# Mindless Reading and Eye Movements

## Theory, Experiments and Computational Modeling

Daniel J. Schad

Doctoral Thesis

Submitted to the Faculty of Human Sciences at the University of Potsdam
(Department of Psychology, Experimental and Biological Psychology)
in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
Advisors: Prof. Dr. Ralf Engbert, Prof. Dr. Reinhold Kliegl

Potsdam, 2012

Gutachter der Dissertation:
Prof. Dr. Ralf Engbert
Prof. Dr. Albrecht Inhoff

Datum der mündlichen Prüfung: 17. 01. 2013

# Acknowledgements

I am very grateful to my advisor Ralf Engbert for the opportunity to be part of an outstanding lab at the University of Potsdam and for the chance to learn and participate in an exciting research program on mathematical modeling of complex dynamical cognitive systems. I thank him for the good and valuable time I had preparing this dissertation, which to me has opened up much new intellectual space and scientific opportunities. I especially thank him for his support, solidarity, promotion, and advice, for teaching me scientific ideals and skills from modeling to scientific writing, and for the freedom and invaluable opportunities he opened up for me for learning and for exploring ideas. Also, I am very grateful to my second advisor Reinhold Kliegl for being a mentor for me. Since I was an undergraduate student, he kept supporting me in my scientific live, and this has been of great importance to me. I also thank him for his visions, strivings, and believes, which to me have provided strong challenges and incentives, and for the chance to learn from his energy to discover, communicate, and encourage new intellectual frontiers. I am deeply indebted to Frank Rösler for his very generous support, both financially and personally, and for invaluable opportunities to participate in a very exciting and enjoyable working group. I thank him a lot for many interesting and enjoyable discussions and a very clarifying view on psychology and neuroscience. I thank Antje Nuthmann a lot for her help with scientific writing, for the opportunity to work in an interesting project, and for her hospitality in Edinburgh. I thank all of my colleagues in a great lab for the positive and productive atmosphere. I especially thank Jochen Laubrock (for his help with networks and programming), Eike Richter, Michael Dambacher, Sven Hohnstein, Odir Rodríguez Villagra, Sven Ohl, Konstantin Mergenthaler, Christiane Wotschack, Olaf Dimigen, André Krügel, and Felix Engelmann for the intellectual exchange about and their great help with diverse questions related to eye movements, statistical/mathematical modeling, experiments, psychology, and many other important and not so important things in life. I thank Kathrin Göthe, Julian Heister, Anja Genth, Mario Bettenbühl, Marco Rusconi, Elke Lange, Anja Sperlich, and others for many enjoyable hours in and outside the lab. I am very thankful to Hans Trukenbrod and Sarah Risse for their friendship, help and support, and for sharing many things in life and science.

I am thankful to the members of the DFG Forschergruppe FOR 868 Research Unit "Computational Modeling of Behavioral, Cognitive, and Neural Dynamics" for a splendid environment enabling very stimulating cross-disciplinary contacts with colleagues in theoretical physics, mathematics and linguistics, for funding, and for a wonderful and exciting stay at Beijing. I thank the Deutsche Forschungsgemeinschaft (grant nos. EN 471/2 "Eye movements during mindless reading: Experiments and computational modeling") for support in a great project, and the Graduiertenkolleg GRK 1668/1 ("Intrapersonale Entwicklungsrisiken des Kindes- und Jugendalters in längsschnittlicher Sicht") and the German National Academy of Sciences Leopoldina for funding. I thank Nicole Düngel, Petra Köhler, and Peggy Sinske for so much help in dealing with bureaucratic questions and for always being so friendly. I thank Petra Schienmann and the eye-tracking lab for their dedication and for help, patience and endurance with collecting data, particularly in a seemingly endless eye-tracking experiment. Moreover, I thank Eva Hasselmann, Max Geisthardt, Peter Haffke, Sandra Beyerman, Linda Weiß, Elisabeth Jünger and others for their help with constructing text material, programming, data collection, and manual pre-processing of massive amounts of data. I thank Kai Würzner and the dlex-project for providing psycholinguistic norms, and Bruno Schwenk for teaching me concepts in Bayesian statistics. I thank the R Development Core Team and many contributors for providing great software. I thank Keith Rayner, Tim Slattery, Bernhard Angele, and the other members of the San Diego lab for being wonderful hosts and for making our visit an extremely enjoyable and insightful time. I very much thank Jonathan Smallwood for his visit in Potsdam, and not only for splendid insights into the nature of and questions about mind wandering, but also for changing some of my basic views of the cooperative aspects of science.

I thank all of my friends for many good times, for always being around, and for so much support and patience in the last months and years. Thank you. I feel lucky that I have met Milena and I thank her for many extremely fruitful discussions and insights during the progress of this thesis, and for her emotional support. I thank Lukas and Noah, and I thank my parents and my family for supporting me in difficult times, and for always being there.

# Abstract

It sometimes happens that we finish reading a passage of text just to realize that we have no idea what we just read. During these episodes of mindless reading our mind is elsewhere yet the eyes still move across the text. The phenomenon of mindless reading is common and seems to be widely recognized in lay psychology. However, the scientific investigation of mindless reading has long been underdeveloped. Recent progress in research on mindless reading has been based on self-report measures and on treating it as an all-or-none phenomenon (dichotomy-hypothesis). Here, we introduce the levels-of-inattention hypothesis proposing that mindless reading is graded and occurs at different levels of cognitive processing. Moreover, we introduce two new behavioral paradigms to study mindless reading at different levels in the eye-tracking laboratory.

First (Chapter 2), we introduce shuffled text reading as a paradigm to approximate states of weak mindless reading experimentally and compare it to reading of normal text. Results from statistical analyses of eye movements that subjects perform in this task qualitatively support the 'mindless' hypothesis that cognitive influences on eye movements are reduced and the 'foveal load' hypothesis that the response of the zoom lens of attention to local text difficulty is enhanced when reading shuffled text. We introduce and validate an advanced version of the SWIFT model (SWIFT 3) incorporating the zoom lens of attention (Chapter 3) and use it to explain eye movements during shuffled text reading. Simulations of the SWIFT 3 model provide fully quantitative support for the 'mindless' and the 'foveal load' hypothesis. They moreover demonstrate that the zoom lens is an important concept to explain eye movements across reading and mindless reading tasks.

Second (Chapter 4), we introduce the sustained attention to stimulus task (SAST) to catch episodes when external attention spontaneously lapses (i.e., attentional decoupling or mind wandering) via the overlooking of errors in the text and via signal detection analyses of error detection. Analyses of eye movements in the SAST revealed reduced influences from cognitive text processing during mindless reading. Based on these findings, we demonstrate that it is possible to predict states of mindless reading from eye movement recordings online. That cognition is not always needed to move the eyes supports autonomous mechanisms for saccade initiation. Results from analyses of error detection

and eye movements provide support to our levels-of-inattention hypothesis that errors at different levels of the text assess different levels of decoupling. Analyses of pupil size in the SAST (Chapter 5) provide further support to the levels of inattention hypothesis and to the decoupling hypothesis that off-line thought is a distinct mode of cognitive functioning that demands cognitive resources and is associated with deep levels of decoupling.

The present work demonstrates that the elusive phenomenon of mindless reading can be vigorously investigated in the cognitive laboratory and further incorporated in the theoretical framework of cognitive science.

## Zusammenfassung

Beim Lesen passiert es manchmal dass wir zum Ende einer Textpassage gelangen und dabei plötzlich bemerken dass wir keinerlei Erinnerung daran haben was wir soeben gelesen haben. In solchen Momenten von gedankenverlorenem Lesen ist unser Geist abwesend, aber die Augen bewegen sich dennoch über den Text. Das Phänomen des gedankenverlorenen Lesens ist weit verbreitet und scheint in der Laienpsychologie allgemein anerkannt zu sein. Die wissenschaftliche Untersuchung von gedankenverlorenem Lesen war jedoch lange Zeit unzureichend entwickelt. Neuerer Forschungsfortschritt basierte darauf gedankenverlorenes Lesen durch Selbstberichte zu untersuchen und als ein Phänomen zu behandeln das entweder ganz oder gar nicht auftritt (Dichotomie-Hypothese). Hier stellen wir die ‚Stufen der Unaufmerksamkeit'-Hypothese auf, dass gedankenverlorenes Lesen ein graduelles Phänomen ist, das auf verschiedenen kognitiven Verarbeitungsstufen entsteht. Wir stellen zudem zwei neue Verhaltensparadigmen vor um verschiedene Stufen von gedankenverlorenem Lesen im Augenbewegungslabor zu untersuchen.

Als erstes (in Kapitel 2) stellen wir das Lesen von verwürfeltem Text vor als ein Paradigma um Zustände von schwach gedankenverlorenem Lesen experimentell anzunähern, und vergleichen es mit dem Lesen von normalem Text. Die Ergebnisse von statistischen Augenbewegungsanalysen unterstützen qualitativ die ‚Unaufmerksamkeits'-Hypothese, dass kognitive Einflüsse auf Augenbewegungen beim Lesen von verwürfeltem Text reduziert ist, und die ‚Foveale Beanspruchungs'-Hypothese, dass die Reaktion der *zoom lens* visueller Aufmerksamkeit auf lokale Textschwierigkeit beim Lesen von verwürfeltem Text verstärkt ist. Wir stellen eine weiterentwickelte Version des SWIFT Modells (SWIFT 3) vor, welches die *zoom lens* der Aufmerksamkeit implementiert, und validieren dieses Modell am Lesen von verwürfeltem und normalem Text (Kapitel 3). Simulationen des SWIFT 3 Modells unterstützen die ‚Unaufmerksamkeits' und die ‚Foveal Beanspruchungs'-Hypothese in einem vollständig quantitativen Modell. Zudem zeigen sie, dass die *zoom lens* der Aufmerksamkeit ein wichtiges Konzept ist um Augenbewegungen in Aufgaben zum Lesen und gedankenverlorenen Lesen zu erklären.

Als zweites (Kapitel 4) stellen wir den *sustained attention to stimulus task* (SAST) vor um Episoden von spontaner externer Unaufmerksamkeit (also Entkopplung der Aufmerksamkeit oder Abschweifen der Gedanken) in einem Paradigma über Verhaltensparameter wie das Übersehen von Fehlern im Text und Signal-Detektions-Analysen von Fehlerentdeckung zu messen. Augenbewegungsanalysen im SAST decken abgeschwächte Einflüsse von kognitiver Textverarbeitung während gedankenverlorenem Lesen auf. Basierend auf diesen Befunden zeigen wir, dass es möglich ist Zustände von gedankenverlorenem Lesen online, also während dem Lesen, aus Augenbewegungen vorherzusagen bzw. abzulesen. Dass höhere Kognition nicht immer notwendig ist um die Augen zu bewegen unterstützt zudem autonome Mechanismen der Sakkadeninitiierung. Ergebnisse aus Analysen von Fehlerdetektion und Augenbewegungen unterstützen unsere ‚Stufen der Unaufmerksamkeit'-Hypothese, dass Fehler auf verschiedenen Textebenen verschiedene Stufen von Entkopplung messen. Analysen der Pupillengröße im SAST (Kapitel 5) bieten weitere Unterstützung für die ‚Stufen der Unaufmerksamkeit'-Hypothese, sowie für die Entkopplungs-Hypothese, dass abschweifende Gedanken eine abgegrenzte kognitiver Funktionsweise darstellen, welche kognitive Ressourcen benötigt und mit tiefen Stufen von Unaufmerksamkeit zusammenhängt.

Die aktuelle Arbeit zeigt, dass das flüchtige Phänomen des gedankenverlorenen Lesens im kognitiven Labor mit strengen Methoden untersucht und weitergehend in den theoretischen Rahmen der Kognitionswissenschaft eingefügt werden kann.

# Table of content

## List of original publications

The dissertation is based on the following original research articles.

- Schad, D. J., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research, 50*(23), 2600-2616. doi: 10.1016/J.Visres.2010.08.005

- Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*(4-5, Special Issue on Computational Approaches to Reading and Scene Perception), 391. doi: 10.1080/13506285.2012.670143

- Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*(2), 179-194. doi: 10.1016/j.cognition.2012.07.004

- Schad, D.J. (2012). *Mental effort during mindless reading? Pupil fluctuations indicate internal processing during levels of inattention.* Unpublished manuscript, University of Potsdam, Potsdam, Germany.

# List of Figures

# List of Tables

# 1 Introduction

„Most readers have probably had the experience of moving their eyes across text while at the same time their mind wandered so that nothing was comprehended from the text. This "daydream mode" would be very difficult to study experimentally."

(Rayner & Fischer, 1996, p. 746)

## 1.1 Motivation and basic concepts

### 1.1.1 Mindless reading

It sometimes happens that we finish reading a passage of text just to realize that we have no idea what we just read. During these episodes of mindless reading or mind wandering our mind is elsewhere yet the eyes still move across the text. The phenomenon of mindless reading is common. Nearly half of our waking lives we spend with thoughts that are unrelated to the events occurring in the present external environment (Kane et al., 2007; Killingsworth & Gilbert, 2010), and even during active performance of the challenging task of reading our minds frequently wander to other thoughts (Grodsky & Giambra, 1990; Schooler, Reichle, & Halpern, 2004), like thinking about how to solve a puzzling problem or how our beloved ones are currently doing.

The phenomenon of mind wandering has long been ignored in the cognitive sciences. For example, in two important reviews on eye movements and attention during reading and other tasks (Rayner, 1998, 2009) there was not a single mention of the phenomenon of mindless reading or mind wandering. Thus, very little is known about what happens in the mind when readers are not focused on understanding the text. This, however, is unfortunate and a better understanding of mindless reading would be desirable.

First, mindless reading is an interesting everyday phenomenon with important practical consequences. Mindless reading is an important factor causing failures of text comprehension (Smallwood, McSpadden, & Schooler, 2008) and was long overlooked in the study of educational performance (Smallwood, Fishman, & Schooler, 2007). Moreover, mind wandering seems to make people unhappy (Killingsworth & Gilbert, 2010) and may help explaining why tests of working memory and fluid intelligence have so powerful predictive utility (Mrazek et al., 2012). Thus, understanding why and how mindless reading occurs, what positive and negative consequences result from this state of mind, and how it can be prevented or optimized are important research questions (Smallwood, in press). To investigate these questions, good experimental measures are needed to assess mindless reading.

Second, studying mindless reading may contribute to an understanding of the general phenomenon of mind wandering (cf. Smallwood, 2011b). It is an interesting aspect of mind wandering that external attention spontaneously ebbs and flows during normal task performance. During episodes of task-focus, attention is directed to the external environment and cognitive processing is coupled to perceptual input. During episodes of mind wandering, however, external attention is reduced and this process of attentional (or perceptual) decoupling impairs the cognitive and neuronal analysis of perceptual information (Schooler et al., 2011; Smallwood & Schooler, 2006). Many previous studies have investigated decoupled processing in simple laboratory tasks, like processing a series of individual digits which are sequentially presented on a computer screen, and important insights have been derived from such studies (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Smallwood, Beach, Schooler, & Handy, 2008). Compared to many simple laboratory tasks, reading engages a large complexity of cognitive processes (e.g., perception, attention, word recognition, syntactic parsing, model building, memory, prediction, motor control) and involves various hierarchical, serial and parallel cognitive processes of varying complexity and automaticity (Engbert, Nuthmann, Richter, & Kliegl, 2005; Graesser, Olde, & Klettke, 2002; Kintsch, 1998; Malmkjaer, 2002; Rayner, 1998, 2009; Rayner & Pollatsek, 1989; Reichle, Warren, & McConnell, 2009). These architectural principles make reading an ideal task to investigate the interplay of different cognitive processes that are involved in decoupling and self-generated thought (cf. Smallwood, 2011b).

Third, a major question and difficulty in the study of mindless reading is to know what happens in a readers' mind when s/he is not paying attention to the text. To approach this fundamental question, I will use eye-tracking technology to record the movements of the eyes during reading. It is a great benefit of eye movements that they reflect moment-to-moment cognitive processing during reading (Rayner, 1998, 2009). Therefore, recording eye movements during mindless reading opens a valuable window to a well-hidden state of mind.

Investigating mindless reading using eye tracking may also shed light on the processes of reading and eye movement control. Various theories have been developed to explain the processes controlling eye movements during reading, and several of these theories have been implemented in mathematical models (e.g., Engbert, et al., 2005; Reichle, Rayner, & Pollatsek, 2003; Reilly & Radach, 2006; Yang, 2006) (for an overview see two special issues on the topic: Henderson, 2012; Reichle, 2006). Overall, these models have been remarkably successful in explaining eye movements during reading (cf. Rayner, 2009), and agreement exists between models concerning many aspects of the reading behavior. With respect to other aspects of eye movement control, however, model assumptions differ strongly. As one important aspect, a long-standing and ongoing debate concerns the role of higher-level cognition and attention as opposed to lower-level visuomotor factors (Inhoff, Topolski, Vitu, & O'Regan, 1993; Nuthmann & Engbert, 2009; Reichle, Pollatsek, & Rayner, 2012; Starr & Rayner, 2001; Vitu, O'Regan, Inhoff, & Topolski, 1995).

Studying mindless reading provides an interesting approach to investigate levels of eye movement control (Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996; Reichle, Reineberg, & Schooler, 2010; Vitu, et al., 1995). During mindless reading the eyes keep moving across the text even when the mind is absent (Reichle, et al., 2010), and this observation suggests that some aspects of reading behavior may be highly automatic. Other aspects of reading, like readers' responses to a very funny or very difficult passage of text, however, may not occur automatically and may quite strongly depend on whether attention is on the text. Studying mindless reading may thus help to distinguish automatic and controlled reading processes, and to inform theories and models of eye movement control.

If mindless reading is indeed common and was acknowledged to be interesting to study previously (Rayner & Fischer, 1996), why then hasn't cognitive science investigated these questions more closely earlier on? Since behaviorism (Skinner, 1986) psychology has often studied the human mind in the stimulus-response paradigm, in the sense that manipulations of external input were related to behavioral output to infer internal processes. Mental states such as daydreaming or mind wandering do not directly arise from specific external events and are not directly linked to obvious external behavior. Together, these characteristics imply that the nature of mind wandering poses challenges to its scientific investigation, and refined experimental techniques are needed to vigorously study the phenomenon.

To investigate mindless reading, researchers have previously followed two principled experimental approaches. One strategy has been to ignore the spontaneous nature of mind wandering thoughts and to approximate mindless reading in non-reading tasks where higher-level cognitive language processing is reduced, but where oculomotor requirements are similar to reading (Inhoff et al., 1993; Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996; Vitu et al., 1995). A second strategy has been to develop techniques that catch episodes of mind wandering when they spontaneously occur during normal reading. These latter techniques were generally based on subjective self-reports, where the phenomenon of mind wandering is described to participants and readers have the task to indicate whether their own mind is or has been wandering (Sayette, Schooler, & Reichle, 2010; Schooler, et al., 2004). Great potential lies in developing methods to catch episodes of spontaneous mind wandering during reading independent of self-reports on subjective mental states, but instead based on objective behavioral recordings (cf. Franklin, Smallwood, & Schooler, 2011).

In summary, studying eye movements during mindless reading is an interesting enterprise from several perspectives. Major current challenges in the field of mindless reading lie in developing and testing (a) good experimental paradigms to study the phenomenon (cf. Smallwood, in press) and (b) a comprehensive theoretical framework for understanding the cognitive processes involved in this elusive state of mind (Smallwood, 2011b). In the present work, together with my colleagues I will make attempts to meet both of these challenges. Moreover, based on insights on mindless reading I will (c) investigate the processes controlling eye movements (Starr & Rayner, 2001). In the upcoming two sections, I will shortly introduce selected concepts and findings from research on mind wandering and eye movement control.

