is required in order to assign the decimal structure correctly to the verbal number format (i.e., inversions, Fayol & Seron, 2005). Hence, I propose bi-directional connections between the visual input lexicon and the phonological input lexicon in my writing model. Once the visual input lexicon has activated the phonological input lexicon, activation is transferred to the graphemic output lexicon by the already established routes for writing number words to dictation. Bi-directional transfer from the Arabic input lexicon to the graphemic output lexicon is theoretically also possible.

ii) Theoretically, sub-lexical visual copying may be assumed, but it is rather artificial. Sub-lexical processing is considered to occur only when a multi-digit number is copied directly from the visual input buffer to the output buffer system (dashed arrows in Figure 7.3) without being internally processed or verbalized. Verbalization would require access to phonological information, and thus, bi-directional access from the visual to the phonological lexicon. This theoretical assumption suggest that central writing and transcoding processes cascade into further peripheral writing processing (Kandel & Perret, 2015).

<u>Copying of number words: i) direct-lexical, lexical-semantic, or ii) sub-lexical processing</u>

The verbal number format is copied in the same way as alphabetic scripts (Damian, 2004). For this reason, common processing modules are proposed for *input identification* (see Figure 3.1 and section 3.2.1 for a detailed description).

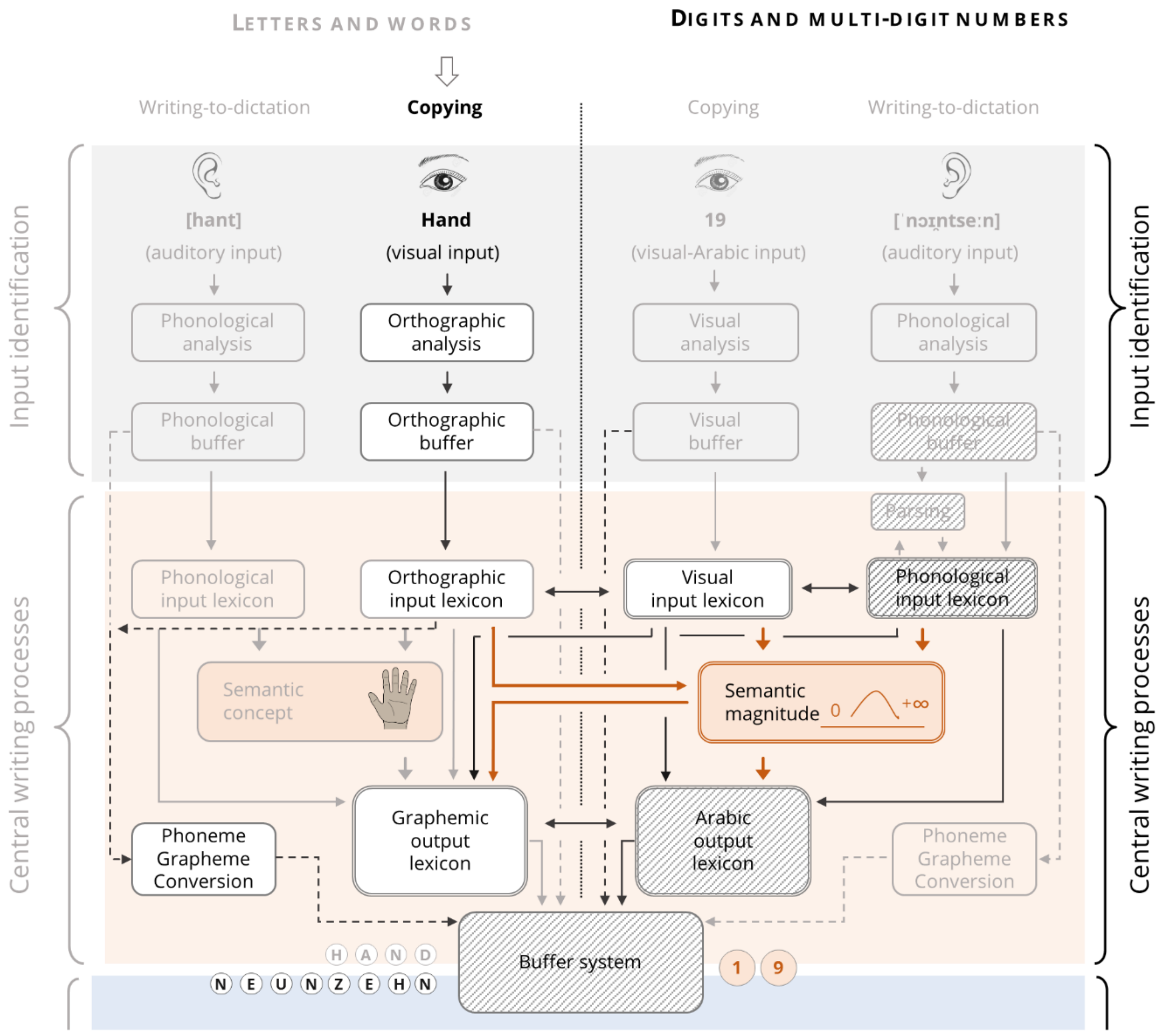


Fig. 7.4: Central writing components for copying and transcoding number words: Model components are given for copying Arabic digits and number transcoding following the visually presented number word 'neunzehn'. Several processing routes are distinguished: copying via direct-lexical processing and number transcoding are given by solid arrows, number transcoding by retrieval of semantic information is depicted in orange, and sub-lexical processing are indicated by dashed arrows. The modules corresponding to the auditory-verbal code and the analog number magnitude of Dehaene's TCM are marked with a double-row frame. The modules adopted from Barroulliet's ADAPT model (2004) are given in boxes filled with a hatched pattern. Curly braces indicate serial processing stages.

At the *central processing level* (lower part in Fig. 7.4 on a light orange background), again i) lexical or ii) sub-lexical processing is possible: i) in lexical-processing, I assume that the entries in the orthographic input lexicon are transferred to the graphemic output lexicon. It is likely that they also access the semantic system (Coltheart et al., 2001). For number transcoding from the verbal number word format to the Arabic digit number format, direct lexical or lexical semantic processing is possible. Direct lexical processing requires bi-directional connections, which cross the border of both domains, between the orthographic input and the visual input lexicon for the number domain (given in solid arrows in Figure 7.4). ii) In sub-lexical processing, information of the orthographic input buffer is forwarded via grapheme-phoneme-conversion into the graphemic buffer. Although this sub-lexical processing is again artificial, it might account for spelling errors occurring in irregular letter-to-sound mappings (i.e., '*siebzehn*' – '*[si:ptse:n]*' – '*siepzeen*') in a copying task, when the stimulus is only shortly presented and the phonologically lexicon is reactivated.

Taken together, the integration of central writing components of numerical scripts into my current writing model seems theoretically feasible. However, each domain is shown to have its own peculiarities. It the following part, I will extend the integrated writing model to different writing media focusing on peripheral components of writing in both domains.

7.3 Extensions to digital media: Writing toward the future

The extension of my integrated writing model to peripheral components of writing alphabetic and numerical scripts is based on the findings of Studies 1 and 2. Model extensions regarding different writing modes (i.e., handwriting and typing) are inspired by results of Studies 3 and 4. Beyond that, model extensions also originated from recent findings on language and number processes involved in writing in either domain. Figure 7.5 depicts the peripheral writing components of handwriting and typing. Figure 7.6 depicts peripheral writing components of alphabetic and numerical scripts. Again, earlier cognitive processes as well as non-required modules are grayed out in these figures.

Generally, similar linguistic representations and neural substrates are assumed for peripheral writing components, whether written by hand or typed with the computer keyboard (Higashiyama, Takeda, Someya, Kuroiwa, & Tanaka, 2015; Purcell, Turkeltaub, et al., 2011). Even though the two writing modes rely on different writing abilities (Mangen & Velay, 2010), they nevertheless share similar, mode specific graphomotor and hand motor programs when writing letters and digits (Lochy et al., 2003). As an example, let us compare handwriting of the letter *'S'* with the digit '*8'*: for both signs almost identical curved strokes are needed. Comparing the letter '*I*' and the digit *'1'*, a vertical stroke is required in writing the letter '*I';* for writing the digit *'1'*, the vertical line has to be joined to a short oblique stroke. Strictly considering the graphomotor plans, one can assume that the movement for the vertical stroke in the letter '*I'* does not deviate from the vertical line in the digit *'1'*. Hand motor programs for typing letters and digits can also be considered to be more or less equal. In both scripts, fingers are assigned to the corresponding keys by spatial memory traces (Ardila, 2012). For example, looking at the typing of an experienced typist, it can be noticed the keys on the keyboard are pressed without always having to observe the hands. For letters, usually both hands are used. For digits, only one hand is active in case a key block for numbers is available on the keyboard. Evidence is provided that typists are able to recapitulate the order of keys pressed by indicating the individual finger (Ardila, 2012). This shows that typing is a highly automatized, embodied process, as is the case with handwriting (Mangen & Velay, 2010). Crucially, both processes are perfected by experience. In the following I will introduce how my integrated writing model extends to writing on different writing media.

7.3.1 Peripheral components of handwriting and typing

In my integrated writing model, I propose that peripheral writing processes start, once the buffer system is activated. The assumption of a (graphemic) buffer for handwriting has been well established (Ellis, 1982, 1988; Rapp, Epstein, & Tainturier, 2002) and was recently proposed for typing by Pinet and co-workers (2016). In terms of the internal structural organization of the (graphemic) buffer (i.e., single graphemes, syllables, etc.), however, the literature is still lacking consensus (Pinet et al., 2016). For the integration of numerical scripts into the current writing model, I suggested a common buffer system for graphemes and abstract digits. The domain-general functionality of the buffer system is depicted by its position directly in the center between writing letters and numbers. Beginning from the buffer system, I propose that the subsequent processes differ between writing modes: for handwriting, the graphemes are externalized into one-handed writing movements through serial processes (Ellis, 1982, 1988; Rapcsak et al., 2002). For typing, graphemes are transferred into two-handed (parallel) finger-movements (Margolin, 1984; Pinet et al., 2016). Figure 7.5 depicts peripheral components of writing for both writing modes.

To briefly summarize the peripheral processes of handwriting (for a detailed description see Section 7.1): Starting from the buffer system, I propose an allographic level to embrace the conversion of the graphemes into allographs. The allographs are then transferred into graphomotor plans which are assumed to be “effector independent in the sense that they are not tied to particular effectors (e.g., the right hand)” (McCloskey & Rapp, 2017, p. 69). The graphomotor plans, however, are simultaneously assumed to be letter-specific (Rapcsak & Beeson, 2000). Letter-specificity means that the motor commands for the letter shape are relatively robust in different writing situations. For example, writing and *‘H’* on a blackboard and on a sheet of paper requires the same strokes to note the letter ‘*H’* (i.e., two more or less parallel vertical lines and a horizontal line which connects both for printed script), but the size of writing movement differs significantly (Papathanasiou & Coppens, 2012). The graphomotor plans are then converted to effector-specific motor programs (e.g., for writing movements with the dominant hand) before being translated into neuro-muscular execution.