### 1.1.2 Mind wandering and the default mode

In the cognitive sciences, a standard approach to understanding the mind and brain is to investigate the impact of external task and stimulus conditions on individuals' behavioral and neural responses to learn about cognitive processes. However, in recent years, converging findings from neuroscience and psychology have demonstrated that the human mind and brain show spontaneous internal and self-generated activity in the form of (*i*) mind wandering and (*ii*) default mode activity that is unrelated to input from the current external environment (Buckner, Andrews-Hanna, & Schacter, 2008; M. F. Mason et al., 2007; Raichle et al., 2001; Schooler, et al., 2011). First, in the psychology of consciousness it has been an early observation that participants in cognitive experiments reported that they were not paying attention to the experimental task and that their minds were instead wandering to other thoughts, like thinking about yesterdays' dinner or an upcoming exam (Giambra, 1995). This phenomenon of mind wandering has been linked to a mindless response style in the performance of a sustained attention task (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, et al., 1997). Second, independent from these observations neuroscientists studying brain activity based on functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) were puzzled by the recurring finding that activity in certain brain areas was systematically increased during periods of rest relative to periods of active task performance (Shulman et al., 1997). A systematic investigation of this phenomenon resulted in the definition of the default mode network (DMN), which is often thought to reflect intrinsic default activity in the brain (Raichle, 2010; Raichle, et al., 2001).

Bringing these lines of research together, several recent studies have demonstrated that mind wandering as assessed via self-reports or impaired task performance is closely related to activity in the default mode network and in frontal control regions (Christoff, et al., 2009; M. F. Mason, et al., 2007; Weissman, Roberts, Visscher, & Woldorff, 2006), suggesting a default mode of the human mind and brain (Buckner, et al., 2008) where conscious thought is decoupled from the external environment and occupied with inner thoughts and feelings.

# 1. Introduction

Research on mind wandering has often tended to conflate two important aspects of the experience (Smallwood, in press): one question about mind wandering is why it occurs (and why it is terminated), and this question concerns the events that lead to the occurrence of spontaneous default activity and may be related to the function or benefit of mind wandering. Second, it is a puzzling question what cognitive and neuronal processes are involved in the process of mind wandering and how the mind sustains internal thought in the presence of external input[1] (Smallwood, in press). In the present work, I am primarily concerned with the question of how the mind wanders, and with the cognitive processes involved in this state of mind.

The process of mind wandering has been described as involving at least two specific alterations in cognitive processing (Schooler, et al., 2011): conscious processing of internal information and reduced processing of external information. First, during mind wandering attention is often directed inwards to internal information derived from memory, and people entertain thoughts that are unrelated to the current task (task-unrelated thought, TUT) or are independent of the current stimulus (stimulus independent thought, SIT) (Giambra, 1995; Schooler, et al., 2011; Smallwood & Schooler, 2006). An important feature of research on TUT/SIT is that they are usually assessed via self-reports. The fact that individuals can report on these internal thoughts suggests that they are conscious (Smallwood, 2010b), that they become globally available to the system and may occupy a global workspace of consciousness (Baars, 1988; Dehaene & Naccache, 2001; Smallwood, 2010b). Accordingly, mind wandering was recently defined as a cognitive state in which "information generated by the default mode becomes available to consciousness" (Smallwood, 2010b, p. 201). Consistent with this definition, many previous studies have investigated whether mind wandering thoughts are available to meta-consciousness (Reichle, et al., 2010; Sayette, Reichle, & Schooler, 2009; Sayette, et al., 2010; Schooler, 2002; Schooler, et al., 2004).

Second, during mind wandering attention is directed away from the external environment. This process of perceptual or attentional decoupling reduces the cognitive

---

[1] Note that external input can sometimes be massive, as for example, when the mind wanders in the middle of the busy downtown streets of a large city, or in an electronic music club (Forster & Lavie, 2009).

and neuronal analysis of external information and leads to errors in the performance of external tasks (Schooler, et al., 2011; Smallwood, 2011b; Smallwood & Schooler, 2006). Such decoupling has been widely demonstrated in diverse tasks and measures, including impaired performance in tasks of sustained attention (Cheyne, Solman, Carriere, & Smilek, 2009; Manly, et al., 1999; Robertson, et al., 1997; Smallwood et al., 2004), reading (McVay & Kane, 2012b; Schooler, et al., 2004; Smallwood, McSpadden, & Schooler, 2008), or tests of general aptitude (Mrazek, et al., 2012), as well as in reductions of event related potentials (ERPs, in the electroencephalogram, EEG) (Barron, Riby, Greer, & Smallwood, 2011; Kam et al., 2011; O'Connell et al., 2009; Smallwood, Beach, et al., 2008) or blood-oxygen-level-dependent (BOLD) activity in fMRI (Raichle, et al., 2001; Weissman, et al., 2006).

The decoupling hypothesis (Schooler, et al., 2011; Smallwood, 2011b; Smallwood et al., 2011; Smallwood & Schooler, 2006) provides a link between reduced external and increased internal processing. It postulates that internal thought processes during mind wandering consume domain general cognitive resources, for example by occupying the global workspace of consciousness (Smallwood, 2010b). For that reason, the decoupling hypothesis predicts that internal stimulus-independent thought (SIT) is associated with reduced stimulus-dependent thought (SDT; i.e., perceptual decoupling) due to resource competition, and that SIT and SDT reflect two distinct modes of cognitive and neuronal functioning (Fox et al., 2005; Smallwood, et al., 2011).

To investigate mind wandering during reading, the present work will primarily focus on studying and measuring attentional decoupling. Moreover, in chapter 5 I will study how attentional decoupling relates to internal cognitive activity by analyzing pupil size during mindless reading.

### 1.1.3 Eye movements

“One great virtue of eye movement data is that they give a good moment-to-moment indication of cognitive processes during reading” (Rayner, 2009, p. 5). Eye movements therefore provide an ideal window to learn about what happens in the mind during mindless reading. In the following, I will introduce basic concepts and measures of research on eye movements during reading.

During reading, the eyes do not move smoothly across the text. Instead, relatively quick movements called *saccades* can be distinguished from *fixations*, that is, phases of relative rest. During saccades no visual input is extracted from the environment and the eyes are factually blind (Matin, 1974). Saccades fulfill the function of moving the eyes from one location to the next, while visual information processing is taking place during fixations. Saccades are necessary because visual acuity is high only in central vision in the fovea (i.e., the central about 2° of the retina with the highest visual resolution) and rapidly drops off toward parafoveal (2° to 5° around fixation) and peripheral (further than 5° from fixation) vision where visual acuity is heavily reduced. Therefore, saccades are needed to bring peripheral visual stimuli into foveal vision, where during fixations fine visual details, like letters in a text (Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981), can be processed.

Reading saccades in alphabetic languages are relatively short in duration (typically 20-30 ms), and often move the eyes over an average distance of about seven letters; saccade amplitudes, however, strongly vary: saccades sometimes displace the eyes for a few letters only, but can also move the eyes across a whole line of text (e.g., in return sweeps). Reading fixations occur at an average rate of four to five per second. Like saccade amplitudes, also fixation durations vary strongly and can last between less than 50 ms and more than 600 ms (Rayner, 1998, 2009).

Eye movements during reading form complex sequences involving different types of saccades and fixations (cf. Rayner, 1998, 2009). Most saccades during reading move the eyes forward from the currently fixated word *N* to the next word *N*+1. A smaller proportion of saccades moves the eyes to word *N*+2 or even further to the right, resulting in skipping of word *N*+1. Skipping primarily occurs for short words and words of length 2-3 are skipped about 75% of the time, but words with 8 letters or more are nearly never skipped. Many words are fixated only once, but refixations occur on about 15% of the words, resulting in more than one fixation on these words. About 5-15% of all saccades are regressions in which the eyes move backward; in languages like German and English regressions thus move the eyes from right to left within a line or back to previously read lines of text.

Given the different types of saccades that occur during reading, different measures for fixation durations can be computed (Rayner, 1998), and I will shortly define those most used in the present work. Firstpass reading comprises all fixations on a word before any regressive saccades occur to this or previously read words in the text and morepass reading comprises all fixations that are not in the first pass. The duration of the first fixation on a word in firstpass reading irrespective of later eye movements is often described as the first fixation duration (Rayner, 1998). Usage of this term has, however, varied in the past (e.g., Engbert, et al., 2005; Rayner, 1998) and to avoid confusions I will here use the term initial fixation duration instead. The duration of the first of multiple fixations on a word is a subset of initial fixation durations, and is also sometimes described as first fixation duration (Engbert, et al., 2005). Single fixation durations are the subset of initial fixation durations where a word is fixated exactly once in firstpass reading. Gaze duration is the cumulative duration of all fixations on a word in firstpass reading. Total viewing time is the cumulative duration of all fixations on a word during first- and more-pass reading.

In broader perspective, during reading the eyes move across the words and sentences in complex sequences of different types of saccades and fixations of variable durations. This process provides an example for how human behavior frequently evolves over time in a highly complex, dynamic and interactive manner (Schöner, 2008). Facing this complexity is very challenging, but also fascinating and essential when trying to understand the process and function of the human mind. Eye movements during reading provide a very good case to approach this complexity. First, the complexity of eye movements can be well captured via high-resolution eye tracking technology, and this provides an extremely rich and accurate data-base for understanding this complexity and dynamic. Moreover, despite being highly complex, dynamic and chaotic, complexity in eye movements is constrained by the spatial arrangement of the text and the task of reading, which involves a directed progression from the beginning to the end of a sentence. In addition, human language has been extensively investigated and well described in diverse fields like linguistics, philosophy, and psychology. Based on these constraints and the excellent data, successful attempts have been made to develop sophisticated mathematical models of the cognitive processes controlling eye movements during reading (for an overview see two special issues on the topic: Henderson, 2012; Reichle, 2006). As a major aim, these models seek to understand how the chaotic external dynamics observed during reading (and other tasks) are linked to a few general principles of the functioning of the cognitive system.

#### Upcoming sections

In the remainder of this introduction, I will first (Section 1.2) introduce some findings and theoretical accounts of how attention turns away from the external environment during mind wandering. Next (Section 1.3), I will introduce research on eye movement control during reading, by presenting selected concepts, findings, models, and theories. Moreover (Section 1.4), I will consider methods that have previously been used to study mindless reading, and will discuss previous eye movement findings. Last (Section 1.5), I will discuss the main aims of the current work, provide a quick overview of the used methods, and introduce the present experimental and theoretical studies. Note that reviews of previous work are not intended to be comprehensive, but will be selective.

## 1.2 Attentional decoupling at different levels

### 1.2.1 Hierarchical levels of cognitive processing

The human mind can be described as a complex interplay between many different cognitive processes. It is an important aim in the cognitive sciences to understand these processes and how they relate to each other and to the external world, for example, in terms of serial processing bottlenecks, parallel processing, or sequential, hierarchical, and interacting dependencies (Anderson, 2005; Anderson et al., 2004; Craik & Lockhart, 1972; Fodor, 1983; Gazzaniga, 2009; McClelland & Rumelhart, 1981; Meyer & Kieras, 1997b; Navon & Gopher, 1979; Pashler, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998; Shiffrin & Schneider, 1977; Spearman, 1914). How the multiplicity of cognitive processes and resources can be incorporated into a comprehensive theory of mind wandering is a fundamental, yet open theoretical question. Answers to this question are currently beginning to be formulated (Cheyne, et al., 2009; Schooler, et al., 2011; Smallwood, 2011b; Smallwood & Schooler, 2006).

As a fundamental organizing principle, the human mind (and brain) is often viewed as a hierarchy (Cohen, 2000; Craik & Lockhart, 1972; Gazzaniga, 2009). Incoming sensory information sequentially engages different lower and higher levels of stimulus processing, where early low-level stages process more superficial stimulus features, whereas later stages involve increasingly abstract, complex, and integrated information processing at higher levels of analysis. “This conception of a series or hierarchy of processing stages is often referred to as ‘depth of processing’ where greater ‘depth’ implies a greater degree of semantic or cognitive analysis” (Craik & Lockhart, 1972, p. 675).

Hierarchical levels of cognitive processing and representation are particularly prominent in reading, where text is processed at many different levels to generate a high-level understanding of the text (Frawley 2003; Malmkjaer, 2002; Rayner, 1998, 2009). Many theories and models of reading have postulated hierarchical levels of text processing (Engbert, Longtin, & Kliegl, 2002; Graesser, et al., 2002; Kintsch, 1998; Reichle, et al., 1998). For example, in theories of discourse analysis (Graesser, et al., 2002), the *surface code* is a rather superficial representation of the text and contains the exact wording and syntactic structure. From this superficial representation, the explicit meaning of the text is derived at the *text level*, where propositions are extracted and combined with few inferences to link propositions. At the highest level, the abstract content of the story is represented in the *situation model*. Hierarchical cognitive processes have also been implemented in mathematical models of eye movement control during reading (e.g., Engbert, et al., 2002; Engbert, et al., 2005; Reichle, et al., 1998; Reichle, Warren, et al., 2009), and these models have included low-level visual and oculomotor processes, medium-level lexical processes, and influences from higher-level language processing (for more details see Section 1.3). Note that processing at different text levels during reading is often considered interactive: in bottom-up processing information flows from lower visual levels to support higher-level comprehension. During top-down processes, to the contrary, higher levels influence lower processing levels, as, for example, when upcoming words in the text are predicted from the preceding context, which facilitates their lexical processing (cf. Balota, Pollatsek, & Rayner, 1985; Calvo & Meseguer, 2002; Dambacher, Kliegl, Hofmann, & Jacobs, 2006; Dambacher, Rolfs, Göllner, Kliegl, & Jacobs, 2009). Numerous studies have documented that various low-level and high-level variables affect eye movements during reading (Rayner, 1998, 2009), and thus, tracking eye movements provides a tool to study different levels of text processing during reading (cf. Section 1.3).

### 1.2.2 Early versus late attentional selection

It has been a very early and longstanding question in the cognitive sciences how attention filters or attenuates processing of perceptual information at early (Broadbent, 1958; Treisman, 1960) versus late (Deutsch & Deutsch, 1963) processing stages (for review see Driver, 2001). In her attenuation theory, Treisman (1960) emphasized that the two views should not be regarded as a dichotomy, but rather as the endpoints of a continuum. Today, evidence exists for both the early and the late selection view, and there is agreement that selective attention sometimes attenuates incoming information at early, and sometimes at late processing stages (Chun, Golomb, & Turk-Browne, 2011; Lavie, 2005). Lavie's load theory (Lavie, 2005; Lavie, Hirst, de Fockert, & Viding, 2004) is a modern framework to integrate these views into a single theory, and provides a solution to the early- versus late-selection debate. The theory suggests that the stage at which attentional selection operates depends on the kind of attentional load applied: perceptual load attenuates processing at early perceptual stages, whereas central load impairs central processing (and can therefore enhance early perceptual distractor processing).

Studies on early versus late attentional selection usually use external manipulations to vary perceptual or central load experimentally. However, external attention also ebbs and flows spontaneously during normal task performance. During mind wandering external attention is reduced and cognitive processing of perceptual information is impaired at various lower and higher processing levels (Schooler, et al., 2011). Here, I will shortly review selected findings this research (for review also see Smallwood, 2011b).

At a high processing level, mind wandering reduces the encoding of stimulus meaning into memory (Smallwood, McSpadden, & Schooler, 2008). Moreover, it interferes with processes of model building during reading, as readers tend to overlook meaningless text passages when their minds are wandering (Smallwood, Fishman, et al., 2007). A recent study recorded eye movements during mindless reading (Reichle, et al., 2010) and argued that influences of linguistic and lexical variables were reduced during mindless reading, providing support for reduced high-level linguistic and also reduced medium-level lexical processing.

At a medium processing level, mind wandering was found to reduce the P300 component in ERPs (O'Connell, et al., 2009; Smallwood, Beach, et al., 2008), and this finding indicates reduced processing and updating of task-relevant representations (Donchin & Coles, 1988). Interestingly, such reduced P300 processing during mind wandering was not only found for task-relevant, but also for task-irrelevant stimuli (Barron, et al., 2011) (Note, that this finding is theoretically relevant as it suggests that mind wandering does not result from failures of executive control processes to shield attentional focus against distractions from external stimuli, but rather from decoupling of attention from external information).

At an even earlier level, some studies suggest that mind wandering reduces early perceptual processing. A recent ERP-study (Kam, et al., 2011) found interesting evidence that mind wandering even reduces very early perceptual stimulus processing within the first 100 ms after stimulus presentation (P1). Kam et al. (2011) also showed that this early effect was present across (visual and auditory) modalities. This finding that mind wandering reduces low-level perceptual processing is consistent with an fMRI study reporting reduced activity in the visual cortex during attentional lapses (Weissman, et al., 2006). Last, during mind wandering the processing of external information may be reduced at an even earlier level due to physical blocking of perceptual input at sensory endings via blinks (Smilek, Carriere, & Cheyne, 2010b).

An important goal in the present work is to understand how this diverse set of findings can be incorporated into a comprehensive theory of decoupling and mind wandering. In the following section, I will shortly review the cascade model of inattention, which explains decoupling in a hierarchical system. Moreover, I will discuss two hypotheses that are compatible with this model: I will review the dichotomy-hypothesis, suggesting that mindless reading is an all-or-none phenomenon (Reichle, et al., 2010; Smallwood, et al., 2011; Vitu, et al., 1995). Moreover, I will introduce the alternative levels-of-inattention hypothesis proposing that decoupling in a hierarchical cognitive system is of a graded nature.

### 1.2.3 Hypotheses on attentional decoupling

#### *1.2.3.1 Cascade model of inattention*

The cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) proposes a mechanism to explain decoupling in a hierarchical cognitive system. According to the model, mind wandering reduces cognitive stimulus processing at a very early perceptual level (cf. Kam, et al., 2011; Weissman, et al., 2006) and this impairs the quality of perceptual representations. The consequences of such low-level decoupling are proposed to "cascade downward through the cognitive system" (Smallwood, Fishman, et al., 2007, p. 233) and to cause processing failures at higher cognitive levels. Thus, because bottom-up input is impoverished, higher-level analysis of the external environment fails. Based on this mechanism, the model successfully explains why decoupling occurs in a wide range of tasks, such as reading, signal detection and memory encoding. Moreover, the model explains why decoupling occurs at various lower and higher levels of cognitive processing (cf. Smallwood, 2011b; Smallwood, Fishman, et al., 2007) by suggesting that decoupling is a cascaded phenomenon in the human mind and brain.

Importantly, the cascade model of inattention is compatible with two alternative hypotheses on the nature of hierarchical decoupling: decoupling could be either a dichotomous or a graded process. I will outline these hypotheses in the following sections.

#### *1.2.3.2 The dichotomy-hypothesis*

Mind wandering is usually treated as a dichotomy (Schooler, et al., 2011; Smallwood, 2010b; Smallwood, et al., 2011). The decoupling hypothesis (Smallwood, 2010b; Smallwood, et al., 2011; Smallwood & Schooler, 2006) proposes that because internal mind wandering thoughts (SIT/TUT) rely on domain general cognitive processes there should be a trade-off between stimulus-independent thought and stimulus-dependent thought. Accordingly, Smallwood et al. (2011) postulated that both are distinct cognitive states and that transitions between states are sharp and step-like rather than continuously graded. An implicitly dichotomous approach to mind wandering has dominated previous research as many previous studies have measured or analyzed mind wandering as an all-or-none state (e.g., Kane, et al., 2007; Killingsworth & Gilbert, 2010; Sayette, et al., 2009). With respect to levels of processing, the decoupling hypothesis predicts that when the mind entertains internal thoughts, then all hierarchical levels of cognitive processing are decoupled from external information, but when attention is directed outwards then all levels of processing are coupled.

The decoupling hypothesis concerns a trade-off between SIT and SDT. To the contrary, the present work mainly focuses on attentional decoupling of SDT (i.e., attentional disengagement from the external task) and often doesn't use measures for SIT. The dichotomy-hypothesis of decoupling assumes that decoupling is an all-or-none phenomenon, where either all levels of cognitive processing are decoupled, or all levels are coupled (see Figure 1-1A for illustration). Many previous studies have treated attentional decoupling as a dichotomous process (for reviews see Schooler, et al., 2011; Smallwood, 2011b). Importantly for the present work, the dichotomous approach was particularly prominent in studies on mindless reading: studies assessing mind wandering during reading have mostly used dichotomous self-reports (e.g., Reichle, et al., 2010; Sayette, et al., 2009; Sayette, et al., 2010; Schooler, et al., 2004) (for an exception see Franklin, et al., 2011). Similarly, research approximating mindless reading via non-reading tasks (like z-string scanning, see Section 1.4.1) have adopted a dichotomous approach by comparing a mindless reading condition, where language processing is completely absent (e.g., z-string scanning), to the reading of normal text, where all levels of text processing are present (Nuthmann, et al., 2007; Rayner & Fischer, 1996; Vitu, et al., 1995). The dichotomy-hypothesis predicts that intermediate states of weak decoupling, where high-level processes are decoupled but low-level processes are intact, should not occur. In terms of traditional attention research, the dichotomy-hypothesis postulates early attentional selection during mindless reading, but no states of late attentional selection.

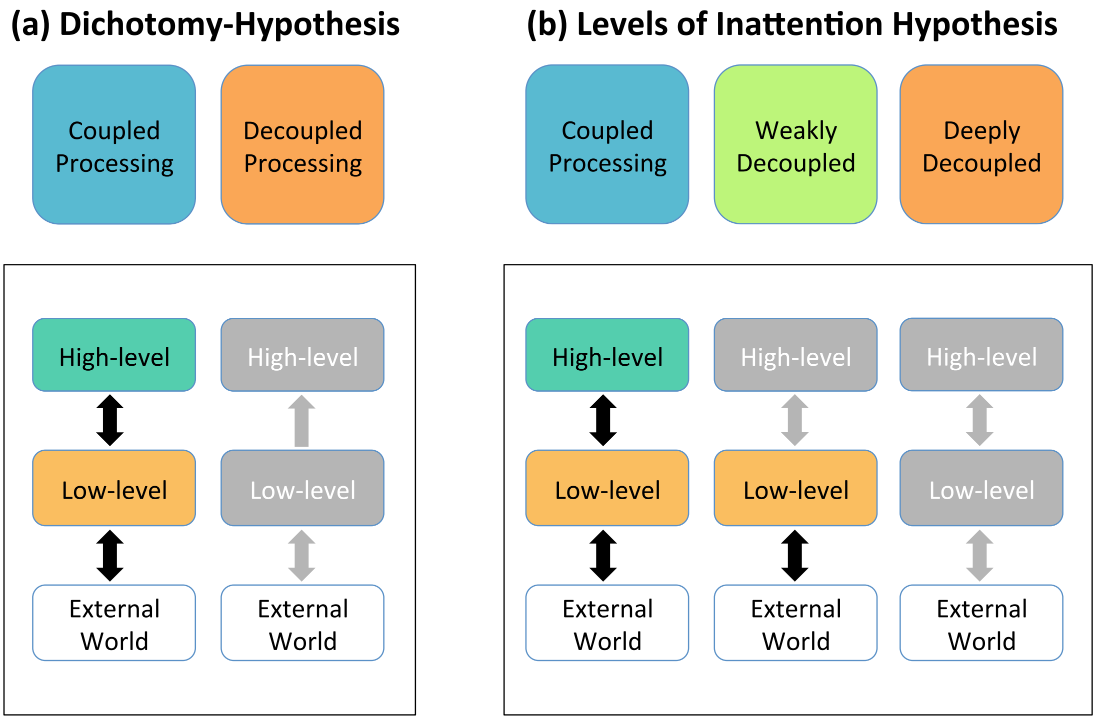


**Figure 1-1**

Schematic illustration of two hypotheses on the nature of attentional decoupling. (a) The dichotomy-hypothesis (left panel) postulates the existence of two discrete attentional states (Smallwood, et al., 2011): during task focus (left column) cognitive processing is fully coupled to the external environment, and low-level (red, middle boxes) as well as high-level (green, upper boxes) processes are coupled (black arrows) to perceptual information (transparent, lower boxes). During mind wandering (right column), to the contrary, high-level and low-level processing is decoupled (grey boxes and arrows) from external input. (b) The levels of inattention hypothesis (right panel) also assumes states of fully coupled processing during task focus (left column) and deeply decoupled processing during mind wandering (right column). Moreover, it assumes different degrees of weak decoupling (middle), where low-level processes are coupled to the external world (red, lower box), but perceptual information is not processed at a high cognitive level (grey upper box). Please note, that even in deeply decoupled states, some low-level processes may remain intact (not shown). Moreover, weak decoupling is not suggested to be a discrete state. Instead, processing failures at many different hierarchical processing levels may constitute different graded degrees of decoupling.

Recent findings on attentional decoupling during mind wandering provide some support to the dichotomy-hypothesis. As I have reviewed above, some very interesting recent studies found that mind wandering attenuates external information from task-relevant and task-irrelevant stimuli, that it attenuates processing across visual and auditory modalities, and that it operates at early perceptual processing stages (Barron, et al., 2011; Kam, et al., 2011; Smilek, et al., 2010b; Weissman, et al., 2006). These findings seem compatible with the view that stimulus-independent thought is generally associated with a state of deep perceptual decoupling (Smallwood, et al., 2011). Moreover, in support of the dichotomy-hypothesis, Smallwood et al. (2011) recently found that a measure of spontaneous cognitive activity (baseline pupil size) and a measure of perceptual decoupling (response times) showed a highly non-linear relationship in the form of a step-function, which is consistent with distinct cognitive states of either off-task or on-task thought.

#### *1.2.3.3 The levels of inattention hypothesis*

As an extension of the dichotomous view on attentional decoupling, I here propose that decoupling and mindless reading are graded phenomena: in the levels of inattention hypothesis (Figure 1-1B) I postulate that cognitive stimulus processing can fail at different hierarchical levels, and that this reflects graded degrees of attentional decoupling. The levels of inattention hypothesis assumes that cognitive processing is sometimes fully focused on the external task, and that it is sometimes deeply decoupled from the external environment. However, it also assumes states of weak attentional decoupling, where low-level processes are fully coupled to the external environment, but higher-level stimulus processing fails and becomes decoupled. The levels of inattention hypothesis is compatible with the existence of early and late attentional selection (Chun, et al., 2011; Lavie, 2005), and suggests that early as well as late attentional selection can occur spontaneously during normal task performance.

This view is consistent with some previous studies suggesting that the phenomenon of decoupling may be more complex than a dichotomy. Cheyne et al. (2009) introduced a model assuming three discrete states of task engagement/disengagement by distinguishing attention to dynamically changing “moment-to-moment stimulus meaning” from attention to the “general task environment”, and from attention to “motor behavior” (Cheyne, et al., 2009, pp. 99-100). According to the model, each of these cognitive processes can be disengaged from the task separately, and this leads to different discrete states of mind wandering. The three-states model has been successful in predicting aspects of task performance in a simple go/no-go task (sustained attention to response task, SART; Cheyne, Carriere, Solman, & Smilek, 2011; Cheyne, et al., 2009). However, it does not consider how inattention occurs at different hierarchical levels of cognitive stimulus processing as they exist – for example – during reading. The notion of graded attentional decoupling may also be consistent with principal component analyses of SART responses (Smallwood, 2010a; Smallwood, McSpadden, Luus, & Schooler, 2008), which have revealed that performance errors were preceded by a continuous fastening of response times.

In the present work I aim at testing predictions derived from the dichotomy-hypothesis against the levels of inattention hypothesis by studying eye movements during mindless reading. Moreover, I will investigate predictions from the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) and from the decoupling hypothesis (Smallwood, 2011b; Smallwood, et al., 2011). Thus, I will study mindless reading to test whether different levels of weak versus deep mindless reading can be distinguished, and whether deep levels of decoupling are associated with internal processing.

## 1.3 Eye movement control during reading

In the upcoming sections, I will provide a selective review of research on eye movement control during reading. I will review findings on cognitive influences on eye movements during reading (Section 1.3.1), theories and models of eye movement control, including a conceptual introduction to one existing mathematical model of eye movement control during reading (the SWIFT model, Engbert et al., 2005) (Section 1.3.2) and some currently debated questions in the field that are relevant for the present work (Section 1.3.3).

### 1.3.1 Cognitive influences on eye movements

Research on eye movements during reading has made much progress in the last decades (for reviews see: Rayner, 1998, 2009). Many previous studies have demonstrated that eye movements are very sensitive to the moment-to-moment cognitive activity during reading, making eye movements an ideal tool to investigate cognitive processing. Evidence for cognitive influences on eye movements comes from a large research field investigating how many variables at different levels of cognitive text processing do (and do not) affect eye movements during reading (for reviews see: Rayner, 1998, 2009), and some of these levels include orthographic, morphological, phonological, lexical, syntactic, semantic, and discourse processing. Cognitive influences have been examined for diverse aspects of eye movements such as measures of fixation durations, saccade amplitudes, word skippings, refixations, regressions, and within-word landing sites. Here, I will shortly introduce a few selected findings and concepts from this research.

#### *1.3.1.1 Influences of linguistic and lexical variables on eye movements*

Much previous research has investigated how variables related to word recognition and lexical processing affect eye movements during reading. Most prominently, the (printed) *word frequency effect* (word frequency denotes how often a word occurs in a language) is one of the most robust and replicated effects in reading research: low frequency words are generally fixated longer than high frequency words (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner & Duffy, 1986) (for reviews see: Rayner, 1998, 2009). Word frequency effects have been found in various tasks involving word recognition (e.g., reading, lexical decision, semantic classification, perceptual identification, spoken word recognition) and are often interpreted as reflecting *lexical processing*, that is, getting access to the entry of a word in the mental lexicon (Monsell, 1991; Norris, 2006; Whaley, 1978).

*Word length* is another variable that strongly and reliably affects eye movements during reading (Just & Carpenter, 1980; Kliegl, et al., 2004; Rayner, Sereno, & Raney, 1996). Readers look longer at long words than at short words, they make more refixations on long words, and they skip long words less often than short words (Brysbaert & Vitu, 1998; Rayner, 1979, 1998). As one explanation for these effects, long words extend further into the visual periphery as compared to short words, which may impair visual and/or lexical processing for long words (Engbert, et al., 2005; Reichle, et al., 1998). It has also been found that word length interacts with word frequency, where frequency effects are stronger in long words than in short words (Kliegl, Nuthmann, & Engbert, 2006; Pollatsek, Reichle, Juhasz, Machacek, & Rayner, 2008; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999).

At higher (post-lexical) levels of processing, words are integrated in their context at syntactic, semantic, and discourse levels, and many studies have investigated how higher-level language processing affects eye movements (Clifton, Staub, & Rayner, 2007; Staub, 2011). An important variable associated with higher-level cognitive processing is *word predictability*, reflecting the fact that readers are able to predict upcoming words in the text from the preceding context. Empirically, word predictability is often assessed via a cloze task (Ehrlich & Rayner, 1981), where the first part of a sentence is shown to a subject and the subject is asked to guess the next word in the sentence, which is not presented. Word predictability affects eye movements during reading, with longer fixations on low-predictable than on high-predictable words (Balota, et al., 1985; Drieghe, Brysbaert, Desmet, & De Baecke, 2004; Ehrlich & Rayner, 1981; Zola, 1984).

Another influence from higher-level language processing on eye movements are *wrap-up effects*. It has been an early observation in research on eye movements during reading that readers look longer at the last word of a phrase or a sentence compared to other words in a text (Just & Carpenter, 1980). This and later findings together suggest that readers take some time to integrate the words from a phrase or sentence into a coherent cognitive representation before moving on in the text (Kuperman, Dambacher, Nuthmann, & Kliegl, 2010; Warren, White, & Reichle, 2009).

#### 1.3.1.2 Deciding ‘when’ and ‘where’ to move the eyes

It is often observed that low-level and high-level factors tend to affect different aspects of eye movements (cf. Starr & Rayner, 2001). Many studies support the assumption that the “decision” of *when* to move the eyes is influenced by higher-level cognitive language processing. As I have discussed above (Section 1.3.1.1), many studies have documented influences from numerous variables related to text processing at different cognitive levels, and these influences are particularly prominent for measures of fixation durations (Rayner, 1998, 2009). More specifically, fixation durations in first-pass reading are often influenced by factors related to word recognition. Measures of later eye movements (starting with gaze durations, but also including regression probability, total reading time, and fixation durations in larger regions of the text) are usually more strongly affected by higher-level language comprehension (Clifton, et al., 2007).

To the contrary, the ”decision” of *where* to move the eyes (particularly within a word) is often more strongly controlled by low-level visual or oculomotor factors. For example, readers most often fixate words at the preferred viewing location (PVL), which is located between the middle and the beginning of a word (McConkie, Kerr, Reddix, & Zola, 1988; Rayner, 1979). The (Gaussian) distributions of within-word landing positions underlie control by visual and/or oculomotor factors (Engbert & Krügel, 2010; Krügel & Engbert, 2010; McConkie, et al., 1988; Rayner, 1998), but are largely independent of lexical or linguistic variables (Rayner, Binder, Ashby, & Pollatsek, 2001; Rayner, et al., 1996) and are highly similar in normal and mindless reading paradigms (Nuthmann, Engbert, & Kliegl, 2005; Rayner & Fischer, 1996; Vitu, et al., 1995), supporting their independence from higher-level processing.

### 1.3.2 Theories and models of eye movement control

Based on the wealth of experimental findings, several theories have been developed and implemented in mathematical models to explain the cognitive processes controlling eye movements during reading (for an overview of models see: Henderson, 2012; Reichle, 2006). These models have been very successful in explaining diverse aspects of eye movement behavior. Despite this success, some questions on the control of eye movements have remained debated (see e.g., Starr & Rayner, 2001, for discussion).

In the following sections, I will discuss different principled theories of the eye-mind link (Section 1.3.2.1) and some benefits of using mathematical modeling in cognitive research (Section 1.3.2.2). Moreover, I will introduce basic concepts from a mathematical model of eye movement control (SWIFT, Engbert, et al., 2005) (Section 1.3.2.3).

#### *1.3.2.1 Theories of the eye-mind link*

An important and open theoretical question in theories and models of eye movement control during reading concerns the relative extent to which eye movements are controlled by lower-level visuomotor factors as opposed to higher-level cognitive processing, and different classes of models or general theories can be distinguished based on their assumptions about the role of cognitive processing (cf. Starr & Rayner, 2001). Theories and models can be viewed as lying on a continuum ranging from *primary oculomotor control* (POC) models, via *indirect cognitive control* models, to *direct cognitive control* models (cf. Rayner & Pollatsek, 1981; Reichle, et al., 2003; Trukenbrod & Engbert, 2009; Trukenbrod & Engbert, submitted).

In *POC* models, visual and oculomotor factors primarily determine the movements of the eyes, and cognition has little local influence on eye movements (Yang, 2006; Yang & McConkie, 2001). *Cognitive processing models* have been classified as 'direct' or 'indirect', and these two terms have been used to describe either the mechanisms or the timing of cognitive control (Schad, Risse, Slattery, & Rayner, 2012). Here, I am concerned with the mechanisms linking cognition to eye movements. *Cognitive trigger theory* describes a mechanism for direct cognitive control models (like E-Z Reader, Reichle, et al., 1998; Reichle, et al., 2012; Reichle, et al., 2003) and assumes that a high-level cognitive process triggers new saccade programs. In the E-Z Reader model, for example, the completion of an early phase of lexical processing triggers new saccade programs (Reichle et al., 1998; Reichle et al., 2003). Cognitive trigger theory proposes that lexical word processing is "the 'engine' driving eye movements during reading" (Reichle, et al., 2003, p. 459). *Cognitive modulation theory*, to the contrary, defines a mechanism for indirect cognitive control models (like SWIFT: Engbert & Kliegl, 2001; Engbert, et al., 2002; Engbert, et al., 2005; or Glenmore: Reilly & Radach, 2006) and thus takes an intermediate position between direct cognitive control and POC models. Cognitive modulation theory postulates the existence of an autonomous motor timer that initiates saccades at regular time intervals "in the absence of lexical processing demands" (Engbert, et al., 2005, p. 792). Cognition (e.g., lexical processing) does not directly trigger new saccades in this framework, but can modulate (e.g., inhibit) the progression of the autonomous timing process.

Mathematical models implementing POC, cognitive trigger theory, and cognitive modulation theory have been successful in explaining diverse aspects of eye movement behavior during reading and other tasks (e.g., Engbert & Kliegl, 2011; Engbert, et al., 2005; Reichle, 2006, 2011; Reichle, et al., 2003). However, which class of models best describes the processes generating eye movements during reading is unclear at present, and further evidence on this question is desirable.

#### *1.3.2.2 Some benefits of mathematical modeling*

„Ich behaupte aber, daß in jeder besonderen Naturlehre nur so viel eigentliche Wissenschaft angetroffen werden könne, als darin Mathematik anzutreffen ist.“

(Kant, 1786, p. A VIII)

[„I assert, however, that in any special doctrine of nature there can be only as much proper science as there is mathematics therein.“, translation: (Friedman, 2004, p. 6)]

Mathematical models provide a very valuable tool for studying the mind for several reasons (May, 2004; McClelland, 2009), and I will discuss a few of these reasons here. First, building mathematical models forces researchers to be highly explicit about the assumptions they make about the cognitive processes in question. It turns out that, sometimes, valid predictions from theoretical concepts for experimental data are possible only when several aspects of a system are specified simultaneously. This is particularly true for highly interactive control processes, as for example when predictions from a mechanism of refixations depend on the processes generating word skipping (Engbert, et al., 2005).

Based on mathematical models it is possible to perform Monte Carlo model simulations to generate consistent theoretical predictions, which can be directly compared to outcomes of cognitive experiments. Generating consistent theoretical predictions is particularly important when the investigated behavior is generated by complex interactions between different cognitive processes, or when the behavioral outcomes themselves are very complex. Such complexity, for example, is present in eye movements during reading, where current eye movements depend on an event history of previous and/or upcoming saccades (e.g., see Chapter 3). Interestingly, models of eye movement control during reading generate the same type of output as human subjects (i.e., sequences of eye movements) and model predictions can therefore be compared with observed data using diverse and advanced statistical measures. Such analyses can therefore test theoretical predictions at levels of detail that can impossibly be obtained by pure verbal reasoning

(Kliegl & Engbert, in press). Therefore, mathematical models are ideal tools to test the explanatory power of theoretical concepts.

Ideally, mathematical models also serve a heuristic function by generating new ideas for experimental hypotheses and by stimulating new research. Indeed, differences between models of eye movement control during reading (e.g., Engbert, et al., 2005; Reichle, et al., 2003) have generated much effort to resolve theoretical discrepancies empirically (e.g., Inhoff, Radach, & Eiter, 2006; Kliegl, et al., 2006; Reichle, Liversedge, Pollatsek, & Rayner, 2009). Thus, because mathematical models are vigorous implementations of cognitive theories they have the potential to elicit, frame and structure scientific debates.

#### *1.3.2.3 SWIFT: A mathematical model of saccade generation*

A prominent model of eye movement control during reading is the SWIFT[2] model of saccade generation (Engbert, et al., 2002; Engbert, Longtin, & Kliegl, 2004; Engbert, et al., 2005; Laubrock, Kliegl, & Engbert, 2006; Nuthmann & Engbert, 2009; Richter, Engbert, & Kliegl, 2006). SWIFT belongs to the class of models assuming a processing gradient (PG) and implementing cognitive modulation theory, two core assumptions that set it clearly apart from its main competitor, the E-Z Reader model (Reichle, et al., 1998; Reichle, et al., 2003; Reichle, Warren, et al., 2009).