Unlike handwriting, I do not assume an allographic level for (Purcell, Turkeltaub et al., 2011, for a similar approach; but see Cameron, Cubelli, & Della Sala, 2002; Margolin, 1984). This assumption was primarily based on results of Study 1 reporting preserved typing but impaired handwriting. On a computer keyboard, letters are given in generic shape. To specify the letter shape, which is particularly the case in capitalization or with special letters (e.g., è, ô, etc.), relatively sophisticated motor plans are required (i.e., parallel pressing of the shift-key and the respective key for the letter intended to generate a capital letter, Ardila, 2012). Specifying letters during typing differs significantly from specifying letters during handwriting at the allographic level. Hence, I do not propose allographic processing for typing. The motor plans are suggested to also be effector-independent; again, no matter of which finger or which hand pressed these keys.

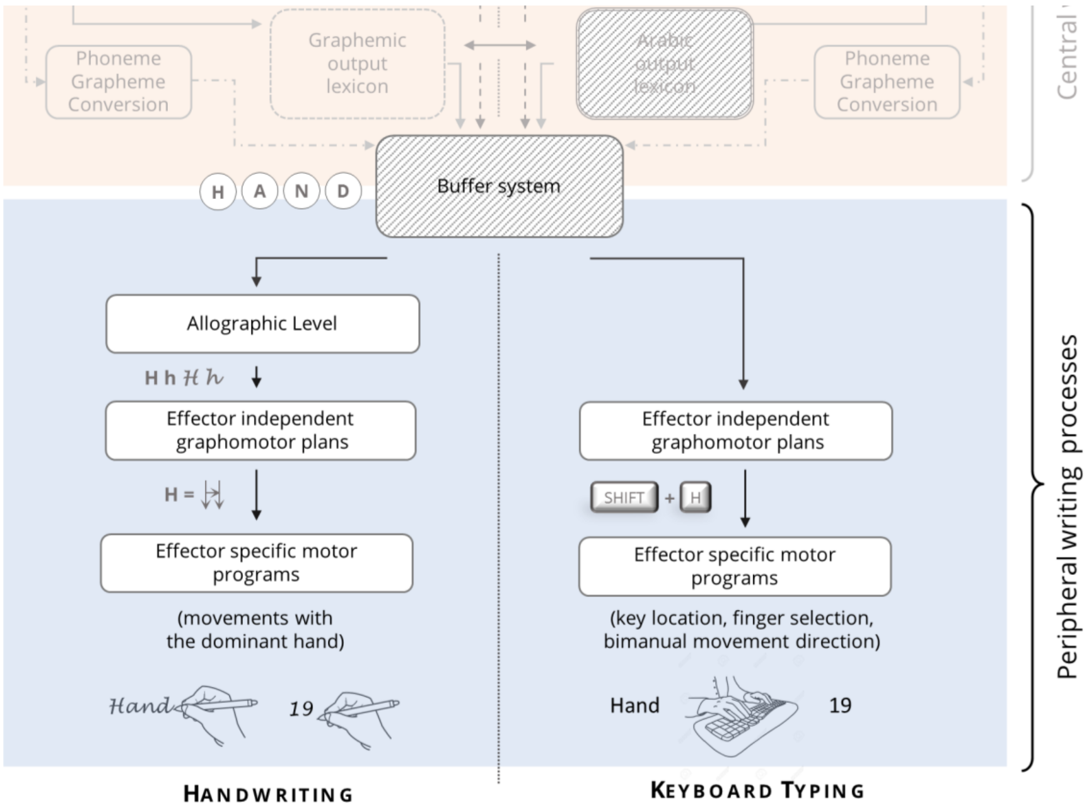


Fig. 7.5: Peripheral components of handwriting and typing: Writing of alphabetic scripts is depicted for handwriting and typing from the buffer system to neuro-muscular execution.

Finally, effector-specific graphomotor programs activate neuro-muscular execution of typing. Accordingly, typing is executed either with only one hand, or both, the index finger or the pinky, and so forth. Because handwriting and typing differ at the level of graphomotor plans and programs, distinct modules are proposed in the writing model. According to these peripheral writing components, an explanation for the dissociation between handwriting and typing of the patient in Study 1 might be the following: By clinical observation, the patient's handwriting disturbances were related to impaired processing at the allographic level making him unable to access letter forms and shapes. For typing, such allographic processing is not required.

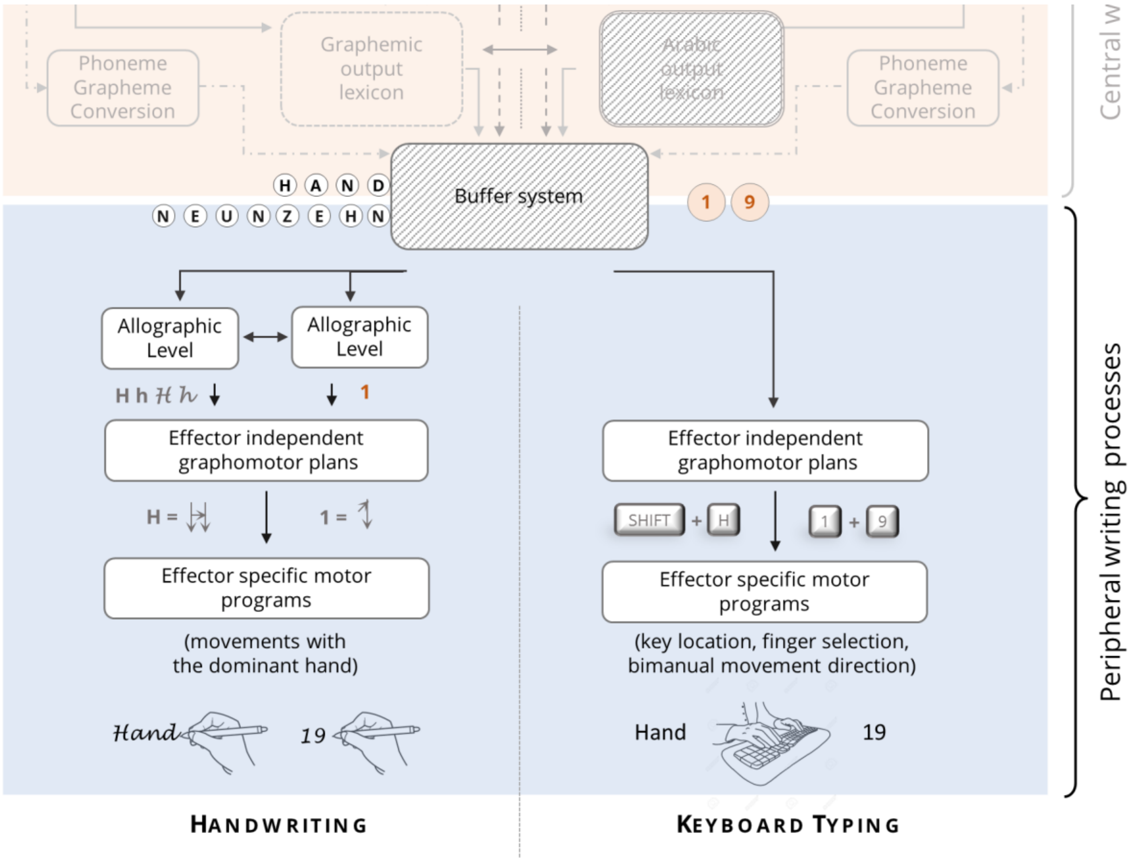
7.3.2 Peripheral writing components of letters and digits.

So far, peripheral components for writing numerical scripts have only been studied to a limited extent. In these studies, domain-general writing processes but also domain-specific differences have been reported (Lochy et al., 2003). Domain-general writing processes might affect writing in both scripts, whereas domain-specific writing differences might lead to dissociations between writing of letters and digits. In the literature, some patients have been described with a prominent dissociation between preserved digit but severely impaired or distorted letter writing (Anderson et al., 1990; Delazer et al., 2002; Tohgi et al., 1995). Possible reasons for the occurrence of such dissociations may be explained as follows considering processing in peripheral writing components.

The first major difference between both scripts can be found in the internal organization of the buffer system: In the buffer, numbers are assumed to be assembled digit-by-digit analogous to grapheme-by-grapheme assembly in words (Lochy et al., 2003). While graphemes are abstract, 'meaningless' representations of letters, representations of single digits are 'meaningful. In this context, 'meaningful' means that each individual digit carries semantic information about numerical magnitude (Beeson et al., 2003). Figure 7.6 highlights in color the differences between the meaningful and meaningless abstract representations held in the buffer (i.e., meaningful abstract semantic information for digits is depicted in orange, abstract meaningless graphemes are depicted in black). Numerical quantity indicated to a certain set size. This indication is intended to make the abstract representation of numbers more imaginable than the abstract representation of letters. Single digits numbers are assigned their semantic meaning because of the numerical magnitude, but single letters (i.e., graphemes) do not carry a semantic meaning per definition. A semantic meaning only arises when the graphemes are composed in a certain order from a word. In language studies, imageability has been found to facilitate word retrieval and lexical decision (Sereno, O'Donnell, & Sereno, 2009) and reduce working memory load (Barrett, Tugade, & Engle, 2004). Furthermore, in a working memory task (i.e., n-back task) by Knops and colleagues (2006), digits but not letters were found to automatically access semantic information. Based on these findings, domain-specific processing differences arise at the level of the buffer system. From the buffer, graphemes and digits are then directed to the allographic level, but only for handwriting.

For typing, I propose the buffer to directly activate effector-independent graphomotor plans (see again Fig. 7.5).

Fig. 7.6: Peripheral writing components of letters and digits: Writing of alphabetic and numerical scripts is depicted for handwriting and typing from the buffer system to neuro-muscular



execution. Semantic meaningful abstract representations of digits are indicated in orange.

At the allographic level, differences between letters and digits have also been described (Lochy, 2003). Unlike letters, numbers do contain less invariants in handwriting like case and style. Moreover, the set size of allographs for digits also differs from the set size for graphemes and the variation of strokes is smaller. Lochy and colleagues (2003) concluded that processing letters and digits differs at the allographic level. In line with this conclusion, I assumed two domain-specific but connected allographic modules. These allographic modules differ in the number of stored allographs and their visual properties. The allographs are assumed to activate the respective effector-independent graphic motor plans, which are further directed to effector-specific motor programs. The effector-unspecific graphomotor plans of digits were considered to be less complex as they contain less changes in direction (Lochy et al., 2003). Lower complexity of digits in graphomotor execution might also facilitate writing numerical scripts and cause dissociations between writing of letters and digits.