A comprehensive description of all model mechanisms (see Engbert, et al., 2005) is clearly beyond the scope of this introduction. Here, I will shortly introduce some selected components of the SWIFT model. For the present work on mindless reading, I will put a special emphasis on mechanisms and parameters in SWIFT that capture higher-level cognitive influences on eye movement control. Since the publication of the last version of the SWIFT model (SWIFT II, Engbert, et al., 2005) an advanced and as hitherto unpublished model version has been developed (SWIFT 3). Ralf Engbert and I will introduce this advanced version of SWIFT in Chapter 3 of the present work. In the following, I will briefly describe some of the mechanisms that have been incorporated in SWIFT 3.

The SWIFT model is based on several core cognitive principles. A central concept is a set of spatially distributed word-based activations, which controls saccade target selection.

---

[2] **S**accade initiation **w**ith **i**nhibition from **f**oveal **t**argets.

Moreover, an autonomous (random) saccade timer initiates new saccade programs, and an inhibitory control process, called foveal inhibition, delays saccade initiation based on foveal processing difficulties. A key motivation for the development of SWIFT was to implement an integrative framework generating all types of saccades (i.e., forward, skipping, refixation saccades, regressions) observed during reading. In analogy to the dynamic field theory (Erlhagen & Schöner, 2002) the word-based activation field determines probabilities for target selection at any point in time. These processes constitute partial independence of the temporal ('when') and spatial ('where') decisions about the movements of the eyes (Findlay & Walker, 1999).

A main objective in the development of SWIFT has been to model the interplay between visual word recognition and lower-level eye movement control processes during reading. So far, assumptions about higher-level language comprehension are rather limited in the model, and the model could and probably will be extended in the future to capture such influences. In SWIFT, the processing of words occurs in a one-dimensional dynamic field of activations, which is interpreted as a dynamic visual saliency map. The activation field dynamically evolves over time under the influence of several cognitive processes like visual attention, current fixation position, word recognition, and post-lexical processing. Activation of a word increases in a preprocessing stage, and for reading this increase is based on lexical processing. After word activation has reached it's maximum, it decreases and finally returns to zero. (During reading this decrease may be associated with post-lexical processing.) Model parameters capture different aspects of this process. The lexical difficulty of a word is estimated from the word's frequency and determines the word's maximum activation. Word predictability is implemented as a separate process that affects the processing rate of a word, that is, how quickly the word activation is rising and falling. An important concept in dynamic field theory (Erlhagen & Schöner, 2002) is global inhibition of activations, and this is implemented in SWIFT 3 as word activations reduce the processing rate for upcoming words.

In the SWIFT model, word processing depends on visual attention. The allocation of visual attention along a gradient is implemented as a processing span, which defines the letter-based processing rate as a function of the eccentricity relative to the current fixation location. In SWIFT 2, the processing span was defined via an asymmetric Gaussian distribution. In SWIFT 3, this assumption was changed and the zoom lens model of attention (Eriksen & St. James, 1986) was incorporated (see Chapter 3 for details): the processing span is now defined as a negative quadratic function, which dynamically changes in size as a function of foveal word activation. During preprocessing of the foveal word, that is, before word activation has reached its maximum, the processing span is symmetric and in a state of a focused zoom lens. After the maximum foveal word activation has been reached, the processing span dynamically extends further to the right and this process results in an asymmetric processing span and in a defocused zoom lens. Based on the zoom lens overall processing resources are assumed to be constant throughout this dynamic process. As an additional assumption, 15 letters to the right of fixation receive minimal preprocessing. In SWIFT, allocation of visual attention according to the processing span defines processing rates for individual letters. These are then combined to determine the word-based processing rate, and word length plays an important part in this integration.

SWIFT explains saccade generation via an autonomous saccade timer, which initiates saccade programs at random time intervals. In SWIFT 3 this process is implemented as a directed discrete-state random walk. The field of word-based activations exerts an influence on this autonomous timing process via foveal inhibition: foveal activations inhibit the progression of the random walk and thus delay the initiation of the next saccade. Moreover, as a central model concept, the relative activation of a word determines the probability that this word is selected as the next saccade target. Based on these mechanisms, the SWIFT model implements highly dynamic interactions between eye movements and cognitive processing. For example, fixation positions strongly affect word recognition, which in turn influences the activation field, which in turn determines visual attention and the ‘where’ and ‘when’ of future eye movements.

The SWIFT model moreover implements various assumptions about lower-level oculomotor processes, and I will just name a few of these assumptions here. SWIFT incorporates (a) two-stage saccade programming with labile and nonlabile levels (Becker & Jürgens, 1979), (b) systematic and random errors in saccade lengths (McConkie, et al., 1988), (c) error correction for mislocated fixations (Nuthmann, et al., 2005), (d) modulation of saccade latency by intended saccade amplitude (e.g., Wyman & Steinman, 1973). As a general expectation, these low-level oculomotor processes should not be strongly affected during mindless reading.

Technically, SWIFT 3 is a highly non-linear dynamical systems model. It is implemented as a set of coupled ordinary differential equations, and stochastic variables are computed as independent parallel discrete-state random walk processes, which are exactly computed in continuous time based on the Gillespie algorithm (Gillespie, 1977, 1978).

### 1.3.3 Some current questions in eye movement control

A main focus of the present work lies on investigating how mindless reading affects cognitive control of eye movements. However, it also makes contributions to other current questions in research on eye movement control during reading. In the following, I will shortly introduce some of these questions.

#### *1.3.3.1 Spatially distributed processing in the perceptual span*

The perceptual span is the region of effective vision from which "skilled readers pick up various types of visual information during a fixation while reading" (McConkie & Rayner, 1975, p. 578). To study the perceptual span, McConkie & Rayner (1975) introduced the moving window paradigm, where a mask that moves with the eyes covers the text such that only the currently fixated part of the text is visible to the reader. Research with the moving window paradigm found that the word identification span extends 7-8 letters to the right and 3-4 letters to the left (McConkie & Rayner, 1975; Rayner, 1998; Rayner, Well, Pollatsek, & Bertera, 1982; Underwood & McConkie, 1985).

An alternative approach to studying the perceptual span has been to investigate how spatially adjacent words influence fixation durations on a fixated word *N* (e.g., Kliegl, 2007; Kliegl, et al., 2006). It has long been known in reading research that lexical properties of the fixated word *N* (like a word's frequency) influence fixation durations on this word (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986), and this finding is consistent with the immediacy-assumption that current-word processing determines fixation durations (Just & Carpenter, 1980; Morrison, 1984). Moreover, evidence suggests that processing of the previous word *N*-1 often spills over into fixations on word *N*, resulting in lag-effects (Rayner & Duffy, 1986; Schroyens, et al., 1999). In recent years, an intense controversy has evolved around the question in how far words to the right of the fixated word influence current fixation durations on the fixated word *N*, reflecting parafoveal-on-foveal (PoF) or successor effects. Several studies have found that lexical properties of word *N*+1 (e.g., Kennedy & Pynte, 2005; Kliegl, 2007; Kliegl, et al., 2006; Pynte, Kennedy, & Ducrot, 2004) and even properties of word *N*+2 (Kliegl, Risse, & Laubrock, 2007; Risse & Kliegl, 2011) influence fixation durations on the fixated word *N*, and these findings suggest spatially distributed word processing during reading. However, some of these findings have been debated due to their corpus-analytic approach (Rayner, Pollatsek, Drieghe, Slattery, & Reichle, 2007) and because other experimental studies did not find the effects (e.g., Henderson & Ferreira, 1993; Rayner, Juhasz, & Brown, 2007) (for review see Rayner, 2009). At present, large-scale studies are desirable that test lexical PoF effects based on experimental manipulations of successor-word frequency, and I present such a study in Chapter 2. Theoretically, successor effects are highly interesting for theories and models of eye movement control assuming serial versus parallel word processing during reading.

#### 1.3.3.2 Serial versus parallel word processing

How is visual attention allocated to the text to support lexical processing during reading? *Serial attention shift* models (SAS) like E-Z Reader (Reichle, et al., 1998; Reichle, et al., 2003; Reichle, Warren, et al., 2009) propose that the spotlight of attention strictly serially focuses on one word at a time. Shifts of the attention spotlight occur when lexical processing of the attended word has been completed. To the contrary, *processing gradient* models (PG) propose that attention is distributed along an attentional gradient, which can span across several words and supports parallel lexical processing of more than one word at a time (e.g., SWIFT: Engbert, et al., 2002; Engbert, et al., 2005; Glenmore: Reilly & Radach, 2006). The question of serial versus parallel attention allocation during reading has attracted much attention and has been strongly debated in recent years (e.g., Inhoff, Eiter, & Radach, 2005; Kliegl, et al., 2006; Pollatsek, Reichle, & Rayner, 2006b; Rayner, Pollatsek, et al., 2007; Reichle, Liversedge, et al., 2009).

In the present work, I will introduce an advanced version of a PG model assuming parallel word processing (SWIFT 3; cf. Engbert, et al., 2002; Engbert, et al., 2005) and use it to derive and test qualitative as well as quantitative predictions for eye movement experiments (see Chapters 2 and 3). These analyses are limited to delineating an explanation based on the SWIFT model. Analyses based on different conceptions of eye movement control – including SAS models – would provide interesting additional perspectives.

#### 1.3.3.3 The zoom lens of attention and the foveal load hypothesis

An important concept of visual attention that has been neglected in mathematical models of eye movement control is the *zoom lens model* (Eriksen & St. James, 1986) (however, see Reilly & Radach, 2003; Reilly & Radach, 2006). According to the zoom lens model, the focus of attention can change in size and vary between sharply focusing on a small region and being widely distributed across a larger area of the visual field. In the zoom lens model, overall processing resources are assumed to be constant. Therefore, focusing the zoom lens of attention increases the locally available cognitive resources and speeds the processing of information presented in the attended region of the visual field.

In eye movement control during reading, the zoom lens of attention is related to the *foveal load hypothesis* (Henderson & Ferreira, 1990). The foveal load hypothesis postulates that foveal processing difficulties dynamically modulate the size of the perceptual span during reading: the perceptual span is assumed to be small for difficult fixated words, and to be large for easy fixated word. Empirical support for the foveal load hypothesis comes from studies using the boundary paradigm (Rayner, 1975). In this paradigm, a target word in a sentence is masked at the beginning of a trial. When the reader's eyes cross an invisible boundary that is located prior to the target word, then the mask is removed and the target word becomes visible. As the main finding, fixation durations on the target word are found to be shorter when no mask was presented prior to fixating the target word, and are found to be longer when the target word was masked, and this differences reflects the preview benefit effect (Rayner, 1975). In support of the foveal load hypothesis, Henderson and Ferreira (1990) found that the preview benefit effect was larger when the pre-target word was easy (i.e., a high frequency word or a sentence with low syntactic complexity) as compared to when it was difficult to process (low frequency word or high syntactic complexity).

The zoom lens of attention provides an explanation for the foveal load hypothesis by assuming that foveal processing difficulties modulate the focus of the zoom lens. This view assumes that during fixations on difficult words attention is narrowly focused and supports processing of only one word at a time. During fixations on easy words, to the contrary, attention is more widely distributed and upcoming words in the text are processed in parallel (see e.g., Kliegl, et al., 2006). To my knowledge, the zoom lens model of visual attention has not been implemented in mathematical models to simulate eye movements during reading (but see Reilly & Radach, 2003; Reilly & Radach, 2006). Together with Ralf Engbert, I will introduce such an implementation in Chapter 3 of the present work. Alternative explanations of foveal load effects based on SAS models have been proposed (see Henderson & Ferreira, 1990; Reichle, et al., 1998), and I will discuss the explanation based on the E-Z Reader model (Reichle, et al., 1998) in more detail in the General Discussion (Section 6.3.2).

#### 1.3.3.4 Skipping costs and benefits

An empirical question that has been tightly linked to and motivated by mathematical eye movement models concerns the durations of fixations prior to skipping saccades versus non-skipping saccades. SAS models like E-Z Reader (Reichle, et al., 2003) and PG models like SWIFT (Engbert, et al., 2005) implement different mechanisms to explain why the eyes sometimes skip words in the text, and differential predictions can be derived from these mechanisms. In E-Z Reader, lexical skipping is generated if a saccade to the next word $N+1$ is cancelled and replaced by a saccade program aiming at word $N+2$. Such cancellation occurs only when word $N+1$ is substantially processed (i.e., when L1 is completed) while the eyes still fixate on word $N$. Therefore, the model predicts longer average fixation durations before word skipping than before non-skipping, and skipping costs have been obtained in simulations of the E-Z Reader model (Reichle, et al., 2003). Due to the dynamical nature of the SWIFT model, predictions from SWIFT for fixation durations before skipping are difficult to derive without performing numerical model simulations. Simulations of the

SWIFT 2 model have shown that SWIFT 2 also predicts skipping costs (Engbert, et al., 2005).

Empirically, many studies have reported mixed results for fixation durations before skipping versus non-skipping: while some studies have observed longer fixation durations before skipping (skipping costs), others have observed no effect of skipping, or even shorter fixations before skipping (i.e., skipping benefits) (Drieghe, et al., 2004; Hogaboam, 1983; Kliegl, 2007; McConkie, Kerr, & Dyre, 1994; Pollatsek, Rayner, & Balota, 1986; Pynte, et al., 2004; Radach & Heller, 2000). Kliegl and Engbert (2005) proposed a resolution to the debate. In a highly controlled statistical analysis they found that word length, frequency, and predictability of the skipped word strongly modulated the observed skipping costs and benefits: they found skipping costs for the skipping of long, low frequency, and low predictable words, but reliable skipping benefits for short, high-frequency, and high predictable words. These results suggest that inconsistencies in previous findings may exist because previous studies did not control for the critical variables modulating the effects.

Current implementations of both SAS and PG models are compatible with the skipping costs observed for long words (Engbert, et al., 2005; Reichle, et al., 2003). However, to my knowledge no eye movement model exists at present that can explain skipping benefits for short words. In the present thesis (Chapter 3), Ralf Engbert and I will show that an advanced version of the SWIFT model (SWIFT 3), incorporating the zoom lens of attention, successfully explains skipping benefits and costs in normal and shuffled text reading.

#### *1.3.3.5 Reading versus non-reading tasks: Toward general mechanisms of eye movement control*

Modeling the cognitive processes controlling eye movements has been quite successful for the task of reading in the last decade, where several models have been developed that can reproduce many aspects of the eye movement data (see, e.g., a special issue in Cognitive Systems Research, Reichle, 2006). It has been argued that modeling eye movements in other tasks, like scene perception or visual search, has been less advanced compared to reading (Rayner, 2009; Reichle, et al., 2012). Therefore, extending reading models to other tasks may be beneficial. Such extensions pose the important question whether mechanisms implemented in models of eye movement control during reading are specific to the reading task, or whether they represent task-independent general mechanisms of eye movement control (for discussions see Engbert, et al., 2005; Nuthmann & Engbert, 2009; Reichle, et al., 2012).

Motivated by these questions, recent work has started to apply models of eye movement control during reading to explain eye movements in other, non-reading tasks (Nuthmann & Engbert, 2009; Reichle, et al., 2012; Salvucci, 2001). These studies showed that models of eye movement control during reading have the capability of simulating aspects of eye movements during non-reading tasks. This demonstrates the feasibility of generalizing reading models to explain eye movements in other tasks, and provides interesting starting points for a more detailed exploration of the task-specific versus domain-general nature of eye movement control processes. In the present work, my colleague and I will contribute to this beginning enterprise: we generalize a model that has been developed to explain eye movements during reading normal text (SWIFT) to explain eye movements during the "reading" or scanning of randomly shuffled text (i.e., random word lists, see Chapter 3) (also see Nuthmann & Engbert, 2009).

## 1.4 Mindless reading: Methods and eye movement findings

The phenomenon of mind wandering during reading is ubiquitous in everyday life, but due to its spontaneous nature it is difficult to induce and study experimentally (cf. Rayner & Fischer, 1996) with legal measures. To investigate mindless reading in the laboratory a variety of paradigms have been developed. In the following, I will discuss the general approaches to approximate mindless reading via non-reading tasks (Section 1.4.1), to design paradigms that catch episodes of mind wandering (or attentional lapses) during normal reading (Section 1.4.2), and to develop objective (behavioral or physiological) online-markers for mind wandering (Section 1.4.3). Eye tracking has previously been used in combination with the first two approaches. Here, I will discuss what conclusions can be drawn from these findings about whether cognition is needed to initiate saccades (Section 1.4.4).

### 1.4.1 Approximating mindless reading via non-reading tasks

It seems intuitively plausible that when the mind wanders away from the text during reading higher-level language comprehension ceases, while low-level oculomotor processes remain relatively unaffected. To approximate this state of mind, several studies have investigated non-reading tasks where higher-level cognitive language processing is reduced, but where low-level visual and oculomotor requirements are similar to normal reading. Some studies have used designs were language processing is removed (or reduced) by impoverishing aspects of the stimulus material. In the influential z-string scanning paradigm (Inhoff et al., 1993; Vitu et al., 1995), all letters in a text are replaced by the letter 'z'. For example, the sentence "The mind, which wanders …" would be replaced by "Zzz zzzz, zzzzz zzzzzzz …". Participants have the task to move their eyes across the z-strings "as if they were reading" (Vitu, et al., 1995, p. 355). Eye movements in this task have been compared to eye movements during normal text reading to learn about cognitive factors in eye movement control (Drieghe, Brysbaert, & Desmet, 2005; Inhoff et al., 1993; Liversedge et al., 2004; Nuthmann et al., 2007; Rayner & Fischer, 1996; Vitu et al., 1995). A second way to approximate mindless reading via non-reading tasks has been to use normal text material, but to change task instructions such that participants focus on aspects of the text that are unrelated to higher-level cognitive language comprehension. Previous paradigms involve repeated reading of the same text (Inhoff et al., 1993; Schnitzer & Kowler, 2006), the visual search for a target letter (Vitu, et al., 1995) or a target word (Rayner & Fischer, 1996; Rayner & Raney, 1996), as well as proof-reading of a text containing spelling errors (Wotschack, 2009).

### 1.4.2 Catching episodes of mind wandering

Several recent studies have used methods to catch episodes when mind wandering spontaneously occurs in the laboratory.

#### *1.4.2.1 Self-report measures*

A prominent way to catch episodes of mind wandering is via self-report (Schooler, et al., 2011; Smallwood & Schooler, 2006). Self-report measures of mind wandering often implement a thought sampling procedure, where participants are intermittently asked to report on their inner experience in a controlled experiment. Thought-sampling has become a standard measure when studying mind wandering. It is employed in a rapidly increasing number of studies on diverse aspects related to mind wandering (for reviews see Schooler, et al., 2011; Smallwood & Schooler, 2006), and has also been the most widely used method for studying mind wandering during reading (Reichle, et al., 2010; Sayette, et al., 2009; Sayette, et al., 2010; Schooler, et al., 2004; Smilek, et al., 2010b; Uzzaman & Joordens, 2011).

However, there are some potential problems associated with using self-reports as the basis for cognitive research. It has long been known that subjective self-reports on one's own mental states and processes (i.e., introspection) are often invalid (Corallo, Sackur, Dehaene, & Sigman, 2008; Johansson, Hall, Sikstrom, & Olsson, 2005; Nisbett & Wilson, 1977). Findings from a classic study (Nisbett & Wilson, 1977) suggest that, for example, participants are sometimes completely unaware of the cognitive processes allowing them to find the solution to a problem, or of the true causes of their choice behavior. Moreover, a recent study suggests that self-reports about why an individual has previously chosen a certain option are highly unreliable as individuals introspectively generate causes even for choices that they have never taken, but are lead to believe to have taken (Johansson, et al., 2005). Thus, when people reason about the causes of their choices, the reasons they come up with may be computed completely independent of the real factors causing the choice. Even if such strong dissociations between actual mental processes and reports about these processes should be the exception rather than the rule, introspective self-reports still seems unreliable. At least, their validity may strongly depend on the specific instructions used in a given experiment (Overgaard, Rote, Mouridsen, & Ramsoy, 2006). As an additional aspect, introspecting on one's own cognitive processes may not only be unreliable – it may also change behavior. Instructing participants to report on their mind wandering may fundamentally change the way these participants perform the task (however, see Schooler, et al., 2004).

These potential problems associated with self-report measures of mind wandering have long been known, and different strategies have been employed to deal with the difficulties (Smallwood & Schooler, 2006). One such strategy has been to validate subjective self-reports of mind wandering using behavioral and neurophysiological measures, and a large number of studies has provided strong validity for the thought-sampling procedure (Christoff, et al., 2009; Reichle, et al., 2010; Sayette, et al., 2009; Sayette, et al., 2010; Schooler, et al., 2011; Smallwood, Beach, et al., 2008; Smallwood & Schooler, 2006). A second strategy has been to derive behavioral measures for mind wandering.

#### 1.4.2.2 Behavioral measures

The most widely used paradigm to measure mind wandering behaviorally is the sustained attention to response task (SART, Manly, et al., 1999; Robertson, et al., 1997), which measures lapses of attention via performance errors in a simple go/no-go task. In the SART, individual stimuli are presented sequentially on a computer screen, and participants have the task to respond to each of the frequent non-target stimuli (e.g., the digits 0 to 9), but to withhold their response to infrequent target stimuli (e.g., the digit 3). Failures to withhold the response to target stimuli in the SART have been associated with attentional lapses during episodes of mind wandering: performance errors in the SART were found to be related to self-reports and behavioral markers of mind wandering (Manly, et al., 1999; Robertson, et al., 1997), to neural signatures of attentional decoupling (Smallwood, Beach, et al., 2008) and to activity in the default mode network and in executive control regions (Christoff, et al., 2009). Performance errors in the SART thus seem to indicate episodes when participants respond in a mindless "stimulus-response, stimulus-response style", without basing their responses on a cognitive analysis of the presented stimuli, and provide a behavioral measure for attentional decoupling (Manly, et al., 1999; Robertson, et al., 1997).

Behavioral measures have also been developed to assess mind wandering during reading. The SART-technology has been adapted to assess attentional lapses during the reading of individual words in the semantic SART task (Smallwood, Riby, Heim, & Davies, 2006). As an interesting approach, Smallwood, Fishman, et al. (2007) report about an unpublished study using self-paced rapid serial visual presentation (RSVP) of individual words during sentence reading, where readers had the task to detect meaningless (gibberish[3]) text that was embedded in a body of normal text. In this study, overlooking the onset of meaningless text was associated with increased self-reports of mind wandering. Moreover, a highly interesting recent study, which was also based on self-paced RSVP,

---

[3] In gibberish text, the order of nouns and pronouns in a sentence are shuffled, resulting in meaningless but grammatically correct sentences. For example, the sentence "*They tried to collect money for the circus*" is replaced by the gibberish version "*They tried to collect circus for the money*".

found that episodes of self-reported mind wandering could be predicted based on reaction times in the RSVP task (Franklin, et al., 2011) (for details, see the next Section 1.4.3).

However, behavioral measures that assess mind wandering during natural text reading are missing to date. This is unfortunate and being able to measure objectively whether readers are currently paying attention to the text in a normal reading task would be desirable. Such measures would provide important tools to study – independent of the problematic aspects of subjective self-reports – for example how eye movements are controlled when the mind is absent during reading.

### 1.4.3 Objective online markers for mind wandering

It is a major obstacle in the investigation of mind wandering that the variable of interest is difficult to induce experimentally. Success in a scientific investigation of mind wandering therefore hinges upon developing good measures for this state of mind (Smallwood, in press). Previous measures (see previous Section 1.4.2) have typically caught episodes of mind wandering by intermittently intervening with task performance, either by asking participants to report about their mental state or by presenting specific target stimuli to assess performance errors (Schooler, et al., 2011). Although this approach has been successfully applied in the study of mind wandering, it has some clear disadvantages: most importantly, in order to learn about a participant's state of mind the experimenter has to intervene with ongoing task performance (or design the task such that these interventions are inherent to the task, as in the SART). Such mind-probing can be executed only intermittently and little is known about what happens in the mind between mind-probes. Moreover, any interference with the ongoing task may also interfere with the phenomenon of interest (Cheyne, et al., 2011; Cheyne, et al., 2009). Critically, it may also change the general way participants perform a task, or may modify ongoing task-related cognitive processes of interest, like processes of text comprehension or memory encoding. Last, and critically, assessing mind wandering by occasionally intervening with task performance provides poor temporal information: it is difficult to know when an episode of mind wandering started or ended, or how long it lasted. This latter limitation is a major roadblock for understanding the cognitive processes that initiate or terminate individual episodes of mind wandering and to clearly distinguish these from the cognitive processes that support the mind wandering state (Smallwood, in press).

An ideal measure of mind wandering, which could avoid these shortcomings, would be to find objective online-markers for mind wandering that are based on continuous behavioral or (neuro)physiological recordings and that allow assessing at each moment in time whether the mind currently wanders or is focused on the task at hand (cf., Franklin, et al., 2011; Reichle, et al., 2010; Smallwood, 2011a; Smallwood, McSpadden, Luus, et al., 2008). Mind wandering is known to influence various continuously recordable variables, like response times (Manly, et al., 1999; Robertson, et al., 1997; Smallwood, 2010a; Smallwood, McSpadden, Luus, et al., 2008), eye movements (Reichle, et al., 2010; Smilek, et al., 2010b; Uzzaman & Joordens, 2011), recordings of EEG (Kam, et al., 2011; Smallwood, Beach, et al., 2008) or fMRI (Christoff, et al., 2009; Weissman, et al., 2006). However, to my knowledge only one previous study has used objective online-markers to actually predict whether participants' minds are on task at a certain moment in time (Franklin, et al., 2011). It is important to note in this context that to reliably predict mind wandering from objective markers at the level of individual trials is very different from (and presumably far more difficult than) the usual finding that mind wandering influences such markers on average across a whole experiment.

Franklin et al. (2011) studied a self-paced word-by-word reading paradigm (RSVP). Based on objective recordings of readers' response times in this reading task they were able to predict the state of self-reported mind wandering at the level of individual trials. Their predictions were successful for passages of difficult text containing many difficult words (i.e., long, multi-syllable, unfamiliar). For these passages, relatively slow responses indicated that readers processed the text at a cognitive level (Just, Carpenter, & Woolley, 1982) and predicted subsequent self-reports of being on-task. Interestingly, fast response times for difficult text passages indicated that readers did not take the time to process the text and predicted states of mind wandering. Based on an independent set of data, Franklin et al. (2011) defined a prediction algorithm that was based on average response times in individual trials. In a second experiment, they used this – a priori defined – algorithm to predict self-reports of mind wandering online, and were able to correctly predict self-reports of mind wandering 72% of the time. In addition, the algorithm predicted participants' performance on questions probing text comprehension. These results are highly interesting and very encouraging as they provide the first demonstration that behavioral markers can be used to predict mind wandering online in a new set of data with a fairly high level of accuracy. This finding, together with findings that mindless reading affects eye movement behavior (Reichle, et al., 2010; Smilek, et al., 2010b; Uzzaman & Joordens, 2011), also suggests that it may be possible to predict states of mind wandering from recording of eye movements during natural reading.

### 1.4.4 Eye movements during mindless reading

During mindless reading, external attention lapses and cognitive text processing is reduced. Studying eye movements during mindless reading provides important information about the cognitive processes controlling eye movements (Reichle, et al., 2010). Current theories and models of eye movement control (see previous Section 1.3.2.1) generally assume that non-cognitive processes trigger some saccades during normal reading, for example when saccades miss their intended target word (Engbert, et al., 2005; Nuthmann, et al., 2007; Reichle, Pollatsek, & Rayner, 2006). However, an important difference between models concerns the question whether higher-level cognition is necessary to generate regular sequences of eye movements during an external task like reading. Strong empirical evidence exists in support of cognitive contributions to eye movement control during normal reading (see previous Section 1.3.1). However, whether the eyes can move across the text without any contributions from higher-level cognition is not clear from the previous findings. Studying eye movements during mindless reading may provide an interesting opportunity to investigate this question.

One approach to study mindless reading in the laboratory has been to approximate mindless reading experimentally via non-reading tasks (like z-string scanning or target letter/word search). In these tasks, higher-level cognitive language processing is reduced, but lower-level visuomotor requirements are similar to normal reading (Inhoff et al., 1993; Nuthmann et al., 2007; Rayner & Fischer, 1996; Vitu et al., 1995). Interestingly, these studies show that many aspects of eye movements are remarkably similar in a 'mindless reading' condition as compared to normal text reading (Nuthmann, et al., 2007; Rayner & Fischer, 1996; Vitu, et al., 1995). For example, distributions of fixation durations, of saccade lengths, of the eyes' initial landing positions within words, the probabilities of skipping and refixating, and the inverted optimal viewing position in these tasks are very similar to those during normal reading. These similarities suggest that some aspects of eye movement control are rather independent of higher-level cognitive language processing. Moreover, based on these similarities it has been argued that low-level visuomotor processing may be the driving force for eye movements during reading, supporting POC models of reading (Vitu, et al., 1995).

However, eye movements in non-reading tasks also differ from normal reading, and these differences point to contributions from higher-level cognitive language processing in the control of eye movements. For example, Rayner & Fischer (1996) found that word frequency effects were absent during target word search. These and other differences between normal reading and non-reading tasks have been taken to support models of cognitive eye movement control (Rayner & Fischer, 1996).

Moreover, from the perspective of cognitive control models, similarities between non-reading tasks and normal reading do not exclude the possibility that higher-level cognition initiates eye movements in these tasks. For example, it may be that an internal high-level controlled and voluntary cognitive process moves the eyes across meaningless stimuli like z-strings. Indeed, simulations of the E-Z Reader model have assumed that an internal cognitive trigger generates new saccade programs during z-string scanning (Reichle, et al., 2012). As an alternative view, however, an autonomous low-level system may move the eyes across the z-strings in the absence of modulations from higher-level cognition (Nuthmann & Engbert, 2009), and the available data does not seem to clearly distinguish between these cognitive and autonomous explanations. The inability of non-reading tasks to differentiate between cognitive and non-cognitive explanations highlights a weakness of this experimental approach: while it successfully ensures that eye movement patterns are independent of external language information, it seems that it cannot exclude the possibility that internal high-level cognitive processing controls eye movements.

An alternative approach to mindless reading has been to catch episodes of spontaneous mind wandering during normal reading. During mind wandering, attention is decoupled from perceptual information, and cognitive control of external behavior ceases

(Schooler, et al., 2011; Smallwood, 2011b). Therefore, studying episodes when attention spontaneously lapses during normal reading may yield better constraints for cognitive mechanisms of eye movement control. A central question for theories and models of eye movement control is whether cognitive processing does or does not control eye movements during such episodes of mind wandering, and in particular whether effects of lexical processing are present or absent (for further discussion see the General Discussion, Section 6.3.1). So far, Reichle et al. (2010) is the only study that has investigated this question.

Reichle et al. (2010) were the first to study eye movements during episodes when the mind wanders during reading. In their study, four subjects read the entirety of the novel Sense and Sensibility by Jane Austen across several experimental sessions. Episodes of mind wandering during reading were assessed via self-reports. Reichle et al. (2010) found longer fixation durations during episodes of mindless reading. This effect was particularly prominent when readers caught themselves mind wandering, indicating they had become aware of their mind wandering thoughts. Two additional studies have since been published assessing eye movements during self-reported mindless reading, and both of these studies did not replicate the finding of prolonged fixation durations (Smilek, et al., 2010b; Uzzaman & Joordens, 2011).

Moreover, Reichle et al. (2010) argued that the influences of lexical (word frequency and length) and linguistic (wrap-up) variables on fixation durations is absent during mindless reading.[4] This argument is consistent with the data from Rayner and Fischer (1996; see also Rayner & Raney, 1996), on target word search. Thus, it is a reasonable hypothesis that the influence of lexical and linguistic variables on fixation durations may be reduced when the mind wanders during reading.

---

[4] To support this argument, Reichle et al. (2010) report results from repeated measures regression analyses of fixation durations: they report significant effects of word frequency, length, and wrap-up effects for normal reading, and report that these effects were not significant during mindless reading. Problematically, however, these null-results are not informative: in their data, there were far less observations of mindless reading than normal reading episodes (2,231 fixations during normal reading versus 307 and 230 fixations in the conditions for aware and unaware mind wandering), and this difference can fully explain why effects were non-significant during mindless reading. Moreover, Reichle et al. (2010) did not report regression coefficients for non-significant effects, nor slope-differences between normal and mindless reading. Based on this information, Reichle et al. (2010) provide no evidence, neither statistical nor numeric, that the influence of lexical and linguistic variables on fixation durations is reduced during mindless reading.

To conclude, previous findings on eye movements during mindless reading suggest that during states of mindless reading low-level visuomotor influences may be intact, but that higher-level lexical and linguistic processes may be reduced. However, direct evidence is missing to date whether lexical processing is absent or present during episodes when external attention lapses during mindless reading. Thus, based on the current state of knowledge it is unclear whether higher-level cognition is needed to move the eyes.

## 1.5 Research aims and overview of the present studies

In the present work, I will study eye movements during mindless reading with three general aims. First, I aim at studying mindless reading as an interesting everyday phenomenon. Most people know the experience of reading, while at the same time they pay no attention to the text and their thoughts drift off towards internal thoughts and feelings. During such states of mind wandering, the eyes can sometimes move across large passages of text, but readers may have no memory whatsoever for these passages (cf., Reichle, et al., 2010; Smallwood, McSpadden, & Schooler, 2008). Here, I aim at contributing to the methodological basis for the scientific investigation of this elusive phenomenon by developing objective behavioral measures for it. Of particular interest, it may (or may not) be possible to use eye tracking to read mindlessness off the readers' eyes in an online fashion without interfering with task performance, and I aim at exploring the feasibility of this possibility. Moreover, based on objective measures I aim at improving the empirical and theoretical understanding of mindless reading and of the cognitive processes involved in this state.

My second major aim in the present work is to contribute to an understanding of the general phenomenon of mind wandering. It is an important aspect of mind wandering that it interferes with the performance of external tasks (Schooler, et al., 2011). The details of this process (known as attentional or perceptual decoupling) are currently beginning to be better understood (e.g., Cheyne, et al., 2009; Schooler, et al., 2011; Smallwood, 2011a; Smallwood, 2011b). In research on eye movement control during reading, a rich empirical and theoretical understanding has been accumulated of the processes that couple the eyes to the text (Rayner, 1998, 2009). This research tradition offers a wealth of knowledge about specific cognitive processes and mechanisms involved in the performance of the external reading task, ranging from diverse low-level visuomotor processes via medium-level word recognition to higher-level text comprehension. By investigating how mind wandering interferes with these different processes (cf. Smallwood, 2011b), I aim at contributing to the conception of the general phenomena of mind wandering and attentional decoupling.

Third, I aim at learning about the processes controlling eye movements during reading. An important interest guiding research on eye movements has been to understand the mechanisms by which eye movements are linked to higher-level cognition, and several theories, implemented in detailed mathematical models, have been developed to explain the eye-mind link (see Henderson, 2012; Reichle, 2006 for special issues on the topic). Studying eye movements during mindless reading provides an interesting possibility to evaluate and constrain assumptions embedded in these models. First, mindless reading provides a baseline to normal reading, where higher levels of cognitive language processing may be reduced, but lower-level visuomotor factors may be similar to normal reading (Nuthmann & Engbert, 2009; Nuthmann, et al., 2005; Rayner & Fischer, 1996; Reichle, et al., 2010; Vitu, et al., 1995). Thus, mindless reading provides an opportunity to learn about the relative influences of low-level and high-level factors. Second, mindless reading allows to investigate whether higher-level cognition is needed to initiate eye movements, a question that has been debated in previous theories. Third, cognitive processing affects the allocation of visual attention during reading, and mindless reading may provide an opportunity to learn about this process. With the present work, I thus aim at contributing to the understanding of the processes controlling eye movements during reading. To conclude, I aim at developing new behavioral measures for mindless reading, and to use these measures to study the nature of mind wandering and of processes of eye movement control.

To approach these aims, I will present four studies containing empirical and theoretical analyses of eye movements during mindless reading, which I conducted in collaboration with my colleagues Ralf Engbert and Antje Nuthmann. Empirically, we will introduce two new paradigms, the reading of shuffled text (Chapters 2 and 3) and the sustained attention to stimulus task (SAST; Chapters 4 and 5), to investigate mindless reading. Both paradigms are designed to study different weak and deep levels of mindless reading. Based on these paradigms, we will report results from three eye movement experiments (two of them involving extensive data collection, see Chapter 4). Theoretically, we will introduce and test the levels of inattention hypothesis (Chapter 4), a new hypothesis on the nature of attentional decoupling, and test predictions from the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) (Chapter 4) and the decoupling hypothesis (Smallwood, 2010b, in press; Smallwood, et al., 2011; Smallwood & Schooler, 2006) (Chapter 5). Moreover, we will introduce an advanced version of a mathematical model of eye movement control during reading (SWIFT 3) incorporating the zoom lens model of visual attention in a model of eye movement control (Chapter 3). We will use the SWIFT model to explain eye movements in normal reading and in shuffled text reading, and investigate mechanisms linking cognitive processing to eye movements and attention (Chapter 3). The general research framework and some central questions in the present work are illustrated in Figure 1-2.

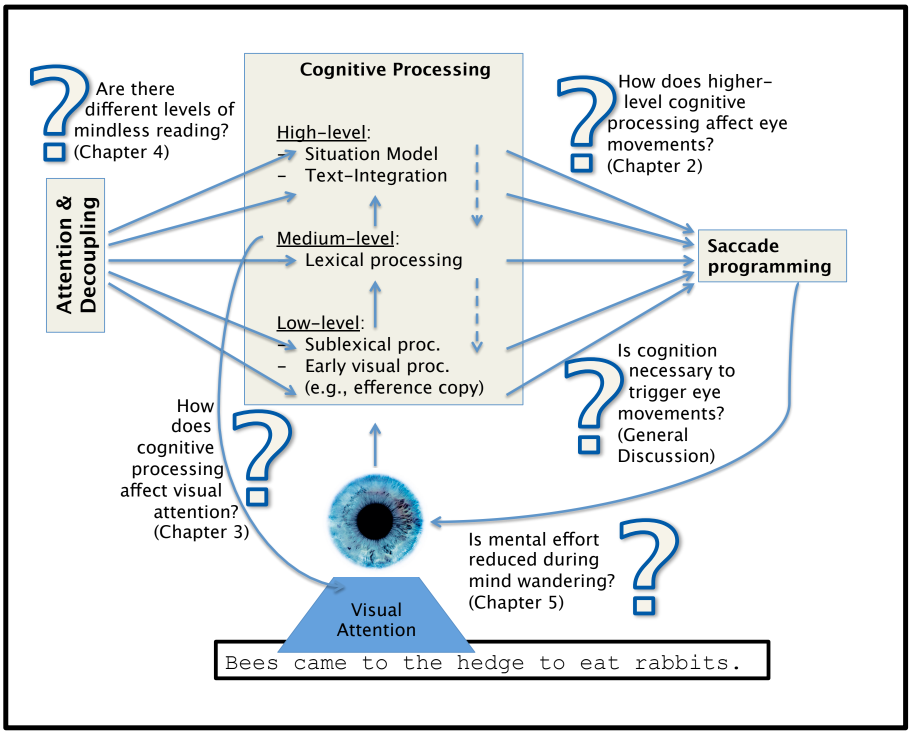


**Figure 1-2**

Schematic overview of research framework.