The transition between allographic level, motoric plans and programs was recently attributed to a circumscribed brain area, namely Exner's area, which was found activated for both handwriting and typing in various imaging studies (Higashiyama et al., 2015; Planton et al., 2013; Purcell, Turkeltaub et al., 2011; Roux, Draper, Köpke, & Démonet, 2010). The relevance of Exner's area and its connectivity within a larger writing network for letters and digits was examined in Study 2. This study demonstrated that Exner's area seems to be relevant not only for writing alphabetic scripts, but also for writing digits, corroborating earlier work from Eberhardt (2010). Based on these results, the functional relevance of Exner's area seems to be general for mode and domain, thus confirming its role at the stage of (access to) effector-unspecific plans and specific motor programs.

To sum up, writing alphabetic and numerical scripts is found to rely on the same cognitive mechanisms for both handwriting and typing at peripheral writing processes. The allographic level was found to be of particular importance: In my model I propose two distinct but connected allographic modules for handwriting in each domain, while I do not consider an allographic level for typing in both domains. In terms of a serial model assumption, the peripheral writing processes are adjoined to the central writing processes in order to provide a comprehensive integrated writing model. The empirical framework for the model was provided by the insights of the studies presented in the second part of my thesis. Figure 7.7 shows the final integrated writing model for alphabetic and numerical scripts. By integrating writing in different domains and writing modes, my extended model acknowledges the complexity of different writing contexts.

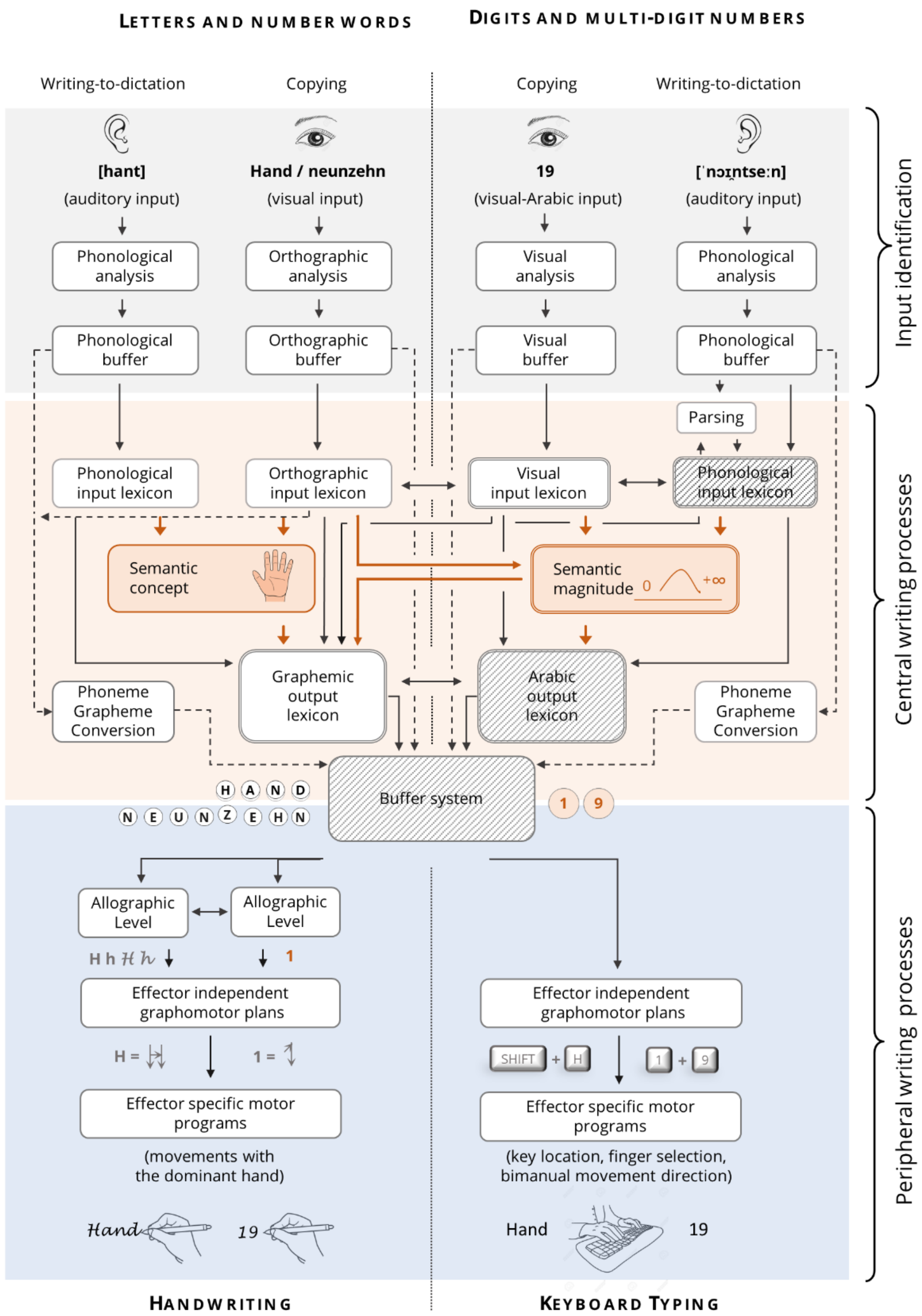


Fig. 7.7: Integrated writing model for handwriting and typing of alphabetic and numerical scripts. Central writing components: Writing-to-dictation and copying: Writing via direct-lexical processing and number transcoding are given in solid lines and arrows, semantic processing is depicted in orange, and sub-lexical processing routes are indicated by dashed arrows. The modules corresponding to the auditory-verbal code and the analog number magnitude of Dehaene's TCM are marked with a double-row frame. The modules adopted from Barroulliet's ADAPT model (2004) are given in boxes filled with a hatched pattern. Peripheral writing components: handwriting and typing is illustrated from the buffer system to neuro-muscular execution. Semantic meaningful abstract representations of digits are indicated in orange. Curly braces indicate serial processing stages.

8 Evaluation of the novel integrated writing model

The development of the integrated writing model was based on the study insights presented in the thesis as well as on empirical observations from previous literature. Both own and previous studies showed that the writing of letters and numbers may be affected simultaneously, or that a deficit may affect writing only in one domain (i.e., letters or numbers) or in one modality (i.e., handwriting or typing). Since current research could not sufficiently explain the observed writing performance by means of existing theoretical models in the respective domain, I developed an integrated writing model for writing in both scripts which considers writing from a cross-domain perspective. In the following, I will evaluate whether the model can sufficiently explain similarities and dissociations of writing in different notational systems. This evaluation is based on findings from the studies presented above and previous studies reported in the first part of this thesis (summarized in Table 2 in the introduction part of this thesis).

Evaluation of my integrated writing model starts for central followed by peripheral components of writing in order to retain the outline used in the previous paragraphs.

8.1 Central writing processes

Recent research suggested shared networks for language and arithmetic in the brain (Baldo & Dronkers, 2007). The commonality of these networks is mainly based on similar phonological processes (Andin et al., 2015). However, within these networks domain-specific processing was also observed (Peters & Smedt, 2018, for a meta-analysis). These studies lend support to the theoretical assumption of interacting domain-specific and domain-general central writing processes required to generate alphabetic and numerical scripts.

8.1.1 Input identification

At the level of *input identification*, encoding processes for alphabetic and numerical scripts are assumed to differ. Very recently, Dotan and Friedmann (2018b, 2018a) showed separate mechanisms for words and Arabic multi-digit numbers already at the level of visual analysis. The authors stated that a word consists of single letters, whereas a multi-digit number has to be converted to multiple words. Thus, the verbal number format resembles a whole phrase. Supporting evidence is provided by neuro-functional studies. In these studies, two separate processing pathways including different brain areas for word reading and reading of multi-digit numbers were described: the visual word form area (VWFA) and visual number form area (VNFA, Hannagan, Amedi, Cohen, Dehaene-Lambertz, & Dehaene, 2015, for a review). Whereas the VWFA was suggested to be connected to the language network (i.e., left temporal cortex and left inferior frontal gyrus), the VNAF has been suggested to connect to regions associated with numerical quantity processing, such as the IPS (Hannagan et al., 2015). Processing of number words in contrast, has been found to proceed also to language processes (Damian, 2004).

In sum, these findings substantiate the assumption of domain-specific modules at the level of input identification: phonological analysis for phonological encoding of (number) words, orthographic analysis for visual encoding of (number) words and visual analysis for the Arabic digit number format. Impairment of these early input processes leads to severe language disorders, such as 'pure word deafness' (Cazzolli et al., 2017; Kussmaul, 1877) or 'pure alexia' (L. Cohen, Dehaene, McCormick, Durant, & Zanker, 2016). Interestingly, Arabic digits can often be better perceived than letters. This dissociation has been attributed to lower similarity between digits (e.g., 1, 2, 4, 6) as compared to letters (e.g., for upper case letters: Z, N, M, W; for lower case letters n, m, u, v; Lochy et al., 2003; Schubert & McCloskey, 2013).

8.1.2 Lexical and sub-lexical processing and their role in writing

Lexical-(semantic) and sub-lexical processing are envisaged in the integrated writing model for writing of alphabetic and numerical scripts. For the language domain, evidence for both processing routes arose from previous studies investigating linguistic effects (i.e., regularity effects: Beauvois & Dérouesné, 1981; Dehaene, 1991; lexical and frequeny effects: Shallice, 1981), further substantiated in Study 3. In this study, a rule-specific approach was applied in studying spelling performance. By this approach, lexical orthographic knowledge could be investigated separately from sub-lexical processing. This investigation is particularly important for the German language system: in German, considerably more exception rules exist for writing than for reading as compared to English (De Bleser, 2005), and thus more graphemic entries are stored in the output lexicon. For example, derivation of written language from spoken language is less transparent than vice versa: the German words *'mehr'*, *'Meer'*, and *'Mär'* are all pronounced as *[ˈmeːɐ̯]* when read. Accordingly, writing *[ˈmeːɐ̯]* to dictation leads to at least three potential different spellings. Correct spelling of [ˈmeːɐ̯], however, is only possible when the respective semantic and syntactic context is known. i) In Study 3, and as regards lexical processing, rule words (i.e., for which a lexical entry has to be activated in order to write it correctly, e.g., *'Bus'*, *'Garage'*, etc.) were used to examine whether there already exists an entry in the output lexicon for these words. ii) As regards sub-lexical processing, minimal pairs of some rule words (e.g., *'Bus - Mus', 'Lok – Lot'*) were used to identify letter-to-sound-mapping strategies (e.g., *'Buss - Muhs', 'Lock – Loht'*). For letter-to-sound-mapping regularization, slower writing times were observed in Studies 3 and 4, corroborating earlier findings (Kandel & Perret, 2015). The rule-specific analysis of spelling performance represented a valuable alternative to traditional spelling diagnostics and allowed for a fine-grained picture of the development of spelling competencies. By presenting an adaptive assessment for orthography and arithmetic, Study 5 revealed that the demand of a task can be adapted to a child's individual performance. For spelling, this adaption is hypothesized to affect the graphemic output lexicon. Evidence for this hypothesis was provided by the training results reported in Study 6. Playful learning of, for instance, morphological aspects, phonology, assembly of words, and language specific exceptions by means of digital learning games significantly enhanced orthographic skills. For arithmetic, training effects of the arithmetic learning games, which focused on arithmetic facts and calculation, were explained by more unspecific cognitive functions such as procedural knowledge or working memory (Roesch et al., 2016). However, adaptive procedures are generally very promising, given the considerable variance in orthography and arithmetic skills observed in the children in Studies 3 and 4.