Methodologically, the research uses (a) state-of-the-art video-based eye tracking technology (using the EyeLink 2 and EyeLink 1000 systems, SR Research) to record saccades, fixations, and pupil size during reading (Chapters 2-5). For statistical analyses, I (b) employ generalized linear mixed effects models [GLMM, Chapters 2, 4, & 5 (Pinheiro & Bates, 2000)] including mixed effects signal detection analyses (Chapter 4) (Wright, Horry, & Skagerberg, 2009) and parametric bootstrapping of a GLMM (Chapter 4). These analyses (c) implement complex design matrices including factor coding via dummy variables, effect coding, Helmert and sliding difference contrasts as well as nesting of factors, covariates, and interactions under experimental conditions (Chapters 2, 4 and 5). I moreover (d) realize a Bayesian logistic regression analysis (Chapter 4) (Gelman, Jakulin, Pittau, & Su, 2008) and (e) use corpus analyses and the analysis of target words (Chapters 2 + 4) (Kliegl, 2007; Kliegl, et al., 2006; Rayner, Pollatsek, et al., 2007). Importantly, I (f) perform Monte Carlo simulations of a (stochastic) nonlinear dynamical system model (Chapter 3) (Engbert, et al., 2005; Glass & Mackey, 1988) [implemented as a system of coupled ordinary differential equations, where several stochastic variables are computed exactly in continuous time via the Gillespie algorithm (Gillespie, 1978)] and (h) estimate model parameters via (restricted) maximum likelihood estimation (for the GLMMs) and a genetic algorithm (for the SWIFT model, Chapter 3) (Goldberg, 1989). The statistical analyses make extensive use of the R System for Statistical Computing (R Development Core Team, 2012), and specifically the packages lme4 (Bates & Sakar, 2008) and ggplot2 (Wickham, 2009). Matlab, including the Psychophysics and Eyelink Toolbox extensions (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002), was used to program eye movement experiments, and for data preprocessing. Last, the SWIFT model – including the model itself, procedures for model simulation, preliminary data analysis, and parameter estimation – is implemented in c (Ritchie, 1993), and a network of Mac Pro computers was used for the parameter estimation. In the following, I will shortly introduce the present studies and summarize the main findings.

### 1.5.1 Shuffled text reading

#### Chapter 2: Eye movements during reading of randomly shuffled text

Based on the levels of inattention hypothesis, during weak mindless reading high-level text comprehension is reduced, but the processing of individual words may be intact. As one potential cause for this dissociation, word processing is highly automatic (Stroop,

1935) and may prevail even when discourse processing fails during weak attentional lapses. In chapter 2, my colleagues and I present the shuffled text reading paradigm as an initial experimental approximation to weak mindless reading. The basic idea behind shuffled text reading is to convert meaningful sentences into meaningless word lists (via random shuffling of word order) and to instruct subjects to read these meaningless lists. During shuffled text reading – as in weak mindless reading – the processing of individual words may be intact, but at the same time higher-level text comprehension is precluded.

In chapter 2, my colleagues and I present results from an eye tracking experiment in which 30 subjects read corpora of shuffled text, and we compare this data to eye movements of 30 subjects reading normal text. To create shuffled text, the order of words in a corpus of normal text was randomly shuffled across all sentences, resulting in random words lists, where all local relations between words were reduced to the level of chance. We compare eye movements in the two tasks using extensive statistical analyses based on generalized linear mixed effects models. We found that effects of word processing on eye movements were reduced during shuffled text reading, supporting the assumption that eye movements during shuffled text reading may be less coupled to cognitive processing. Moreover, we found a dissociation between slightly reversed immediacy effects and standard effects of spatially distributed processing during shuffled text reading and developed an explanation for these findings based on the zoom-lens of attention: we postulate that in shuffled text reading the perceptual span is more strongly dynamically modulated by foveal word processing as compared to normal text reading. Last, we discuss limitations of the current interpretations due to potential confounds, like specific memory processes.

### Chapter 3: The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model

In chapter 3, we test our hypotheses about eye movement control during shuffled text reading by performing fully quantitative simulations of the SWIFT model of saccade generation. We introduce an advanced version of the SWIFT model (SWIFT 3) incorporating the zoom lens model of visual attention via a dynamical modulation of the processing span. Moreover, the SWIFT model contains mechanisms linking eye movements to cognitive word processing. We use these model features to test our hypotheses about (a) reduced cognitive influences in eye movement control and (b) a stronger zoom lens response in shuffled text reading by estimating SWIFT model parameters for eye movements observed in both tasks. Parameter estimates are based on an elaborate split-half validation procedure. The estimated model parameters support our predictions and generate new hypotheses about how eye movement control may differ between normal and shuffled text reading. Based on parameter estimates, we perform Monte Carlo simulations of the SWIFT model to generate model predictions for eye movements during normal and shuffled text reading. As a result, the SWIFT 3 model successfully explains diverse eye movement findings for the reading of normal text, and generalizes to explain eye movements during shuffled text reading. We report simulation experiments where we manually manipulated the zoom lens parameters in SWIFT 3 to generate predictions for shuffled text reading. The results revealed that the zoom lens contributes to explain effects of immediate and spatially distributed word processing, and contributes to the first mathematical explanation of skipping benefits. In summary, the simulations reported in Chapter 3 demonstrate that the SWIFT model of saccade generation, incorporating the zoom lens of attention, provides a viable model of eye movement control during reading.

### 1.5.2 Sustained attention to stimulus task (SAST)

#### Chapter 4: Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels

In chapter 4, my colleagues and I introduce a behavioral paradigm to catch episodes of attentional decoupling (i.e., attentional lapses or mind wandering) during normal reading. The sustained attention to stimulus task (SAST) operationalizes mindless reading as overlooking of errors in the text, and uses psychophysical methods to measure the propensity to mindless reading via readers' sensitivity for these errors (where a low sensitivity indicates a high propensity for mindless reading). We report results from a large-scale eye tracking study where 30 subjects performed an average of three hours of continuous reading in the SAST. As a main result, we found that the influences of lexical and linguistic variables on fixation durations were reliably reduced before errors were overlooked, which validates the SAST as a measure of mindless reading.

As a theoretical contribution, we introduce the levels of inattention hypothesis – proposing that decoupling is a graded phenomenon that occurs at different levels of processing – and test it against the dichotomy hypothesis. To assess decoupling at different levels of processing empirically, we used the SAST and included errors at different levels of the text. Signal detection analyses of error detection and analyses of eye movements prior to overlooking of errors support the levels of inattention hypothesis and suggest that attentional decoupling is a graded phenomenon, where different levels of weak and deep decoupling can be distinguished. Moreover, the results support the prediction from the cascade model of inattention that low-level decoupling is associated with decoupling of higher processing levels (Smallwood, 2011; Smallwood, et al., 2007).

It is a major current challenge and chance for research on mind wandering to develop objective measures (e.g., behavioral, neurophysiological) to detect episodes of mind wandering and decoupling online without interfering with task performance (Franklin, et al., 2011; Smallwood, in press). Previous findings (Reichle, et al., 2010; Smilek, et al., 2010b; Uzzaman & Joordens, 2011) suggest that recordings of eye movements may be one candidate for such an objective measure. However, eye movements show a large variance and it is unclear at present whether enough information is contained in eye movements to predict states of mindless reading reliably at the level of individual eye movements or trials. In chapter 4, we demonstrate in a Bayesian analysis that it is possible to use gaze durations on target words to predict overlooking of lexical errors average five seconds before they occur in the text. Using eye tracking to detect episodes of mindless reading online may trigger the development of new techniques to study mind wandering.

### Chapter 5: Mental effort during mindless reading? Pupil fluctuations indicate internal processing during levels of inattention

In chapter 5, I report additional analyses of the SAST using pupil size as a physiological marker of internal processing load and mental effort (Beatty, 1982b; Hyönä, Tommola, & Alaja, 1995). The aim of these analyses was to test whether mental effort is reduced in the SAST when errors are overlooked during different levels of decoupling. As one possibility, readers may overlook errors during mindless reading because they maintain task-unrelated thoughts (TUT), which are unrelated to the reading task (Schooler, et al., 2011). The decoupling hypothesis proposes that off-line thought (like TUT) is a distinct mode of cognitive function that draws on cognitive resources and engages domain general processes (Smallwood, 2010b, in press; Smallwood & Schooler, 2006), and predicts that while attentional decoupling may be graded, off-line thought is an all-or-none phenomenon (Smallwood, et al., 2011).

The findings reported in chapter five support the prediction by the decoupling hypothesis: weak and medium levels of decoupling in the SAST were associated with small pupil size, which indicates reduced cognitive activity. However, overlooking low-level errors did not affect pupil size. Thus, it seems that during deep attentional lapses internal processing is not reduced. Instead, TUT may demand cognitive resources and may generate some mental effort during deep decoupling. These findings provide support for the decoupling hypothesis that internal thought like TUT is a distinct mode of cognitive functioning, which is associated with deep levels of perceptual decoupling (Smallwood, et al., 2011).

# 2 Eye movements during reading of randomly shuffled text

Daniel J. Schad[1*], Antje Nuthmann[2], and Ralf Engbert[1]

[1] *Department of Psychology, University of Potsdam, Potsdam, Germany*

[2] *Department of Psychology, University of Edinburgh, Edinburgh, UK*

[*] Corresponding author. Email: Daniel.Schad@uni-potsdam.de

*Running head:* Reading of randomly shuffled texts.

Published

Schad, D. J., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research, 50*(23), 2600-2616.

doi: 10.1016/J.Visres.2010.08.005

## Abstract

In research on eye-movement control during reading, the importance of cognitive processes related to language comprehension relative to visuomotor aspects of saccade generation is the topic of an ongoing debate. Here we investigate various eye-movement measures during reading of randomly shuffled meaningless text as compared to normal meaningful text. To ensure processing of the material, readers were occasionally probed for words occurring in normal or shuffled text. For reading of shuffled text we observed longer fixation times, less word skippings, and more refixations than in normal reading. Shuffled-text reading further differed from normal reading in that low-frequency words were not overall fixated longer than high-frequency words. However, the frequency effect was present on long words, but was reversed for short words. Also, consistent with our prior research we found distinct experimental effects of spatially distributed processing over several words at a time, indicating how lexical word processing affected eye movements. Based on analyses of statistical linear mixed-effect models we argue that the results are compatible with the hypothesis that the perceptual span is more strongly modulated by foveal load in the shuffled reading task than in normal reading. Results are discussed in the context of computational models of reading.

*Key words:* Reading; Eye movements; Linear mixed-effect model; Perceptual span; Foveal load hypothesis; Parafoveal-on-foveal effects

## 2.1 Introduction

Reading represents a very complex task because some of the key cognitive systems (e.g., vision, attention, word recognition, memory, oculomotor control, higher-level language comprehension) must interact to move the eyes across the text. Measurement of eye movements represents a powerful approach to investigate the cognitive subsystems involved in reading as eye movements provide a sensitive online-measure for these processes (Rayner, 1998, 2009). One of the most important problems in current research on the control of eye movements concerns the relative importance of low-level visuomotor processes vs. higher-level cognition related to language processing (Starr & Rayner, 2001). This research problem extends to other aspects of active vision, where eye movements are needed for visual information uptake (Liversedge & Findlay, 2000).

Computational models of reading implement theories about how different cognitive processes act in concert to control the movements of the eyes (for an overview of current models, see the 2006 special issue of *Cognitive Systems Research*). It is undisputed that low-level processes like visual perception and oculomotor control affect eye movements during reading. *Primary oculomotor control* models (POC) focus on such low-level processes and ignore direct cognitive influences on eye movements (e.g., Reilly & O'Regan, 1998). *Cognitive models*, to the contrary, assume that higher-level cognition related to language processing plays an important part in controlling the eyes (e.g., E-Z Reader: Reichle, et al., 1998; Reichle, et al., 2006; Reichle, Warren, et al., 2009) (SWIFT: Engbert, et al., 2002; Engbert, et al., 2005).

Up to now, computational models have mainly considered two kinds of cognitive influences on eye movements. The first one is the lexical processing of words, i.e., the type of processing that is needed to get access to a word’s entry in the mental lexicon (e.g., Engbert, et al., 2002; Morrison, 1984; Reichle, et al., 1998; Reilly & Radach, 2006). The second cognitive influence concerns the predictions that readers make about upcoming words in a text (e.g., Engbert, et al., 2002; Reichle, et al., 1998). Recently, a first attempt has been made to also include some effects of higher-level language processing in a computational model of reading (Reichle, Warren, et al., 2009).

Two general strategies have been used to test hypotheses about how higher- and lower-level factors influence eye movements. First, processes can be tied to the influence of certain variables that modulate these effects. For example, word length is regarded as a low-level variable affecting visual processing. Typically, readers look longer at long words than at short words (e.g., Just & Carpenter, 1980; Kliegl, et al., 2006; Rayner, et al., 1996). Effects of word frequency and word predictability, to the contrary, are thought to result from higher-level cognitive influences on eye movements. Low-frequency words are fixated longer than high-frequency words (e.g., Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kliegl, et al., 2006; Rayner & Duffy, 1986). This is mainly because word frequency affects lexical processing, i.e., it takes longer to recognize words that do not occur very often in a given language. Words that are highly predictable from the context receive shorter fixations and more word skippings (see e.g., Balota, et al., 1985; Calvo & Meseguer, 2002; Kliegl, et al., 2006; Rayner, 1998, 2009; Rayner, Ashby, Pollatsek, & Reichle, 2004). Many cognitive processes contribute to this effect, ranging from rather low-level priming effects to high-level language comprehension (see Rayner, 1998, for a review).

A second strategy to test assumptions on the interplay of different cognitive processes in reading has been to develop tasks, which involve similar visual and oculomotor processes as reading but differ with respect to the higher-level cognitive processing that is necessary to complete the task. In the zzz-string scanning task, originally introduced as *mindless reading* (Vitu, et al., 1995), participants read sentences in both their normal version as well as a transformed (or mindless) version where each letter is replaced with a *z* (see also Nuthmann, et al., 2007; Rayner & Fischer, 1996). *z*-String scanning has similar visuo-oculomotor requirements as reading but shares none of the language-related processes. Mindless reading thus approximates reading without lexical and post-lexical processing (see Nuthmann & Engbert, 2009, for a simulation study).

In target-word search (Rayner & Fischer, 1996; Rayner & Raney, 1996), participants search through passages of text for a target word. All linguistic information, like word frequency and predictability of words, is present in the text. However, processing this information is not necessary to complete the task. Instead, the target can be detected based on superficial visual or orthographic analysis of words. Rayner and colleagues have investigated eye movements during target-word search and found no effect of word frequency on eye movements, contrary to robust frequency effects when reading the same text for comprehension. This finding suggests that lexical processing influences eyes movements during reading, but not in visual search for a target word.

Here, we combine these two approaches to add to our knowledge on eye-movement control in reading. We present a new paradigm, the reading of shuffled text, and we compare the influence of various variables on eye movements in this task to reading normal text. The basic idea underlying the shuffling of words is to convert meaningful sentences into meaningless word lists. We used the Potsdam Sentence Corpus (PSC), which consists of 144 single sentences (Kliegl, et al., 2004; Kliegl, et al., 2006). Based on this sentence corpus, the order of words was randomly shuffled across the whole corpus, yielding randomly shuffled word lists, e.g.,

Affen Vorschlag Armen schmale Giebel Kanzler dem besser.

*Monkeys suggestion poor narrow*$_{[FEM/PL]}$ *gable*$_{[SG/PL]}$ *chancellor*$_{[SG/PL]}$ *the*$_{[DAT]}$ *better.*[1]

Jede ihrer Förster im Jahr Hunde meisten Gräfin Bauern.

*Each*$_{[FEM]}$ *her/their foresters [in the] year dogs most*$_{[PL]}$ *countess countrymen.*

In the randomization process, words were not shuffled within sentences, but for each word list words were randomly drawn from all original sentences in the PSC (cf., Morton, 1964, for a different approach to manipulate the context in English text). Readers were instructed to read these random lists of words. To ensure that readers would indeed process the shuffled and normal sentences, some trials were followed by a comprehension question or a word recognition probe. For shuffled word lists, participants were presented with a word triple and asked to indicate which word they recognized as part of the previous

---

[1] Note that languages differ from each other in various aspects. For example, nouns in German are always capitalized.

list; only content words were queried. For normal sentences, readers had to answer an easy three-alternative multiple-choice question pertaining to the content of the sentence.

How are eye movements controlled during reading of shuffled text? In the remainder of Section 2.1, we will derive specific predictions about how readers' eye movements might be affected by random shuffling of words. We will discuss: (1) basic visuomotor processes, (2) whether effects of lexical processing should occur, (3) differences in the predictability of words, (4) memory and post-lexical processes. Lastly (5) we will derive predictions about how theoretical models of reading can explain differences in word-frequency effects between normal and shuffled-text reading.

When reading shuffled text, low-level visuomotor requirements are similar to the ones in normal text reading. Therefore, similar visuomotor effects should be expected in eye movements. Linguistic information on single words, like their frequency, is also available in shuffled texts. Whether and to what degree this information will be relevant for eye guidance is unclear a priori and may depend on the strategy participants adopt to solve the task. In principle, superficial orthographic or phonological analysis can suffice to remember the words.[2] The use of such a strategy would predict that lexical processing does not influence eye movements in shuffled texts, similar to eye movements during target-word search (Rayner & Fischer, 1996; Rayner & Raney, 1996).

However, we expected that readers process words lexically when reading shuffled text and that this should affect their eye movements, in a similar manner as in normal reading. This is plausible (a) because lexical processing is highly automatic (see the Stroop effect, MacLeod, 1991) and (b) because readers were instructed to read the words (and not, for example, to scan them). In addition, (c) encoding the lexical identity of words should aid readers to do well in the word recognition queries and (d) readers may want to use post-lexical processing of, for example, semantic word information to memorize words. In sum, we expected that word frequency should affect eye-movement parameters during reading of shuffled text.

---

[2] Thanks to Keith Rayner for pointing this out.

Further, we expected specific differences between the reading conditions. In randomly shuffled texts, upcoming words cannot be predicted based on their preceding context. Lacking word predictability should lead to a reduced word-skipping rate and increased fixation durations in reading of randomly shuffled texts compared to normal reading. This effect should be quite strong, because in normal text unpredictable words are often neighbored by predictable words, whereas in shuffled text none of the words are predictable. Although shuffled word lists are essentially free of meaning, readers may try to actively construct some meaning to better remember the words in the list (cf., R. A. Mason & Just, 2004; Myers, Shinjo, & Duffy, 1987). Also, we cannot exclude the possibility that some of the random word sequences may partially make sense and trigger automatic semantic or syntactic analyses. In the present study, however, we will focus on effects of lexical word processing, which is often assumed to be the primary cognitive process controlling eye movements during reading (e.g., Engbert, et al., 2002; Engbert, et al., 2005; Reichle, et al., 1998; Reichle, et al., 2003).

In any case, the shuffling of words does not only manipulate overall sentence meaning and the predictability of individual words, but is likely to affect other factors like the ease of retention of words. Shuffled text has no real meaning, which should make it more difficult to remember the words and may invoke different memory-related processes than normal reading. These could contribute to a slower reading pace when reading shuffled text. Further, the specific instruction given to the participants, combined with the occasional word recognition probes, may cause differences in how readers construe their task when reading shuffled as opposed to normal text. Most importantly, only (low-frequency) content words are probed in the recognition test. It is possible that readers are aware of this and focus more strongly on the processing of salient low-frequency content words when reading shuffled text. In contrast, when reading normal text (high-frequency) function words and content words are equally important to construct meaning. We will outline more specific predictions that build upon this basic idea below.

To summarize, lexical processing should principally affect eye movements in both tasks. Therefore, we hypothesize that some basic mechanisms controlling the eyes when reading single unrelated words for recognition are not fundamentally different from the ones acting during normal text reading. However, post-lexical (especially memory-related) processes should differ between reading conditions. Task differences might lead to specific differences in how certain variables, most notably word frequency, modulate fixation times in shuffled text as opposed to normal reading. Such differences will be discussed on the basis of existing models of eye-movement control, with a focus on architectural principles embedded in our own SWIFT model (Engbert, et al., 2005).

Cognitive models of eye guidance in reading make different assumptions about the nature of lexical processing and how attention is allocated to support such processing. According to sequential attention shift (SAS) models, most importantly the E-Z Reader model, attention is allocated serially to support lexical processing of only one word at a time (e.g., Reichle, Liversedge, et al., 2009; Reichle, et al., 1998; Reichle, et al., 2003). Another group of models assumes guidance by a processing gradient (PG). In PG models, attention is distributed continuously as a gradient, which supports the processing of two or more words in parallel (e.g., Engbert, et al., 2002; Engbert, et al., 2005; Reilly & Radach, 2006). Empirical support has been provided for both kinds of models, and aspects of the empirical findings and their theoretical implications are the subject of considerable debate (see e.g., Engbert & Kliegl, 2011; Inhoff, et al., 2005; Kliegl, et al., 2006; Kliegl, et al., 2007; Pollatsek, Reichle, & Rayner, 2006a; Rayner, Pollatsek, et al., 2007; Rayner, White, Kambe, Miller, & Liversedge, 2003; Reichle, Liversedge, et al., 2009).

As stated above, it could be that readers of shuffled text focus more strongly on the processing of low-frequency content words to better remember these words when reading shuffled text. Thus, shuffled text might influence allocation of attention during reading: It could change how the attentional gradient is dynamically modulated in response to foveal load (Henderson & Ferreira, 1990). In the following, we will outline this hypothesis in more detail. The *perceptual span* can be defined as the „region of the visual field from which useful information can be acquired during a given eye fixation“(Henderson & Ferreira, 1990, p. 417). It was studied in the *moving window paradigm* (McConkie & Rayner, 1975), where text is covered with a mask (e.g., XXX) and only the fixated words or letters are visible to the reader. The window of visible text moves with the readers' eyes, and covering parts of the text slows reading down. At a certain window size (about 14–15 letters to the right and 3–4 letters to the left), however, reading with a window proceeds at the same speed compared to when all text is visible, indicating the size of the perceptual span. In the SWIFT model (Engbert, et al., 2005), the concept of a processing or attentional gradient combines the concept of a perceptual span with the notion of parallel processing of words in a sentence. The rationale here is that words within the perceptual span are processed in parallel, at rates decreasing with distance from the current fixation location.

*Does shuffling of words change the dynamical modulation of the perceptual span by foveal load?* The foveal load hypothesis (Henderson & Ferreira, 1990) postulates that the width of the perceptual span is modulated by foveal load (i.e., foveal processing difficulties). If foveal load is low the perceptual span is wide and attentional resources can be distributed across neighboring words. When foveal load increases, the perceptual span gets narrower and the resources left for processing parafoveal information decrease. Empirically, an incorrect preview for word $n + 1$ during fixations on word $n$ interferes with reading word $n + 1$ more strongly if word $n$ is of high-frequency, due to increased parafoveal processing in this condition (Henderson & Ferreira, 1990) (see also Balota, et al., 1985; Inhoff, Pollatsek, Posner, & Rayner, 1989; Inhoff & Rayner, 1986; Rayner & Pollatsek, 1987; White, Rayner, & Liversedge, 2005). In corpus analyses, the same mechanism is visible. Here, high-frequency words $n - 1$ increase preview for word $n$ during fixations on word $n - 1$. Because part of the processing of word $n$ could already be finished while still fixating word $n - 1$, the fixation on word $n$ is then shorter and the effect of frequency of word $n$ on fixation durations is weaker (Kliegl, et al., 2006).

As outlined above, concerning its theoretical interpretation the foveal load hypothesis naturally adheres to the parallel processing assumption in reading. In PG models, low foveal load would lead to a widening of the attentional gradient. High foveal load, to the contrary, would narrow the attentional gradient such that only the fixated word would be processed. The basic foveal load finding (reduced preview benefit in case of increased foveal load) can also be accounted for within the SAS framework. The E-Z Reader model explains the effect by assuming that the second stage of lexical processing (L2) is a function of word frequency (Pollatsek, Reichle, & Rayner, 2006c; Reichle, et al., 1998; Reichle, et al., 2006; Reichle, et al., 2003; Reingold & Rayner, 2006). In the model, L2 takes longer to complete for low-frequency words, which leads to less preview of the next word (and can even produce spill-over effects). Thus, the key signature finding of the foveal load hypothesis is compatible with both parallel and serial accounts of attention allocation during reading.

We now derive further, more specific predictions based on the assumption that foveal load modulates the perceptual span (Henderson & Ferreira, 1990) based on the PG framework. The basic assumption is that foveal load modulates the width of the attentional gradient. In addition, we assume that the processing resources are limited (i.e., that the total processing rate is constant at any time), such that capture of attentional resources by the fixated word would result in reduced processing of the neighboring word *n* + 1 (see Figure 2-1 for an illustration).

As a first prediction for shuffled text, the effect of current-word frequency should be reduced if the modulation of the perceptual span by foveal load is strong, because low-frequency words capture more attentional resources compared to high-frequency words due to the contraction of the perceptual span. Second, a parallel processing account predicts that the influence of the upcoming word *n* + 1 on fixation durations depends on the frequency of the currently fixated word n. Because the amount of preprocessing of the next word depends on the width of the perceptual span (which in turn depends on the frequency of the fixated word), we expect parafovea-on-fovea effects to be modulated by foveal load (cf., Kliegl, et al., 2006). Third, the current-word frequency effect should depend on the length of the currently fixated word n. A long word, be it of high or low-frequency, will fill more or less the whole perceptual span (Figure 2-1a and b). Therefore, the current-word frequency effect should be fully visible. The effect might be weaker for short words, as they can benefit strongly from focusing of the perceptual span (Figure 2-1c and d). Fourth, to the degree that short words *n* benefit from focusing of the perceptual span, processing of successor words *n* + 1 should suffer from it. A short word *n* with a low frequency should attract all attentional resources. Accordingly, parafoveal processing of word *n* + 1 should be strongly reduced and fixation durations on word *n* + 1 should be enhanced (compared to good preview during a short high-frequency word n). A long word n, again, will fill the whole perceptual span independent of its frequency. For that reason, preprocessing of and fixation durations on word *n* + 1 should not strongly depend on the frequency of word n.

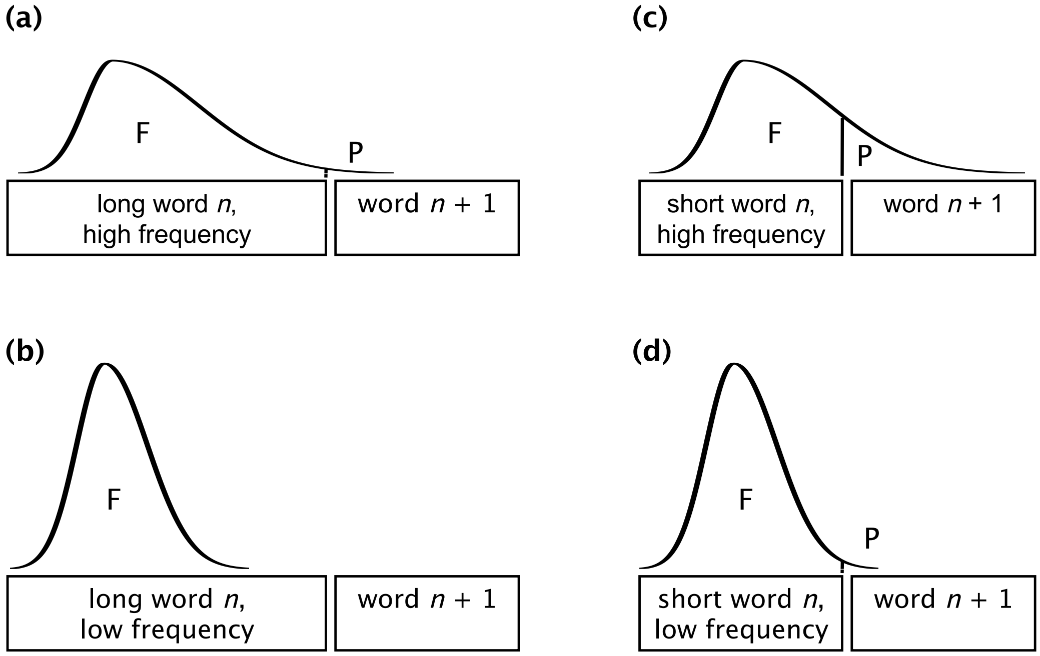


**Figure 2-1**

Processing rate over foveal eccentricity; peak indicates fixation location. Predictions of the foveal load hypothesis for long words (left plots) vs. short words (right plots) with high (top row) vs. low (bottom row) frequency. Low word frequency equates to high foveal load. (1) Long word n: narrowing the perceptual span in response to a low-frequency word does not increase the processing resources available for the fixated word *n* (F) much (compare (b) with (a)). (2) Short word n: narrowing the perceptual span in response to a low-frequency word strongly increases the processing resources available for the fixated word *n* (compare (d) with (c)). F = processing resources available for the foveal word n; *p* = processing resources available for the parafoveal word *n* + 1.

Deriving these four specific predictions is rather straightforward from the perspective of PG models supporting parallel word processing in reading. Notably, the predictions are derived based on one single mechanism, that is, the modulation of the perceptual span by foveal load.

## 2.2 Methods

### 2.2.1 Participants

Sixty university students participated in the study. Thirty readers took part in the shuffled reading condition. Their eye-movement data were compared with data generated by participants who read the Potsdam Sentence Corpus (PSC, normal sentence reading, *n* = 30), an age-matched subsample from a large set of data that has previously been reported in Kliegl et al. (2006). Both groups were tested in the same lab, using the same technical equipment. The two groups did not differ in age (shuffled-text reading: *M* = 22.8, *SD* = 3.4; normal reading: *M* = 22.6, *SD* = 3.6) and in psychometric tests of vocabulary (shuffled-text reading: *M* = 31.8, *SD* = 2.7; normal reading: *M* = 32.7, *SD* = 1.6), and digit-symbol substitution (shuffled-text reading: *M* = 61.7, *SD* = 9.6; normal reading: *M* = 59.2, *SD* = 9.4).

### 2.2.2 The Potsdam Sentence Corpus (PSC) and shuffled texts

The PSC comprises 144 German single sentences. They range from 5 to 11 words (*M* = 7.9, *SD* = 1.4), and there are 1138 words in total. Norms on psycholinguistic variables such as word length, printed word frequency (Geyken, 2006), and predictability norms from an independent cloze-task study are available for each word in the PSC. For details of materials and experimental procedure for the normal PSC data we refer to Kliegl et al. (2004; 2006). To create shuffled text, each single sentence in the PSC was replaced by a shuffled word list. For each sentence, each word was replaced by a different word that was randomly drawn without replacement from the pool of all words that occur in the PSC. In this randomization procedure, the first word of an original PSC sentence was always the first word in a shuffled sentence; the same was true for the last words in sentences. All other words were drawn from random locations in a sentence. Using this constrained randomization procedure a separate set of 144 word lists was generated for each participant. As a consequence of this procedure, words in one word list were randomly drawn from many different sentences in the PSC.

### 2.2.3 Apparatus, materials and procedure

One group of participants read the original 144 PSC sentences, while the other group read a set of 144 random word lists. Sentences and word lists were presented in random order at a distance of 60cm on the centerline of a 21-in. EYE-Q 650 Monitor (832 × 632 resolution; frame rate 75 Hz; font: regular New Courier 12; visual angle: 0.38° per character). A chinrest was used to minimize participants' head movements. Both eyes were monitored with an EyeLink II system (SR Research, Osgoode, ON, Canada) with a sampling rate of 500 Hz and an instrumental spatial resolution of 0.01°. Minimal head movements were corrected automatically by the EyeLink II system.

In order to motivate participants to read the word lists and/or sentences, simple questions occurred after 27% of the sentences and after one third of the word lists. In sentence reading, participants were asked questions pertaining to the meaning of the sentence. As response alternatives, a word triple was presented with the question and participants were required to indicate the correct word, which was always part of the sentence. In shuffled-text reading, participants were again presented with a word triple and were asked to decide which of the three words had been part of the list seen before. In both conditions, only nouns, verbs, or adjectives were queried in order to avoid changing the experiment into a (difficult) memory task. (Preliminary tests had shown that asking for prepositions, adverbs, etc. was difficult.) Participants were not informed about this particularity.

### 2.2.4 Data selection

An initial screening excluded the records of sentences with blinks or loss of measurement from the data. Data from a maximum of 27 (*Median* = 3) sentences were excluded per participant. A binocular velocity-based algorithm for saccade detection (Engbert & Kliegl, 2003; Engbert & Mergenthaler, 2006) was used to identify saccades and fixations. To adjust for the reading situation, only fixations with a minimal duration of 10 ms and saccades with a minimal amplitude of 0.75° were detected. Fixations were assigned to letters within words. Sentences with less than three fixations and fixations left or right of the sentence borders were removed. This procedure resulted in a total number of 73,858 fixations (see Table 2-1 for separate numbers for the shuffled vs. normal text reading groups).

We excluded fixations according to the following criteria: (1) the first or last fixation in a sentence as well as fixations on the first or last word ($N$ = 20,944), (2) fixations longer than 750 ms and fixations bordered by a saccade amplitude of 25 letters or longer (*N* = 313). The remaining fixations are valid fixations (*N* = 52,601). Among these we identified fixations that were not in first-pass reading[3] (*N* = 7051). Given that we wanted to examine influences from neighboring words, we only considered fixations where the left and right eye fixated on the same word. We thus excluded cases where the left and right eye fixated on different words (N = 5715; see Nuthmann & Kliegl, 2009, for an investigation of disparity between eyes). All measures of fixation durations or fixation probabilities were determined using the right eye. We further distinguished cases in which a word was fixated exactly once (binocularly reliable single fixation cases; *n* = 24,433) from cases in which a word had been fixated more than once during first-pass reading (multiple fixation cases; *n* = 15,402). In sum, the single fixation durations analyses reported below consider first-pass fixations where both eyes fixated on the same word.

**Table 2-1**

Number of fixations for various types of fixations in shuffled and normal text reading.

| | | shuffled text | normal text | total |
|---|---|---|---|---|
| 1 *N of fixations* | | 41,873 | 31,985 | 73,858 |
| 2 first/last word; first/last fixation | *N* | 11,075 | 9,869 | 20,944 |
| | *%* | 26 | 31 | 28 |
| 3 long fixation or amplitude | *N* | 195 | 118 | 313 |
| | *%* | 0.5 | 0.4 | 0.4 |
| 4 *N of valid fixations* | | 30,603 | 21,998 | 52,601 |
| 5 not in first pass | *N* | 4,575 | 2,476 | 7,051 |
| | *%* | 15 | 11 | 13 |
| 6 different words | *N* | 2,784 | 2,931 | 5,715 |
| | *%* | 9 | 13 | 11 |
| 7 multiple fixations | *N* | 10,272 | 5,130 | 15,402 |
| | *%* | 34 | 23 | 29 |
| 8 single fixations | *N* | 12,972 | 11,461 | 24,433 |
| | *%* | 42 | 52 | 46 |

*Note*. Row 1 = 2+3+4; row 4 = 5+6+7+8. Data are from 30 readers in the shuffled, and 30 readers in the normal text condition. Data are from right eye.

---

[3] First-pass reading comprises all fixations on a word that occur before the first regression has originated from this word or a word following later in the sentence.

In valid sentences, readers made first-pass fixations on a total of 55,323 words. When reading shuffled text, more words were fixated in first pass ($N$ = 29,704) than when reading normal text ($N$ = 25,619). For fixated words, all first-pass fixations were summed up to obtain gaze durations. For a given subject, words on which at least one invalid fixation (first/last word; first/last fixation; long fixation or saccade amplitude) was identified (shuffled: $n$ = 8372; normal: $n$ = 8156), as well as gaze durations that were longer than 1000 ms (shuffled: $n$ = 46; normal: $n$ = 11) were excluded from analysis. This procedure resulted in a total of 38,738 gaze durations (shuffled text: $n$ = 21,286; normal text: $n$ = 17,452).

## 2.3 Results

### 2.3.1 Global summary statistics

Reading shuffled text resulted in a higher overall number of fixations than reading normal text. This also translated into a higher number of valid as well as first-pass fixations (see Table 2-1), and also more valid gaze durations. Accordingly, readers of shuffled text made more fixations per trial than normal text readers [10.1 vs. 7.8; $t(51) = 5.14$, $p < 0.001$; see Appendix 2-A, Table A-1, for descriptive statistics of eye movements]. Amplitudes for forward saccades were on average shorter in shuffled-text reading as compared to normal reading [6.1 vs. 7.6 letters; $t(55) = -5.85$, $p < 0.001$; see Figure A-1b) for the corresponding distributions of saccade lengths]. Shorter saccade lengths in shuffled text compared to sentence reading were associated with a strong reduction of skipping rate [0.10 vs. 0.21; $t(55) = -6.45$, $p < 0.001$] and an increase in refixation probability [0.16 vs. 0.08; $t(48) = 5.41$, $p < 0.001$]. Refixations were not only more frequent in the shuffled text condition but they were also more often rightward-oriented [90% vs. 79% of refixations in first-pass reading; $t(58) = 3.2$; $p < .01$]. The decrease in skipping probability and increase in refixation probability canceled each other out such that the probability of single fixation was similar for the two groups [0.70 vs. 0.67; $t(58) = 1.69$; $p = 0.10$].

The percentage of regressions was exactly the same (0.06 vs. 0.06). Likewise, the distribution of backward-oriented saccade amplitudes did not differ between reading conditions [Figure A-1b]. The number of fixations in second- and more-pass reading was largely enhanced in shuffled-text reading (4575 vs. 2476 fixations; $\chi^2(1) = 625$; $p < .001$; see Table 2-1).

In shuffled-text reading, readers initially fixated further to the left in a word compared to normal text reading. This difference was significant for single fixation cases [initial fixation on letter 2.5 vs. 2.7; $t(53) = -3.22$, $p < 0.01$], while there was a trend for the first of multiple fixations [letter 2.0 vs. 2.2; $t(53)= -1.85$, $p = 0.07$].

Fixation durations were generally longer in readers of shuffled compared to normal text. This effect showed as a moderate shift in mean and skew in the corresponding global fixation duration distribution [Figure A-1a]. The difference in fixation durations was observed across all types of fixations; it was significant for single [254 vs. 213 ms; $t(55) = 4.37$, $p < 0.001$], first [227 vs. 199 ms; $t(58) = 3.50$, $p < 0.001$], and second [197 vs. 172 ms; $t(58) = 2.74$, $p < 0.01$] fixations, as well as for gaze durations [293 vs. 231 ms; $t(50) = 4.89$, $p < 0.001$]. As a result of the higher number of fixations and the longer fixation durations, the reading rate was strongly reduced in readers of shuffled text as compared to normal text.

Memory performance was close to perfect for readers of normal text (97.5% of the questions, *SD* = 3.6, were answered correctly). Readers of the shuffled text answered 85% of the questions correctly (*SD* = 3.1).