For the number domain, I separated the input lexicon from an output lexicon, which are required in number transcoding (but see Barrouillet et al., 2004). The major advantage of this approach is the possibility to explain intrinsic 'semantic' writing which is not preceded by any input. For example, when writing a shopping list or the receipt of a new dish, the required number or quantity of goods has to be noted. In this case, the semantic system (e.g., 'I need 500g flour, three eggs, etc.') is proposed to activate the particular output lexicons in order to retrieve the correct Arabic digit number format (i.e., Arabic output lexicon) and the correct spelling of the (number) words (i.e., graphemic output lexicon). Accordingly, my integrated writing model might be generalizable to other context beyond writing-to-dictation and copying.

For arithmetic processing, the most important limitation of my writing model is fact that writing down the solution of an arithmetic problem represents the last step of written calculation. Therefore, it seems difficult to disentangle the origin of an incorrect answer. An in-depth error analysis could reveal whether incorrect answers might have occurred because of calculation errors or because of violations to the place-value principle in number transcoding. Violations to the latter principle are more likely related to writing processes, in particular to the Arabic output lexicon. Barroulliet and colleagues (2004) used such an error analysis to evaluate the ADAPT model. In particular, the authors referred, for instance, to empirical data of patient CK initially presented by

Delazer and Denes (1998): The patient showed a severe agraphia for both alphabetic (Schonauer & Denes, 1994, for a detailed description of agraphia symptoms) and numerical scripts following a left hemispheric stroke. Writing of two-digit numbers was preserved but syntactic errors occurred for increasing numbers. According to the ADAPT model, transcoding errors resulted specifically from an impaired lexicon (see Barroulliet et al. 2004, for a more detailed analysis). In writing of alphabetic scripts, the patients produced only neologisms (i.e., meaningless letter strings) even though handwriting execution was preserved. As neologisms also affected typing and oral spelling, his writing deficit was related to central components of writing. Both Barroulliet (2004) and Schonauer (1994) hypothesized that these writing deficits were caused by impaired lexical processing and deficits in working memory capacity. This hypothesis implied that impairments to central processing stages can cause domain-general writing deficits. However, number processing was less impaired, which might be attributed to the different contribution of semantic information to language and number processing (Eger et al., 2003; Klein et al., 2010; Lochy et al., 2003).

8.1.3 Concept and quantity: The contribution of semantic information to writing

In the writing model, I suggested that semantic processing is executed within two separated semantic modules, one for each domain, because these modules differ in their internal organization and kind of processing (cf. Dehaene, 1992; Dehaene & Cohen, 1995; Dehaene et al., 2003; Ellis, 1982, 1988): for the language domain, and in most languages, a word (e.g., *'the hand'*, *'die Hand'*, or *'la main'*) denotes a descriptive basic semantic meaning (i.e., 'body part', 'extremity', 'four fingers and a thumb', etc.). This basic meaning can be superimposed by individually different emotional, stylistic and other meaning components (i.e., connotations: *'Pranke'* vs. *'Klaue'* vs. *'Patsche'* vs. *'Pfote'*; Löbner, 2015). For the number domain, a number denotes always and stable the same numerical quantity. Occasionally, numbers denote some other nominal meanings, such as '112' is associated with the fire department in Germany, or '911' associated with the US emergency or the terrorist attack in the USA on September 11th in 2001. However, the basic meaning of a word is not accessible when the word is divided into single letters or graphemes ([h] [a] [n] [d]. The same applies to the associated meaning of number strings ([1] [1] [2]). However, number magnitude remains accessible, even in single digits.

On the behavioral level, access to number magnitude was suggested to facilitate writing of numbers (Lochy et al., 2003). In this vein, Delazer and colleagues (2002) reported on patient JS who suffered a left-hemispheric ischemic stroke followed by severe agraphia for alphabetic but not numerical scripts. During a writing training the patient learned to associate letter forms with real objects. As in Study 1, this training was successful. Accordingly, semantic information was concluded to enhance writing of alphabetic scripts. More recent evidence for this conclusion, is provided from a study of aphasic patients (Rath et al., 2015), in which semantic meaning not only facilitated processing; semantic demands on a task also increased the accuracy in solving the task, also in language processing.

On the neural level, findings of Study 1 and 2 also provide evidence for domain-specific networks in the brain: for the language domain, a left-hemispheric parieto-frontal network for writing letters was identified which is consistent with previous results (e.g., Beeson et al., 2003; Longcamp, Anton, Roth, & Velay, 2003; Planton et al., 2013; Roux, McKeeff, Grosjacques, Afonso, & Kandel, 2013). Within this network the left angular gyrus was confirmed to be significant for semantic activation and "integrating individual concepts into a larger whole" (Binder, Desai, Graves, & Conant, 2009, p. 2776). Moreover, retrieval of semantic mental images seemed to be reflected by left hemispheric precuneus activation for writing letters in Study 1. Activation of the left precuneus and its co-activated homologue area in the right hemisphere have been identified to support retrieval of visual–spatial information from memory (Cavanna & Trimble, 2006; Ganis et al., 2004). This concurred well with the retrieval of abstract semantic information for single digits, which was unimpaired in the patient CU. For the number domain, a bilateral fronto-parietal network was found activated for writing Arabic digits (e.g., Dehaene et al., 2003), including the horizontal part of the intraparietal sulcus (hIPS). The hIPS has been associated with processing numerical quantity, and was suggested as a candidate region for number specificity (Dehaene et al., 2003). Crucially, lesions to the hIPS have been observed in adults and children with (developmental) Gerstmann syndrome (Dehaene & Cohen, 1997; Gerstmann, 1940; Kinsbourne & Warrington, 1963). Deficits in Gerstmann syndrome involve severe acalculia and dyscalculia, respectively. Acalculia co-occurs with impairments in spatial left-to-right orientation. It is likely that the physical impairment in spatial orientation might affect the mental association of numbers and space in the semantic system, and thus cause calculation deficits.

In summary, the results of Studies 1 and 2 substantiated the presumption of two domain-specific modules which may be involved in writing. As argued above, these modules differ in terms of their internal organization and in their neural substrate located in two domain-specific but connected brain networks. However, although semantic magnitude information was suggested to be automatically retrieved (Eger et al., 2003; Klein et al., 2010), it has not yet been established whether semantic processes are necessary or not for successful number transcoding. This issue may be addressed in a study suggested in the future perspectives of my thesis.

8.1.4 A shared buffer (memory) system?

Working memory capacity was shown to be relevant for writing across domains. For this reason, I proposed a shared multimodal buffer system following Baddeley (2000). The buffer system is considered to bridge central and peripheral writing processes (Purcell, Turkeltaub, et al., 2011), and to play a significant role in writing alphabetic scripts (Baddeley, 2003; Caramazza et al., 1987) and in number transcoding (Camos, 2008; Moura, 2014). For the language domain, the buffer system was assumed to be of multi-dimensional structure providing abstract information about letter identity and their sequential position in the (syllabic) consonant-vowel structure of a word (Domahs, 2016). Such coding of a consonant-vowel structure does not apply to Arabic digits. For the number domain, serial order of digits in the buffer system was found to be of specific importance (Lochy et al., 2003) for correct number transcoding (Barrouillet et al., 2004).

Lesions in left frontal and parietal brain areas, in particular to the superior prefrontal area, were repeatedly observed to lead to impairments of the buffer system (Planton et al., 2013). Impairments to the buffer system provoke length effects as expressed by the number of letters and digits (Delazer & Denes, 1998) and serial order effects (i.e., permutation, substitutions, omission, additions of letters, and digits) reported in healthy subjects (Power and Dal Martello, 1997) and brain-injured patients (e.g., Cipolotti, Butterworth, & Warrington, 1994; Macoir, Audet, & Breton, 1999). Patient CK presented by Delazer and Denes (1998), for instance, showed length effects in both alphabetic and numerical scripts, which corroborates the idea of an integrated buffer system.

Lochy and colleagues (2003) hypothesized that deficits in the buffer system cause mixed syntactic and lexical problems ('32104' – '31024'). An impaired integrated buffer system for graphemes and abstract digits, might also provide an explanation for mixed writing of both scripts. Mixed writing in patient DK was reported by Patterson and Wing (1989). The patient produced substitutions between digits and letters '1S', '1N', '1D4' when he was asked to write down his house number '14'. Although the authors assumed the deficit to originates from a later processing stage, an impaired buffer might also explain substitutions between letters and digits. To give a further example: following an old but still used book printing rule, numbers from 1-12 have to be written as number words within a text (e.g., 'I read one or two books per month'). Numbers larger than 12 should be given in the Arabic number format (e.g., 'I read about 13 books per year'). Quite often, however, it happens that for the numbers up to 12 the Arabic digit number format is automatically used (e.g., 'I read 1 or 2 books per month'). The verbal number word format has to be activated by a controlled retrieval process. Unlike the verbal number word format, activation of the Arabic digit number format is suggested to proceed automatically due to the inherent semantic information of digits which is closely related the decimal structure of the Arabic digits. Consequently, conflicting domain-specific and domain-general processes might occur when the buffer system is impaired.

Existence of the buffer system was proposed for both handwriting (Ellis, 1982, 1988) and typing (Margolin, 1984; Pinet et al., 2016). In this buffer system abstract graphemes and digits are kept active for further processing (Will, Nottbusch, & Weingarten, 2007, for a discussion the of role of the graphemic buffer in typing). Even though, temporal dynamics of handwriting and typing differ significantly, graphemes and digits have to be arranged in a specific sequential order to achieve the correct writing or typing of a word or a multi-digit number. Further investigations are needed to provide more evidence for the theoretical assumption of one integrated, mode independent buffer system.

8.2 Peripheral writing processes

Peripheral writing components are required to produce an overt written response (Ellis, 1982, 1988; Margolin, 1984; Van Galen, 1991). A critical brain region specifically involved in the process of written response generation across domains is Exner's area (Planton et al., 2013; Roux et al., 2010). The role of Exner's area for typing, however, is still discussed controversially (cf. Purcell, Napoliello, & Eden, 2011; Planton et al., 2013). In Study 2, Exner's area was found to integrate domain-specific information for writing letters and digits, which are conveyed via two distinct networks for language and number processing. Therefore, I proposed domain-general graphomotor plans and programs. However, these programs are suggested to be mode-specific, because motor execution of writing and typing differ significantly.

8.2.1 Set, case, style - domain-specific differences at the allographic level

Based on the present results and previous literature, I proposed two distinct domain-specific allographic modules for writing of letters and digits (Beeson et al., 2003; Lochy et al., 2003). For the language domain, impairments to the allographic level might lead to substitutions of similar allographs (Cameron, Cubelli, & Della Sala, 2002; Rapp & Caramazza, 1997), *case* PREFERENCES (Menichelli et al., 2008;; Patterson & Wing, 1989), or to writing of MiXEd CaSeS (De Bastiani & Barry, 1989). For example, Rapp and Caramazza (1997) described two patients who produced meaningless spelling due to severe letter substitutions, whereas oral spelling was relatively intact. Letter substitutions typically involved letters that are visually but more likely graphomotorically similar to the target letter (Beeson & Rapcsak, 2015, for a review). High similarity of allographs impeded processing at the allographic level which was more pronounced for writing letters than digits (Lochy et al., 2003). Letters were assumed to show more similarities than digits due to the larger set of at least 26 alphabetic letters as compared to 10 digits. This issue might explain why mere allographic writing impairments have not been described for Arabic digits.

For the number domain, empirical evidence further substantiates the assumption of distinct allographic modules: Patterson and Wing (1989) presented patient DK who suffered from a left parietal stroke of the cerebrovascular artery. Language comprehension, reading and oral spelling were well preserved, whereas arithmetic skills were initially severely impaired as was handwriting in both domains. Interestingly, italics lower-case writing of letters was significantly more impaired than writing of upper-case letters. These case-preferences led Patterson and Wing (1989) to conclude that "the first candidate locus for his [the patient's] deficit is the physical letter code" (p.19), which is stored at the allographic level. During recovery, numerical impairments almost disappeared, and DK recovered completely for the writing of single- and multi-digit numbers. Writing of letters and words, however, was still impaired. Briefly, after the stroke, handwriting difficulties were comparable. Over the course of the disease, handwriting improved differently in time and degree for the two domains, which again argues for domain-specific processing at allographic level.

However, DK also produced mixed-errors between letters and digits *'1S', '1N', '1D4',* when he was asked to write *'14'* (Patterson & Wing, 1989). Although mixed errors were previously explained by impaired working memory capacity (see previous paragraph), impaired access to the allographic level also seemed probable for two reasons: First, the unit digits of '14' was replaced inconsistently by various graphemes. Inconsistent errors are related to an access disorder whereas constant errors might indicate an impairment of the representation itself (Patterson & Wing, 1989). An impairment of the buffer system was assumed less likely because mixed errors would also occur when writing alphabetic scripts. An impairment at the allographic level itself might also be excluded, because DK was able to write upper-case letters more or less correctly. Writing of lower-case letters, however, was severely impaired. This might have occurred due to case-sensitive selection procedures or case "switching" difficulties (Menichelli et al., 2008). Accordingly, switching between the language and the number domain might also be highly likely implying that the allographic modules are interconnected. This is yet another issue regarding the relationship between writing letters and digits that needs to be clarified in the future.

8.2.2. From domain-general graphomotor plans to specific motor programs

Two successive domain-general but mode-specific modules are assumed in my integrated writing model for the conversion of graphomotor plans into specific motor programs (Figures 7.5 and 7.6). Domain-general processing is proposed based on results of Studies 1 and 2. In these studies, Exner's area was identified to be specifically associated with graphomotor commands in handwriting for both single letters and Arabic digits corroborating previous findings (Eberhardt 2010; Planton et al., 2013). This specific functional role of Exner's area argues for a domain-specific integration of the two brain networks involved in writing letters and digits and thus for dissociations in writing of alphabetic and numerical scripts (Anderson et al., 1990; Delazer et al., 2002; Keller & Meister, 2014; Starrfelt, 2007). Mode-specification was based on results from Study 3 in which differences in performances in handwriting and typing (i.e., different writing times) were observed for peripheral writing processes.

Impairments to other peripheral, specific hand motor programs might cause the production of malformed letters. In this context, brain damage observed in left dorsal premotor regions (Beeson et al., 2003) was suggested to affect writing of letters and digits differently. Empirical evidence for this hypothesis was found in the work by Delazer and colleagues (2002). The authors describe patient JS, who showed an impressive dissociation when asked to write the letter 'O' and the digit '0'. Although almost identical graphomotor plans and hand motor programs can be assumed, JS had severe difficulties to write the alphabetic letter. His difficulties were associated with domain-specific differences at processing stages prior to domain-general graphomotor planning.

In addition, evidence for shared graphomotor plans and programs for both scripts can be found in the development of writing. In particular, mirror writing for letters and digits has also been described in typically developing children. For instance, my 4-year old daughter wrote her first name Thea as . In this example, she wrote from left-to-right and the letter *'E'* was also mirror-inverted. In the early stages of writing, children learn to identify various sets of invariant features specifically associated to a certain character (Fischer, 2017). These features are then stored as effector independent graphomotor plans for each letter of the alphabet and each of the ten digits (Schubert, 2017, for a systematic analysis of the graphomotor plan for capital letters and digits). Through practice and experience mirror errors have been found to vanish. Once acquired, retrieval of graphomotor plans as well as hand motor programs is highly automatized.

Difficulties in the automation of handwriting were also been found to be associated with both making more errors and taking more time during writing (Berninger, 2015). The lack of automation might cause differences in handwriting and typing performance reported in Study 3: for all children, writing times were slower for keyboard typing than for handwriting. Although spelling performance finally did not differ between handwriting and typing, except for the capitalization of letters, typing was much more error-prone (see also Feng, Lindner, Ji, & Malatesha Joshi, 2019; Goldberg et al., 2003). The different graphomotor programs for handwriting and typing as well as the different degree of automation provide evidence to assume distinct processing pathways. These pathways have been found to be influenced by specific linguistic aspects such as spelling rules, double letters or morpheme boundaries (Weingarten, 2004). This finding suggests that central writing processes might have a comparable effect on peripheral writing components in writing alphabetic and numerical scripts. However, there is still an ongoing debate about the integration of central and peripheral writing processes (Roux et al., 2013).

Taken together, the integrated writing model proposed in the present thesis suggests a novel approach to investigate writing of alphabetic and numerical scripts from a cross-domains perspective including different writing modalities. Model evaluation based on previous empirical findings and the studies enclosed support the theoretical assumption of central (cf. Studies 1,2 and 4) and peripheral processing stages (cf. Study 3) involved in writing letters and numbers as well as the underlying neural correlates (cf. Studies 1 and 2). The integrated writing model was feasible to serve as a diagnostic and interventional foundation for typically (cf. Studies 5 and 6) and atypically developed writing skills (cf. Study 3) using different writing modalities. Thereby, the proposed integrated writing model makes a contribution to our understanding of writing in different contexts.

9 Future Perspectives

The development of my integrated writing model for alphabetic and numerical scripts constitutes one step toward future research on the cognitive mechanisms of writing. However, evaluation of the model points to several open research questions: The first question addresses the contribution of semantic information in transcoding letters and numbers. The second question addresses the impact of the writing mode on writing. In the following, I will briefly elaborate on these in turn.

9.1 The contribution of semantic information for writing

For writing alphabetic scripts, semantic information was repeatedly found to facilitate writing performance in the present thesis and in previous studies (e.g., Bub & Kertesz, 1982). With respect to letter transcoding, however, the influence of semantic information still needs to be examined in more depth. For this purpose, the following neuro-functional study might be carried out: In letter transcoding, visually presented lexical and non-lexical verbal material (i.e., words and non-words) has to be transcoded from UPPERCASE to lowercase letters and vice versa[12]. Based on the results of this thesis, semantic information should facilitate transcoding of word material. On the behavioral level, shorter transcoding times might indicate facilitation. On the neural level, the left angular gyrus should be specifically activated for transcoding real words (Binder et al., 2009, for a review) as compared to transcoding non-words. Moreover, semantic 'meaningfulness' might be graduated by using function words (e.g., *'VOR'*, *'BEI'*), abstract words (e.g., *'MUT'*, *'TOD'*) and concrete highly imageable words (e.g., *'ZEH'*, *'HUT'*). In this way, the influence of different degrees of semantic meaningfulness (i.e., pictorial words are more meaningful than abstract and function words) on the cognitive mechanisms underlying writing can be evaluated and associated with the respective brain network. Moreover, copying of real words and pseudo words may be compared to the writing of multi-digit numbers to disentangle whether semantic information is activated in this task (cf. Artemenko et al., 2018; for a comparison of meaningless letter strings, e.g., '*HMB*' with multi-digit numbers).

For number transcoding, the debate about semantic and asemantic processing (Nuerk et al., 2011) cannot be resolved through this thesis. Contradictory evidence is provided for both processes by functional imaging studies suggesting automatized semantic processing (Eger et al., 2003; Klein et al., 2010) and by computational modeling suggesting purely asemantic processing

[12] Transcoding abilities for auditorily presented stimuli cannot be investigated by this experiment.

(Barrouillet et al., 2004). Insights from the present thesis indicate that it might be misleading to pose this as an ‘either/or’ question. The question may rather be whether semantic information can facilitate number transcoding and processing. Further neuro-cognitive studies are needed to explicitly investigate the activation of semantic number magnitude information in number transcoding. For this purpose, the following experiment might be conducted: Number transcoding may be combined with a semantic magnitude comparison task. Two multi-digit numbers (e.g., 123 vs. 321) are presented either a) auditorily or b) visually.

a) The task design follows a writing-to-dictation task. But auditory stimuli are manipulated at the surface, this means, in terms of loudness or pitch. First, in the experimental condition, the auditorily presented numbers have to be encoded and simultaneously ordered by their numerical quantity. Second, the same number pairs have to be ordered by loudness or pitch and not according to their quantity. In the control condition, stimulus pairs have to be transcoded. From the results of the experiment the following conclusions can be drawn: first, different properties of the presented numbers (i.e., semantic number magnitude information and acoustic information at the surface level) of the numbers must be considered during transcoding, whereby reaction time differences may be expected. Second, evaluation of the superficial information (i.e., loudness or pitch) slows down transcoding times RT because this information has to be processed in addition to the number transcoding task. Third, if transcoding times do not differ between control task and semantic condition, it might be assumed that semantics do not need to be processed additionally, because it is automatically activated in number transcoding. However, investigating neuronal processes in the brain provides more precise information on the type of processing.

b) The task design follows number transcoding based on a copying task. Again, surface properties of the visually presented number words are manipulated (e.g., in terms of differently colored numbers, i.e., 123 vs. 321). First, in the experimental condition, number words of the visually presented Arabic digit number pairs have to be written according to their surface information (i.e., sorting the number words by brightness) Second, in number transcoding the number pairs have to be simultaneously ordered by their numerical magnitude. In the control condition, both tasks have to be compared to the traditional task including the same stimuli (i.e., 123 vs. 321). On the behavioral level, differences in reaction times might indicate differences in semantic processing. In this vein, it is of particular interest whether semantic information is also activated when surface information has to be specifically retrieved. On the neural level, differential activation in the hIPS might reflect the differences between the contribution of actually required or automatically activated semantic information (Dehaene et al., 2003). These results might then be compared to the experimental results of the language domain. A very recent study by Karimpoor and colleagues (2018) provided empirical evidence that handwriting of up to 10 symbols (i.e., letters, numbers) and even multiple words can be obtained in the fMRI by means of a digital touch-typing device. Hence, the proposed experiment is assumed to be well feasible.

In both domains, an influence of phonological processing on writing and transcoding processes might be of particular interest. Phonological skills have been shown to be related to verbal representations as well as to manipulations of numbers (Andin et al., 2015; Dowker & Nuerk, 2016) and to be directly involved in writing words (Bonin et al., 2015; Planton et al., 2013). To investigate phonological influences in both domains, the transcoding experiments introduced above could also be conducted under articulatory suppression (i.e., by singing a constant tone or saying PATAKA over and over again), which should have an impact on performance (i.e., inversion errors) and reaction times during transcoding.

9.2 The impact of writing mode on writing

Handwriting and typing were found to be sensitive to specific linguistic aspects (Weingarten, 2004), and to different spelling rules as demonstrated in Studies 3 and 4. However, it has not yet been clarified whether these aspects are language specific. In both studies, disregarding capitalization rules was the most common source of errors. The use of initial upper-case letters within sentences distinguishes German from all other alphabetic scripts (Günther & Nünke, 2005). For this reason, it is most likely that the present results are not easily generalizable to other languages. In terms of generalizability, the following experiments might provide further insights: First, cross-linguistic comparison of writing and typing for different languages might be fruitful to disentangle language specific effects on peripheral components of writing. These studies should move beyond writing on single words because the mode effect was found specifically for higher writing skills (Feng et al., 2019; Goldberg et al., 2003; Russell, 1999).

Second, the mode effect was found to impact language and number processing differently (i.e., writing mode affected reading and writing bot not number processing; Wang et al., 2007; Wang, Hong Jiao, Young, Brooks, & Olson, 2008; for meta-analyses). These findings do not fully agree with in the present thesis. Future work should address this issue by intra-individual comparisons in writing and typing in each domain. In this way, it may be investigated whether typing is more error-prone not only in writing alphabetic scripts but also in other cognitive domains. This is of particular interest as digital media are increasingly used for diagnostics and remediation of cognitive functions. Provided that digital media are not only attractive and motivating (Boyle et al., 2016; Ninaus et al., 2015), but also comparable to traditional methods, they may revolutionize not only writing itself but also its investigation.

10 Conclusions

Taken together, using the example of writing in different notational systems (i.e., alphabetic and numerical scripts), I showed in this thesis that it is possible to bring together various theoretical approaches of different research domains (i.e., language and number domain) providing a joint framework for the cognitive study of writing across these domains. Within this framework, I examined domain-specific processes necessary to generate alphabetic and numerical scripts as well as domain-general processes required for the execution of writing by means of traditional and digital writing media (i.e., handwriting and keyboard typing). By combining insights from both domains, I proposed a novel integrated writing model which allows to represent typical and atypical writing processes. Furthermore, my model allows to theoretically assess changes to or the development of these writing processes and assign them to specific processing stages. Consequentially, this thesis did not only provide an integrated theoretical framework to better understand the cognitive processes and neural correlates involved in writing in different contexts. It further serves as a theoretically substantiated and empirically supported foundation for diagnosis and intervention of writing, within or across domains and writing modes. Beyond writing, the present thesis demonstrates that the understanding of human cognition can be broadened by looking beyond the boundaries of the respective field of research.

My integrated writing model is developed for writing-to-dictation and copying of different scripts only. Consequently, the model represents writing processes which are particularly required in this specific writing context. Conclusions might not be generalizable to higher writing processes such as writing texts. As the model assumes serial writing processes, it adopts a rather artificial perspective on writing. However, this perspective facilitates the understanding of individual processing steps involved in writing. Nevertheless, they are only basic components for higher writing skills. Moreover, by the development of my integrated writing model, previous handwriting models have been extended to keyboard typing for various cognitive domains. This extension suggests that, nowadays, writing using digital media may build on other cognitive mechanisms than only a few decades ago (Ardila, 2012). These cognitive mechanisms may also develop atypically or may be impaired, for instance, because of brain-injury. In this way, new technologies always entail the need for state-of-the-art research.

References

Alexander, M. P., Friedman, R. B., Loverso, F., & Fischer, R. S. (1992). Lesion localization of phonological agraphia. *Brain and Language*, *43*(1), 83–95.

Amalric, M., & Dehaene, S. (2018). Cortical circuits for mathematical knowledge: evidence for a major subdivision within the brain's semantic networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20160515.

Anderson, S. W., Damasio, A. R., & Damasio, H. (1990). Troubled letters but not numbers. *Brain*, *113*, 749–766.

Andin, J., Fransson, P., Rönnberg, J., & Rudner, M. (2015). Phonology and arithmetic in the language-calculation network. *Brain and Language*, *143*, 97–105.

Ardila, A. (2012). Neuropsychology of writing. *Writing: A Mosaic of New Perspectives*, (1998), 325–344.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, *54*(3), 2382–2393.

Assaf, Y., & Pasternak, O. (2008). Diffusion tensor imaging (DTI)-based white matter mapping in brain research: A review. *Journal of Molecular Neuroscience*, *34*(1), 51–61.

Artemenko, C., Coldea, A., Soltanlou, M., Dresler, T., Nuerk, H. C., & Ehlis, A. C. (2018). The neural circuits of number and letter copying: an fNIRS study. *Experimental brain research*, *236*(4), 1129-1138.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Baldo, J. V, & Dronkers, N. F. (2007). Neural correlates of arithmetic and language comprehension: A common substrate? *Neuropsychologia*, *45*(2), 229–235.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*(4), 553–573.

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, *44*(4), 625–632.

Beattie, G., & Ellis, A. W. (2017). The Psychology of Language And Communication. In *The Psychology of Language And Communication*.

Beauvois, M.-F., & Dérouesné, J. (1981). Lexical or orthographic agraphia. *Brain: A Journal of Neurology*, *104*(Pt 1), 21–49.

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., ... Thomas, S. M. (2013). The mental wealth of nations. In *From Stress to Wellbeing* (Vol. 2, pp. 271–279).

Beeson, P. M., & Rapcsak, S. Z. (2015). Neuroanatomical Correlates of Spelling and Writing. In *The Handbook of Adult Language Disorders* (pp. 103–132). Psychology Press.

Beeson, P., Rapcsak, S., Plante, E., Chargualaf, J., Chung, A., Johnson, S., & Trouard, T. (2003). The neural substrates of writing: A functional magnetic resonance imaging study. *Aphasiology*, *17*(6–7), 647–665.

Berninger, V. W., Abbott, R. D., Augsburger, A., & Garcia, N. (2009). Comparison of Pen and Keyboard Transcription Modes in Children with and without Learning Disabilities. *Learning Disability Quarterly*, *32*(3), 123–141.

Berninger, V., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology*, *34*(1), 23–52.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.

Bloechle, J., Huber, S., Bahnmueller, J., Rennig, J., Willmes, K., Cavdaroglu, S., … Klein, E. (2016). Fact learning in complex arithmetic???the role of the angular gyrus revisited. *Human Brain Mapping*, *37*(9), 3061–3079.

Bonin, P., & Meot, A. (2002). Writing to dictation in real time in adults: What are the determinants of written latencies. *Advances in Psychology Research*, *16*, 139–165.

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *Quarterly Journal of Experimental Psychology*, *68*(7), 1268–1294.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., … Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, *94*, 178–192.

Bub, D., & Kertesz, A. (1982). Deep agraphia. *Brain and Language*, *17*(1), 146–165.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47.

Cameron, A., Cubelli, R., & Della Sala, S. (2002). Letter assembling and handwriting share common allographic code. *Journal of Neurolinguistics*, *15*(2), 91–97.

Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, *99*(1), 37–57.

Caramazza, A., Miceli, G., Villa, G., & Romani, C. (1987). The role of the graphemic buffer in spelling: Evidence from a case of acquired dysgraphia. *Cognition*, *26*(1), 59–85.

Catani, M., Dell'Acqua, F., Bizzi, A., Forkel, S. J., Williams, S. C., Simmons, A., … Thiebaut de Schotten, M. (2012). Beyond cortical localization in clinico-anatomical correlation. *Cortex*, *48*(10), 1262–1287.

Cazzolli, G., Maffei, C., Miceli, G., Catani, M., Dell'Acqua, F., Capasso, R., … Piludu, F. (2017). Pure word deafness following left temporal damage: Behavioral and neuroanatomical evidence from a new case. *Cortex*, *97*, 240–254.

Cipolotti, L., & Butterworth, B. (1995). Toward a multiroute model of number processing: Impaired number transcoding with preserved calculation skills. *Journal of Experimental Psychology: General*, *124*(4), 375.

Cipolotti, L., Butterworth, B., & Warrington, E. K. (1994). From "one thousand nine hundred and forty-five" to 1000, 945. *Neuropsychologia*, *32*(4), 503–509.

Cohen, L., & Dehaene, S. (1991). Neglect dyslexia for numbers? A case report. *Cognitive Neuropsychology*, *8*(1), 39–58.

Cohen, L., Dehaene, S., McCormick, S., Durant, S., & Zanker, J. M. (2016). Brain mechanisms of recovery from pure alexia: A single case study with multiple longitudinal scans. *Neuropsychologia*, *91*, 36–49.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204.

Connelly, V., Gee, D., & Walsh, E. (2007). A comparison of keyboarded and handwritten compositions and the relationship with transcription speed. *British Journal of Educational Psychology*, *77*(2), 479–492.

Damian, M. F. (2004). Asymmetries in the processing of Arabic digits and number words. *Memory & Cognition*, *32*(1), 164–171.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, *39*(6), 507–514.

De Bastiani, P., & Barry, C. (1989). A cognitive analysis of an acquired dysgraphic patient with an “allographic” writing disorder. *Cognitive Neuropsychology*, *6*(1), 25–41.

De Bleser, R. (2005). Dyslexie und Dysgraphie. *Kognitive Neurologie*, 65.

De Smedt, B., Taylor, J., Archibald, L., & Ansari, D. (2010). How is phonological processing related to individual differences in children’s arithmetic skills? *Developmental Science*, *13*(3), 508–520.

Dehaene, S. (1991). Neglect dyslexia for numbers? a case report. *Cognitive Neuropsychology*, *8*(1), 39–58.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42.

Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. *Annals of the New York Academy of Sciences*, *1156*, 232–259.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, Vol. 1, pp. 83–120.

Dehaene, S., & Cohen, L. (1997). Cerebral Pathways for Calculation: Double Dissociation between Rote Verbal and Quantitative Knowledge of Arithmetic. *Cortex*, *33*(2), 219–250.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3–6), 487–506.

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*(6), 1330–1340.

Delazer, M., & Bartha, L. (2001). Transcoding and calculation in aphasia. *Aphasiology*, *15*(7), 649–679.

Delazer, M., & Denes, G. (1998). Writing arabic numerals in an agraphic patient. *Brain and Language*, *64*(2), 257–266.

Delazer, M., Lochy, A., Jenner, C., Domahs, F., & Benke, T. (2002). When writing 0 (zero) is easier than writing O (o): *Neuropsychologia*, *40*, 2167–2177.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language Production and Serial Order: A Functional Analysis and a Model. *Psychological Review*, *104*(1), 123–147.

Deloche, G., & Seron, X. (1982). From one to 1: An analysis of a transcoding process by means of neuropsychological data. *Cognition*, *12*, 119–149.

Deloche, G., & Seron, X. (1987). Numerical transcoding: A general production model. In *Mathematical disabilities: A cognitive neuropsychological perspective*.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184–192.

Domahs, F. (2016). Gestörte graphematische Verarbeitung: Alexie und Agraphie. In U. Domahs & B. Primus (Eds.), *Handbuch Laut, Gebärde, Buchstabe* (Vol. 2). Walter de Gruyter GmbH & Co KG.

Donders, F. C. (1868). Die schnelligkeit psychischer processe: Erster artikel. *Archiv Für Anatomie, Physiologie Und Wissenschaftliche Medicin*, 657–681.

Dotan, D., & Friedmann, N. (2018a). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, *101*, 249–281.

Dotan, D., & Friedmann, N. (2018b). Separate mechanisms for number reading and word reading: Evidence from selective impairments. *Cortex*, 1–31.

Dowker, A., & Nuerk, H.-C. (2016). Linguistic influences on mathematics. *Frontiers in Psychology*, *7*, 1035.

Eberhardt, K. (2010). *Exner reloaded – a contribution to Exner ' s Area in writing letters and numbers Exner reloaded – a contribution to Exner ' s Area in writing letters and numbers*. Diploma thesis.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A., & Kleinschmidt, A. (2003). 2003 A Supramodal Number Representation in Hunan Intraparietal Cortex .pdf. *Neuron*, *37*, 719–725.

Eickhoff, S. B., & Grefkes, C. (2011). Approaches for the integrated analysis of structure, function and connectivity of the human brain. *Clinical EEG and neuroscience*, *42*(2), 107-121.

Ellis, A. W. (1982). Spelling and writing (and reading and speaking). *Normality and Pathology in Cognitive Functions*, 113–146.

Ellis, A. W. (1988). Normal writing processes and peripheral acquired dysgraphias. *Language and Cognitive Processes*, *3*(2), 99–127.

Ellis, A. W., & Young, A. W. (2013). *Human cognitive neuropsychology: A textbook with readings*. Psychology Press.

Fayol, M., & Seron, X. (2005). About numerical representations: Insights from neuropsychological, experimental, and developmental studies. In *Handbook of mathematical cognition.* (pp. 3–22). New York, NY, US: Psychology Press.

Feng, L., Lindner, A., Ji, X. R., & Malatesha Joshi, R. (2019). The roles of handwriting and keyboarding in writing: a meta-analytic review. *Reading and Writing*, *32*(1), 33–63.

Galen, G. P. Van. (1991). Handwriting : psychomotor Issues for a theory *. *Human Movement Science*, *10*, 165–191.

Gaskill, J. L., & Marshall, M. (2007). *Comparisons Between Paper-and Computer-based Tests: Foundation Skills Assessment-2001 to 2006 Data*. Society for the Advancement of Excellence in

Education.

Gentner, D. R. (1983). Keystroke timing in transcription typing. In *Cognitive aspects of skilled typewriting* (pp. 95–120). Springer.

Gerstmann, J. (1940). Syndrome of finger agnosia, disorientation for right and left, agraphia and acalculia: local diagnostic value. *Archives of Neurology & Psychiatry*, *44*(2), 398–408.

Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, *2*(1), 2–51.

Graham, S., & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review*, *8*(1), 7–87.

Granà, A., Girelli, L., Gattinoni, F., & Semenza, C. (2001). Letter and number writing in agraphia: A single-case study. *Brain and Cognition*, *46*(1–2), 149–153.

Gross, J. (2006). The long term costs of literacy difficulties: KPMG Foundation.

Gross, J., Hudson, C., & Price, D. (2009). The long term costs of numeracy difficulties. Every Child a Chance Trust and KPMG, London.

Grudin, J. T. (1983). Error patterns in novice and skilled transcription typing. In *Cognitive aspects of skilled typewriting* (pp. 121–143). Springer.

Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, *19*(7), 374–382. h

Higashiyama, Y., Takeda, K., Someya, Y., Kuroiwa, Y., & Tanaka, F. (2015). The neural basis of typewriting: a functional MRI study. *PloS One*, *10*(7), e0134131.

Imbo, I., Vandierendonck, A., & De Rammelaere, S. (2007). The role of working memory in the carry operation of mental arithmetic: *Quarterly Journal of Experimental Psychology*, *0218*(0), 1–48.

Kandel, S., Lassus-Sangosse, D., Grosjacques, G., & Perret, C. (2017). The impact of developmental dyslexia and dysgraphia on movement production during word writing. *Cognitive Neuropsychology*, *34*(3–4), 219–251.

Kandel, S., & Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition? *Cognition*, *136*, 325–336.

Karimpoor, M., Churchill, N. W., Tam, F., Fischer, C. E., Schweizer, T. A., & Graham, S. J. (2018). Functional MRI of Handwriting Tasks: A Study of Healthy Young Adults Interacting with a Novel Touch-Sensitive Tablet. *Frontiers in human neuroscience*, *12*, 30.

Kast, M., Baschera, G. M., Gross, M., Jäncke, L., & Meyer, M. (2011). Computer-based learning of spelling skills in children with and without dyslexia. *Annals of Dyslexia*, *61*(2), 177–200.

Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159–180.

Keller, C., & Meister, I. G. (2014). Agraphia caused by an infarction in Exner's area. *Journal of Clinical Neuroscience*, *21*(1), 172–173.

Kinsbourne, M., & Warrington, E. (1963). The Developmental Gerstmann Syndrome. *Archives of Neurology*, *8*(5), 490–501.

Klein, E., Moeller, K., Nuerk, H.-C. C., & Willmes, K. (2010). On the neuro-cognitive foundations of basic auditory number processing: an fMRI study. *Behavioral and Brain Functions*, *6*(1), 42.

Knops, A., Nuerk, H.-C., Fimm, B., Vohn, R., & Willmes, K. (2006). A special role for numbers in working memory? An fMRI study. *Neuroimage*, *29*(1), 1–14.

Kussmaul, A. (1877). Word deafness and word blindness. *Cyclopaedia of the Practice of Medicine. New York, NY: William Wood*, 770–778.

Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology*, *103*(3), 309–324.

Libertus, M. E., Brannon, E. M., & Pelphrey, K. A. (2009). Developmental changes in category-specific brain responses to numbers and letters in a working memory task. *Neuroimage*, *44*(4), 1404–1414.

Löbner, S. (2015). *Semantik: Eine Einführung*. Walter de Gruyter GmbH & Co KG.

Lochy, A., Zoppoth, S., Domahs, F., & Delazer, M. (2003). Peripheral processes in writing: Why are numbers so often preserved? A comparative study of letters vs. digits kinematics. *Le Langage et l'homme*, *38*(2), 25–46.

Logan, G. D., & Crump, M. J. C. (2011). Hierarchical control of cognitive processes: The case for skilled typewriting. In *Psychology of learning and motivation* (Vol. 54, pp. 1–27). Elsevier.

Longcamp, M., Anton, J.-L., Roth, M., & Velay, J.-L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *NeuroImage*, *19*(4), 1492–1500.

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., & Velay, J.-L. (2008). Learning through Hand- or Typewriting Influences Visual Recognition of New Graphic Shapes: Behavioral and Functional Imaging Evidence. *Journal of Cognitive Neuroscience*, *20*(5), 802–815.

Longcamp, M., Zerbato-Poudou, M. T., & Velay, J. L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, *119*(1), 67–79.

Lopes-Silva, J. B., Moura, R., Júlio-Costa, A., Geraldi Haase, V., Wood, G., Haase, V. G., & Wood, G. (2014). Phonemic awareness as a pathway to number transcoding. *Frontiers in Psychology*, *5*(JAN), 1–9.

MacArthur, C. A. (2009). Reflections on Research on Writing and Technology for Struggling Writers. *Learning Disabilities Research & Practice*, *24*(2), 93–103.

Macoir, J., Audet, T., & Breton, M.-F. (1999). Code-dependent pathways for number transcoding: evidence from a case of selective impairment in written verbal numeral to Arabic transcoding. *Cortex*, *35*(5), 629–645.

Mangen, A., & Velay, J.-L. (2010). Digitizing Literacy: Reflections on the Haptics of Writing. *Advances in Haptics*, 385–403.

Margolin, D. I. (1984). The Neuropsychology of Writing and Spelling: Semantic, Phonological, Motor, and Perceptual Processes. *The Quarterly Journal of Experimental Psychology Section A*, *36*(3), 459–489.

Margolin, D. I., & Goodman-Schulman, R. (1992). Oral and written spelling impairments. *Cognitive Neuropsychology in Clinical Practice*, 263–297.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157.

McCloskey, M., & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive Neuropsychology*, *34*(3–4), 65–82.

Menichelli, A., Rapp, B., & Semenza, C. (2008). Allographic agraphia: A case study. *Cortex*, *44*(7), 861–868.

Menon, V. (2016). Memory and cognitive control circuits in mathematical cognition and learning. *Progress in Brain Research*, *227*, 159–186.

Moeller, K., Nuerk, H.-C., & Willmes, K. (2009). Internal number magnitude representation is not holistic, either. *European Journal of Cognitive Psychology*, *21*(5), 672–685.

Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading*, *13*(1), 1–25.

Morken, F., & Helland, T. (2013). Writing in dyslexia: Product and process. *Dyslexia*, *19*(3), 131–148.

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive Approaches in Spelling* (pp. 117–134). London: Academic Press.

Moura, R., Wood, G., Pinheiro-Chagas, P., Lonnemann, J., Krinzinger, H., Willmes, K., & Haase, V. G. (2013). Transcoding abilities in typical and atypical mathematics achievers: The role of working memory and procedural and lexical competencies. *Journal of Experimental Child Psychology*, *116*(3), 707–727.

Nieder, A., & Dehaene, S. (2009). Representation of Number in the Brain. *Annual Review of Neuroscience*, *32*(1), 185–208.

Ninaus, M., Kiili, K., McMullen, J., & Moeller, K. (2016). A game-based approach to examining students' conceptual knowledge of fractions. *International Conference on Games and Learning Alliance*, 37–49.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Wood, G., … Salvo, P. (2015). Game elements improve performance in a working memory training task INESC-ID and Instituto Superior Técnico , Universidade de Lisboa Laboratory of Brain – Computer Interfaces , Institute for Knowledge Discovery,. *Institutional Journal of Serious Games*, *2*(1), 3–16.

Nottbusch, G. (2008). *Handschriftliche Sprachproduktion: sprachstrukturelle und ontogenetische Aspekte* (Vol. 524). Walter de Gruyter.

Nuerk, H., Moeller, K., Klein, E., Willmes, K., & Fischer, M. H. (2011). Extending the Mental Number Line A Review of Multi-Digit Number Processing. *Journal of Psychology*, *219*(1), 3–22.

Papathanasiou, I., & Coppens, P. (2012). Aphasia and related neurogenic communication disorders: basic concepts and operational definitions. *I. Papathanasiou, P. Coppens, C. Potagas. Aphasia and Related Neurogenic Communication Disorders. Burlington: Jones & Barlett Learning*.

Parsons, S., & Brynner, J. (2005). Does numeracy matter more? *NRDC - National Research and Development Centre for Adult Literacy and Numeracy, Institute of Education - University of London,*.

Patterson, K., & Wing, A. M. (1989). Processes in Handwriting: A Case for Case. *Cognitive Neuropsychology*, *6*(1), 1–23.

Peters, L., & De Smedt, B. (2018). Arithmetic in the developing brain: A review of brain imaging studies. *Developmental Cognitive Neuroscience*, *30*, 265–279.

Pinet, S., Ziegler, J. C., & Alario, F. X. (2016). Typing is writing : Linguistic properties modulate typing execution. *Psychonomic Bulletin & Review*, *23*(6), 1898–1906.

Planton, S., Jucla, M., Roux, F. E., & Démonet, J. F. (2013). The “handwriting brain”: A meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex*, *49*(10), 2772–2787.

Poupon, C., Clark, C. A., Frouin, V., Régis, J., Bloch, I., Le Bihan, D., & Mangin, J. F. (2000). Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *NeuroImage*, *12*(2), 184–195.

Power and Maria Felicita Dal Martel, R. (1997). From 834 to Eighty Thirty Four: The Reading of Arabic Numerals by Seven-year-old Children. *Mathematical Cognition*.

Purcell, J. J., Napoliello, E. M., & Eden, G. F. (2011). A combined fMRI study of typed spelling and reading. *NeuroImage*, *55*(2), 750–762.

Purcell, J. J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, *2*, 1–16.

Rapcsak, S. Z., & Beeson, P. M. (2000). Agraphia. In S. E. Nadeau, L. J. Rothi, & B. Crosson (Eds.), *Aphasia and language: Theory to practice* (Vol. 1, pp. 184–220). Guilford Press.

Rapcsak, S. Z., Beeson, P. M., & Hillis, A. (2002). Neuroanatomical correlates of spelling and writing. *The Handbook of Adult Language Disorders*, 71–99.

Rapp, B., & Caramazza, A. (1997). From graphemes to abstract letter shapes: levels of representation in written spelling. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 1130.

Rapp, B., Epstein, C., & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology*, *19*(1), 1–29.

Rath, D., Domahs, F., Dressel, K., Claros-Salinas, D., Klein, E., Willmes, K., & Krinzinger, H. (2015). Patterns of linguistic and numerical performance in aphasia. *Behavioral and Brain Functions*, *11*(1), 1–24.

Roesch, S., Jung, S., Huber, S., Artemenko, C., Bahnmueller, J., Heller, J., … Moeller, K. (2016). Training arithmetic and orthography on a web-based and socially-interactive learning platform. *International Journal of Education and Information Technologies*, *10*, 204–217.

Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature Reviews Neuroscience*, *5*(10), 812–819.

Roux, F. E., Draper, L., Köpke, B., & Démonet, J. F. (2010). Who actually read Exner? Returning to the source of the frontal “ writing centre” hypothesis. *Cortex*, *46*(9), 1204–1210.

Roux, S., McKeeff, T. J., Grosjacques, G., Afonso, O., & Kandel, S. (2013). The interaction between central and peripheral processes in handwriting production. *Cognition*, *127*(2), 235–241.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, *6*(1), 1–36.

Russell, M. (1999). Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, *7*(20).

Russell, M., & Plati, T. (2002). Does it matter with what I write? Comparing performance on paper, computer and portable writing devices. *Current Issues in Education*, *5*(4), 1–29.

Saffran, E. M. (1982). Neuropsychological approaches to the study of language. *British Journal of Psychology*, *73*(3), 317–337.

Sage, K., & Ellis, A. W. (2004). Lexical influences in graphemic buffer disorder. *Cognitive Neuropsychology*, *21*(2–4), 381–400.

Schmandt-Besserat, D., & Erard, M. (2008). Origins and forms of writing. *Handbook of Research on Writing: History, Society, School, Individual, Text*, *30*, 7–22.

Schonauer, K., & Denes, G. (1994). Graphemic Jargon: A case report.pdf. *Brain and Language*, *47*, 279–299.

Schubert, T. M. (2017). Why are digits easier to identify than letters? *Neuropsychologia*, *95*(Ccd), 136–155.

Schubert, T., & McCloskey, M. (2013). Prelexical representations and processes in reading: Evidence from acquired dyslexia. *Cognitive Neuropsychology*, *30*(6), 360–395.

Sereno, S. C., O'Donnell, P. J., & Sereno, M. E. (2009). Size matters: Bigger is faster. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1115–1122.

Shallice, T. (1981). Phonological agraphia and the lexical route in writing. *Brain: A Journal of Neurology*, *104*(3), 413–429.

Starrfelt, R. (2007). Selective alexia and agraphia sparing numbers-a case study. *Brain and Language*, *102*(1), 52–63.

Tainturier, M.-J., & Rapp, B. (2001). The spelling process. *The Handbook of Cognitive Neuropsychology: What Deficits Reveal about the Human Mind*, 263–289.

Teulings, H.-L. (1996). Chapter 10 Handwriting movement control. In *Handbook of Perception and Action*.

Tohgi, H., Saitoh, K., Takahashi, S., Takahashi, H., Utsugisawa, K., Yonezawa, H., … Sasaki, T. (1995). Agraphia and acalculia after a left prefrontal (F1, F2) infarction. *Journal of Neurology Neurosurgery and Psychiatry*, *58*(5), 629–632.

Verdú, E., Regueras, L. M., Verdú, M. J., De Castro, J. P., & Pérez, M. Á. (2008). An analysis of the research on adaptive learning: the next generation of e-learning. *WSEAS Transactions on Information Science and Applications*, *5*(6), 859–868.

Verguts, T., & Fias, W. (2006). Lexical and syntactic structures in a connectionist model of reading multi-digit numbers. *Connection Science*, *18*(3), 265–283.

Wang, S., Hong Jiao, Young, M. J., Brooks, T., & Olson, J. (2007). A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, *67*(2), 219–238.

Wang, S., Hong Jiao, Young, M. J., Brooks, T., & Olson, J. (2008). Comparability of Computer-Based and Paper-and-Pencil Testing in K–12 Reading Assessments. *Educational and Psychological Measurement*, *68*(1), 5–24.

Warrington, E. K., & Shallice, T. I. M. (1979). Semantic access dyslexia. *Brain: A Journal of Neurology*, *102*(1), 43–63.

Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the ventral pathway got lost--And what its recovery might mean. *Brain and Language*, *118*(1–2), 29–39.

Will, U., Nottbusch, G., & Weingarten, R. (2007). Linguistic units in word typing: Effects of word presentation modes and typing delay. *Written Language & LiteracyWritten Language and Literacy*, *9*(1), 153–176.

Wilshire, C. E. (2008). Cognitive neuropsychological approaches to word production in aphasia: Beyond boxes and arrows. *Aphasiology*, *22*(10), 1019–1053.

Wollscheid, S., Sjaastad, J., & Tømte, C. (2016). The impact of digital devices vs. Pen(cil) and paper on primary school students' writing skills - A research review. *Computers & Education*, *95*, 19–35.

Wollscheid, S., Sjaastad, J., Tømte, C., & Løver, N. (2016). The effect of pen and paper or tablet computer on early writing - A pilot study. *Computers & Education*, *98*, 77–80.

Zhang, J., & Norman, D. A. (1995). The Representation of Numbers. *Cognition*, *57*(8), 271–295.

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H. C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77.

Danksagung

Allen, die mich im Laufe meiner Dissertation begleitet und auf die ein oder andere Weise unterstützt haben, möchte ich an dieser Stelle danken:

Großer Dank gebührt Jürgen Heller, der mir als klinisch-tätige Sprachwissenschaftlerin einen Platz in der sehr theoretischen mathematischen Psychologie-Welt gewährt hat und damit das methodische Fundament für diese interdisziplinäre Arbeit bereitete. Ich danke ihm auch für die Freiheit bei der Ausgestaltung meiner Dissertation.

Mein aufrichtiger Dank gilt Korbinian Möller dafür, dass er mir das Tor zur Welt der Zahlen weit aufgemacht hat. Ich danke ihm für seine fachliche Wertschätzung auch in holprigen Zeiten, besonders für seine unvergleichbar schnellen, konstruktiven sowie stets motivierenden Rückmeldungen und Hinweise weit über die Dissertation hinaus.

Klaus Willmes danke ich fürs Brückenbauen und seine Bereitschaft für meine Verteidigung den weiten Weg aus Aachen auf sich zu nehmen. Damit hat er mich nicht nur beim Abschluss meines Studiums sondern auch meiner Dissertation begleitet.

Elise Klein, die ich nun schon so lange kenne, danke ich für viele gemeinsame Erfahrungen im Publizieren, Antrag schreiben, auf Tagungen fahren, Vortrag proben, Tränen trocknen, Daumen drücken und noch so viele Dinge mehr.

Nora Umbach danke ich für die Zeit, die sie sich für Gespräche rund um die Dissertation und weit darüber hinaus genommen hat, für ihre wertvollen Kommentare zu einem ersten Entwurf dieser Arbeit und auch für die Feststellung, dass Emanzipation ein langer Prozess sein kann.

Mein tiefster Dank gilt meiner Familie, deren Halt und Unterstützung ich mir stets auch aus großer Ferne bewusst bin.

Zuletzt danke ich meinem lieben Mann Thomas, mein unerschütterlicher Fels in der Brandung, der unaufhörlich daran geglaubt hat, dass diese Dissertation zu einem guten Ende kommt. Er hat Recht behalten.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise in einer anderen Prüfung noch nicht vorgelegt worden; auch wurde mit dieser Arbeit oder einer anderen Dissertation noch kein Promotionsversuch unternommen.

Tübingen, 27.06.2019 Stefanie Jung