### 2.3.2 Linear mixed-effects models

We used gaze duration and single fixation duration as dependent measures in our analyses. Gaze durations and single fixation durations were log-transformed to avoid problems with heteroscedasticity. To determine the impact of various predictors on log-fixation durations in shuffled text vs. sentence reading, a linear mixed-effects model (LME; e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000) (see also Kliegl, 2007; Kliegl, Masson, & Richter, 2010; Kliegl, et al., 2007) was tested, using the lmer program of the lme4 package (Bates & Sakar, 2008). Plots were created using the ggplot2 package (Wickham, 2009). The packages and programs are supplied in the R system for statistical computing (R Development Core Team, 2012; under the GNU General Public License, Version 2, June 1991).

Fixed effects in LME terminology correspond to regression coefficients in standard linear regression models. They can also estimate slopes or differences between conditions. A number of fixed effects were entered into the model. We tested the influence of visual and lexical factors characterizing the currently fixated word *n* by including its length (i.e., 1/length) and its frequency, with linear and quadratic (cf., Kliegl, 2007) effects, as well as their multiplicative interaction (cf., Pollatsek, et al., 2008; Schroyens, et al., 1999). To test for lag effects of the previous word *n* – 1 on fixation durations on the fixated word n, we used word *n* – 1 length (1/length; cf., Pollatsek, et al., 2008) and frequency as predictors (cf., Rayner & Duffy, 1986; Schroyens, et al., 1999). Likewise, successor effects were tested by including word *n* + 1 length (1/ length) and frequency (cf., e.g., Kennedy & Pynte, 2005; Vitu, Brysbaert, & Lancelin, 2004). We further added the length of the incoming (cf., Pollatsek, et al., 1986; Radach & Heller, 2000; Vitu, McConkie, Kerr, & O'Regan, 2001) and outgoing saccades as model predictors. To capture the inverted-optimal viewing position effect for fixation durations (IOVP, Nuthmann, et al., 2005, 2007; Vitu, Lancelin, & d'Unienville, 2007; Vitu, et al., 2001) the relative fixation position within a word (i.e., fixated letter number divided by word length) was included as a linear and as a quadratic effect.

In addition, three further predictors involving multiplicative interaction terms of continuous variables were added to the model. We tested whether the influence of current-word frequency was modulated by the frequency of the prior word (a prediction derived from the foveal load hypothesis, Henderson & Ferreira, 1990; Kliegl, et al., 2006). Likewise, we examined whether the influence of the frequency of the parafoveal word *n* + 1 depended on limits of visual acuity (i.e., on the length of the fixated word; cf., Kennedy & Pynte, 2005) and on attentional constraints (i.e., on the frequency of the fixated word; a second prediction derived from the foveal load hypothesis). Except for the quadratic effect of current-word frequency (Kliegl, 2007) and lacking effects of word predictability, this set of predictors was identical to the set of predictors tested with repeated measures multiple regression analysis (rmMRA) reported by Kliegl et al. (2006) (see also several random-subject lme models in Kliegl, 2007).

For statistical modeling we used two complementary approaches. First, we tested whether the fixation-level fixed effects differed between the shuffled and the normal reading group (i.e., we tested cross-level interactions). This was done by simultaneously including all of the fixation-level effects as well as their interactions. Experimental condition was included as a dummy factor, using the shuffled text condition as the reference group.[4] In addition, we estimated how strongly mean fixation durations varied with participants and words by fitting crossed random intercepts for participants and words (if the same word occurred more than once in the corpus, the same random effect was used for all of these occurrences, yielding unique word ID). Instead of estimating a slope or a difference between conditions, random effects estimate the variance that is associated with the levels of a certain factor. After including these effects into the model, non-significant predictors were dropped. The results for this final model are reported in the text below; for an overview see Appendix 2-B, Table B-1. Values of $t > 1.96$ indicate significance of a predictor, while effects with $t > 1.645$ indicate marginal significance. Second, we tested whether the fixation-level fixed effects described above are significant in each of the reading conditions separately. To do so, we included each of these predictors twice within one model: once nested under shuffled and once nested under normal text reading.[5] In this post hoc model, we again used the same random effects and the same procedure for dropping predictors. In the following we report the effects of word frequency when reading normal and shuffled text.

---

[4] Consequently, if the interaction of a fixation-level fixed effect with experimental condition is kept in the model (e.g., frequency of word n × experimental condition), the coefficient estimating the fixation-level fixed effect itself (i.e., in this case the main effect of frequency of word n) tests the influence of this variable in the shuffled text condition. If the same interaction is, however, removed from the model because it does not reach significance, the fixation-level fixed effect (e.g., the main effect frq. n) represents the average effect of the variable (frq. n) for both reading conditions.

[5] Nesting a covariate (e.g., word frequency) under the level of an experimental factor (e.g., under shuffled-text reading) can be done by means of setting all values of the covariate for the other factor levels (in this case for normal sentence reading) to zero and to center the covariate within the critical factor level. As a result, the effect of the covariate is estimated and tested only within the specified factor level (Kliegl, 2007).

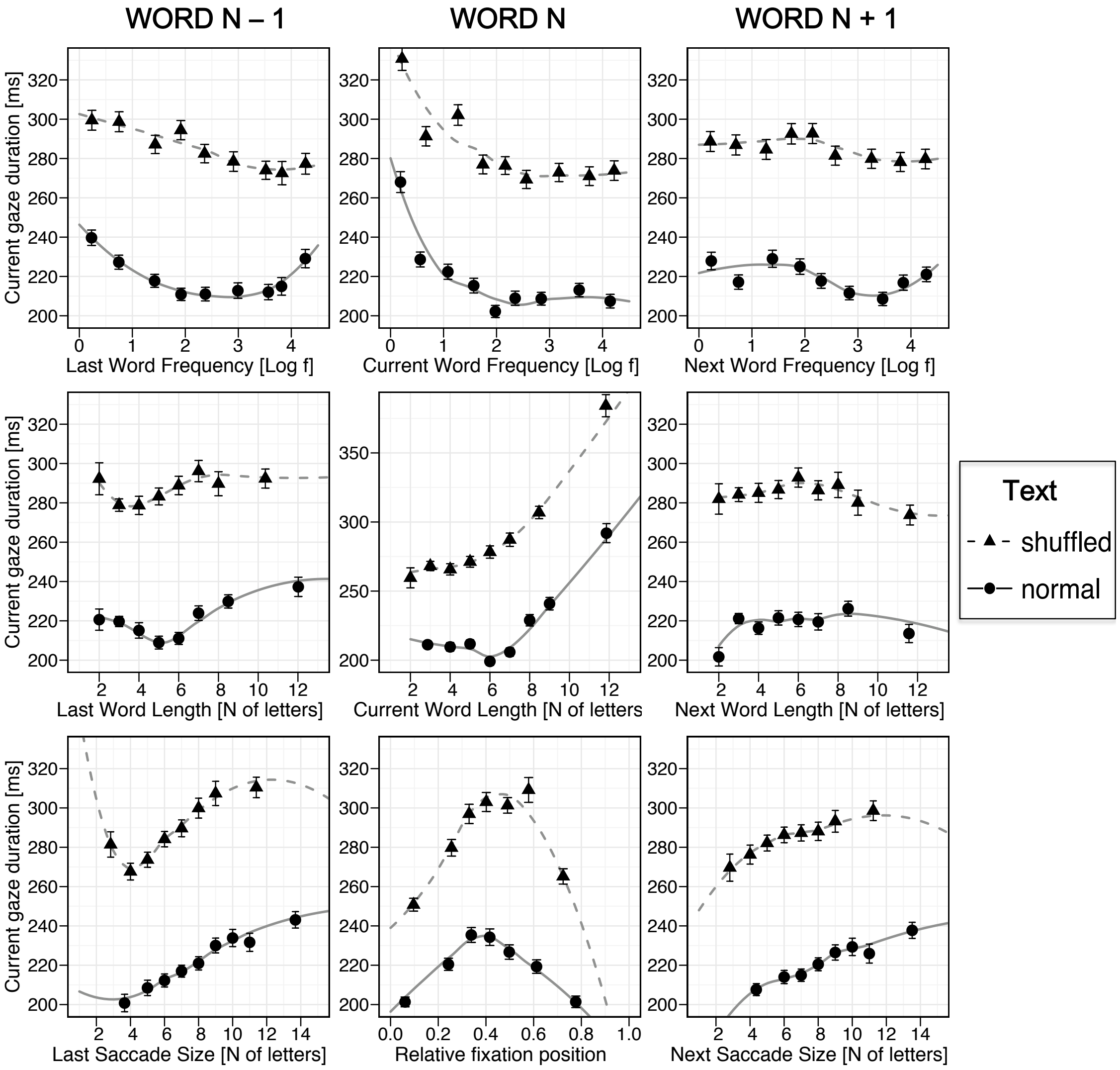


**Figure 2-2**

Nine main effects for gaze durations on word n for reading shuffled (*N* = 30, triangles and dashed lines) and normal text (*N* = 30, circles and solid lines). Predictors are frequency and length of words *n* – 1, *n*, and *n* + 1 (first two rows), the amplitude of the incoming saccade, the relative fixation position (rfp) in the word (linear + quadratic trend), and the amplitude of the outgoing saccade (last row). For each predictor, fixations were binned into categories with a minimum of 800 fixations. Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005). In addition, the predictions from a least squares local regression model, applied to the full set of ungrouped data, are plotted for each effect.

#### 2.3.2.1 Effects of current-word frequency

##### Main effect of word frequency

The word-frequency effect on fixation durations is one of the most basic and best-replicated findings in reading research: low-frequency words are fixated longer than high-frequency words (e.g., Inhoff & Rayner, 1986). Accordingly, fixation durations should decrease with increasing current-word frequency. Such an inverse relationship will be referred to as a negative effect of a variable (indicated by a negative fixed effect coefficient), while we will use the term positive effect (with a positive coefficient in the model) for cases in which fixation durations increase with higher values in the predictor variable.

For the log-frequency of the fixated word n, we found the expected negative influence on gaze durations (see Figure 2-2; for normal sentence reading: $b$ = -0.032, $SE$ = 0.006, $t$ = -5.23). For readers of shuffled text, however, the linear effect of word frequency disappeared (in Figure 2-2, low-frequency words show somewhat longer gaze durations because word frequency is confounded with effects of word length. The LME model controls for such effects and reveals a null-effect of word frequency: ($b$ = 0.004, $SE$ = 0.066, $t$ = .67; for the difference between conditions: $b$ = -0.026, $SE$ = 0.004, $t$ = -6.7). The quadratic current-word frequency effect did not significantly differ between the two conditions and was overall significant ($b$ = 0.020, $SE$ = 0.005, $t$ = 4.3).

For single fixation durations, the linear current-word frequency effect also significantly differed between the two reading conditions ($b$ = -0.028, $SE$ = 0.004, $t$ = -7.0). Like in gaze durations, it was weaker in readers of shuffled text. However, the effect actually changed its sign for single fixation durations. In sentence readers, low-frequency words were fixated significantly longer than high-frequency words ($b$ = -0.028, $SE$ = 0.005, $t$ = -5.8). For readers of shuffled text, this traditional negative frequency effect numerically turned positive, such that low-frequency words were fixated for less time than high-frequency words. This positive frequency effect was marginally significant ($b$ = 0.010, $SE$ = 0.005, $t$ = 1.84). The quadratic frequency effect on single fixation durations did not significantly differ between the two conditions and was overall significant ($b$ = 0.017, $SE$ = 0.004, $t$ = 4.5).

### Interaction of word frequency and word length

The current-word frequency effect on log gaze durations was modulated by word length, as there was a stronger frequency effect for long compared to short words. This was true for normal reading (see also Kliegl, et al., 2006) as well as reading of shuffled text (for the overall interaction of word length and frequency: $b$ = 0.401, $SE$ = 0.055, $t$ = 7.3). This interaction did not significantly differ between the two reading groups. For sentence readers, the current-word frequency effect was negative for both long and short words (Figure 2-3a). For readers of shuffled text, this effect changed its sign. Low-frequency words were actually fixated shorter than high-frequency words, if the words were of short length. These word length dependent linear frequency effects were significant in a post hoc analysis.[6]

While the current-word frequency effect on log single fixation durations was significantly modulated by word length for readers of normal sentences ($b$ = 0.160, $SE$ = 0.041, $t$ = 3.9), this modulation was not significant for participants reading shuffled text ($t$ = -1.38; for the condition-difference: $b$ = 0.177, $SE$ = 0.037, $t$ = 4.8). However, we again tested the same post hoc model as reported for the corresponding interaction in the gaze duration analysis and again found current-word frequency effects to be significantly positive only for short words among readers of shuffled text ($b$ = 0.015, $SE$ = 0.005, $t$ = 2.95). In normal sentence reading, however, the frequency effects were significantly negative in both word length conditions ($b$s < -0.011, $t$s < -2.0).

To summarize, during reading of shuffled text the current-word frequency effect on gaze and single fixation durations was overall strongly reduced. It disappeared for gaze durations and was actually reversed for gaze durations on short words and for single

---

[6] To test these effects, the word length-variable was dichotomized (median-split; short words had five letters or less) and word frequency was nested under long and under short words in the shuffled and in the normal text reading condition (yielding four linear effects of word frequency for these four conditions). The new current-word length and frequency variables were used in an additional post-hoc mixed-effects model that lacked the overall linear effects of word length, frequency, and their interaction, and that was otherwise identical to the first post-hoc model (i.e., testing fixation-level effects nested under experimental condition). The linear frequency effect was significantly negative in three conditions [in short ($b$ = -0.013; $SE$ = 0.007; $t$ = -1.98) and long ($b$ = -0.071; $SE$ = 0.010; $t$ = -7.2) words for normal text reading and in long words ($b$ = -0.040; $SE$ = 0.011; $t$ = -3.8) for shuffled-text reading], but was significantly positive for short words among readers of shuffled text ($b$ = 0.013, $SE$ = 0.006, $t$ = 2.0) [effect of word length (short vs. long words): $b$ = 0.098; $SE$ = 0.011; $t$ = 8.9; (Exp) × (word length): $b$ = -0.046; $SE$ = 0.009; $t$ = -5.3].

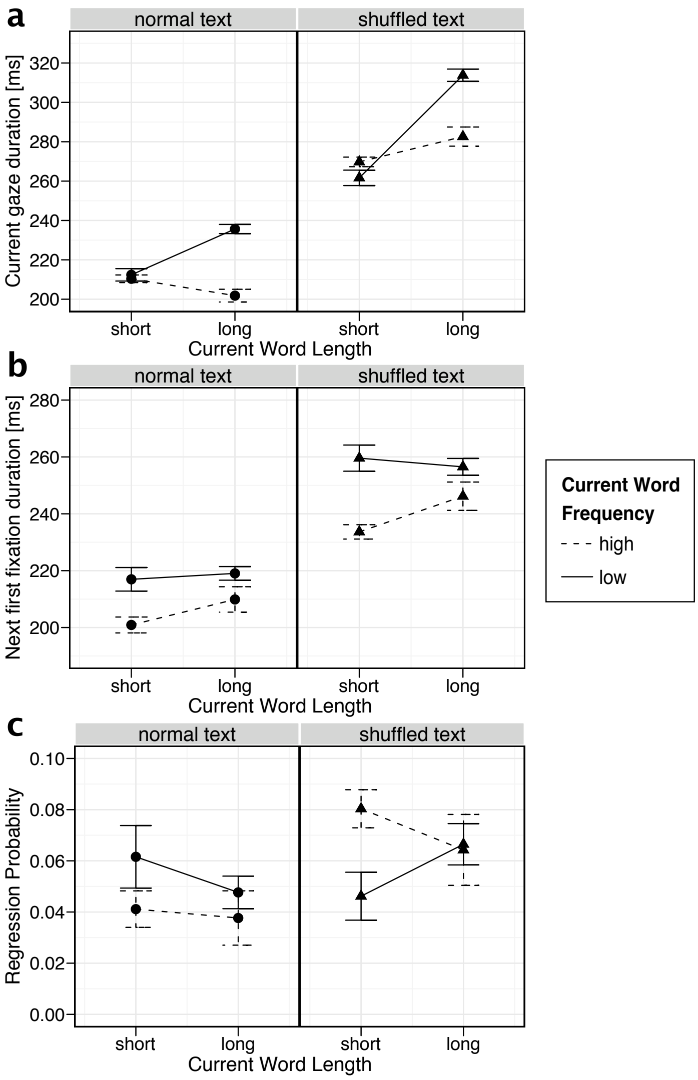


**Figure 2-3**

Interaction between length and frequency of word *n* for normal (left plots, circles) vs. shuffled (right plots, triangles) text reading. (a) Effects on gaze duration on word *n*. (b) Modulation of first fixation duration on word *n* + 1, defined as the duration of the next fixation after having made one first-pass single fixation (see Section 2.2 for selection criteria) on word *n* and given that this next fixation is on word *n* + 1. (c) Effects on regression probability to word *n*, defined as the probability of regressing to word *n* after having made one first-pass single fixation on word *n* and one fixation on word *n* + 1. Short words are five or fewer letters long; DWDS frequencies were split on medians (calculated across both groups). Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005).

fixation durations, yielding longer fixations on high- compared to low-frequency words. However, the standard effect of word frequency, with longer fixations on low-frequency words, was observed on long words in the gaze duration analysis. Also, the quadratic effect of word frequency was present during reading of normal as well as shuffled text.

#### 2.3.2.2 Effects of distributed processing: lag and successor frequency

##### Lag effects

The effect of the frequency of word $n - 1$ on gaze durations did not significantly differ between the two groups of readers ($t = -1.2$). It was significant and negative in both groups (shuffled text: $b = -0.033$, $SE = 0.003$, $t = -13.2$; normal text: $b = -0.040$, $SE = 0.003$, $t = -11.9$). For single fixation durations, the lag effect of word $n - 1$ frequency was numerically weaker in readers of shuffled text. However, the condition-difference for the slope of word $n - 1$ frequency only approached significance ($t = -0.008$, $SE = 0.004$, $t = -1.7$). The effect was still strong and highly reliable in readers of shuffled text ($t = -0.034$, $SE = 0.003$, $t = -12.5$).

##### Successor effects

The effect of the frequency of the upcoming word $n + 1$ on gaze durations did not significantly differ between shuffled and normal text reading, however there was a trend towards a stronger effect in shuffled text readers ($b = 0.006$, $SE = 0.003$, $t = 1.86$). Gaze durations were generally shorter before high-frequent words $n + 1$ (shuffled PSC: $b = -0.015$, $SE = 0.003$, $t = -5.9$; normal PSC: $b = -0.011$, $SE = 0.003$, $t = -3.3$). The same was true for the successor effect on single fixation durations: there was a significant effect for shuffled ($b = -0.015$, $SE = 0.003$, $t = -5.5$) and for normal text readers ($b = -0.010$, $SE = 0.003$, $t = -3.2$), but no significant slope-difference ($b = 0.005$, $SE = 0.003$, $t = 1.6$). Thus, we found strong, consistent, and highly reliable effects of lag and successor-word frequency on gaze and single fixation durations during normal and shuffled-text reading.

#### 2.3.2.3 Interactions of frequencies of neighboring words

##### Lag effects

For gaze durations, the interaction between word $n$ and word $n - 1$ frequency was significant in the normal sentence reading condition ($b = -0.004$, $SE = 0.002$, $t = -2.1$). It was also significant for readers of shuffled text ($b = 0.015$, $SE = 0.001$, $t = 10.4$). However, the coefficient was opposite in sign and higher in absolute value (for the difference: *b* = -0.019, *SE* = 0.002, *t* = -7.9). Among readers of shuffled text, gaze durations were especially prolonged if word *n* and word *n* – 1 were both low in frequency (see Figure 2-4a). For readers of normal text, on the other hand, gaze durations were particularly shortened in the case of high-frequent words *n* and *n* – 1.

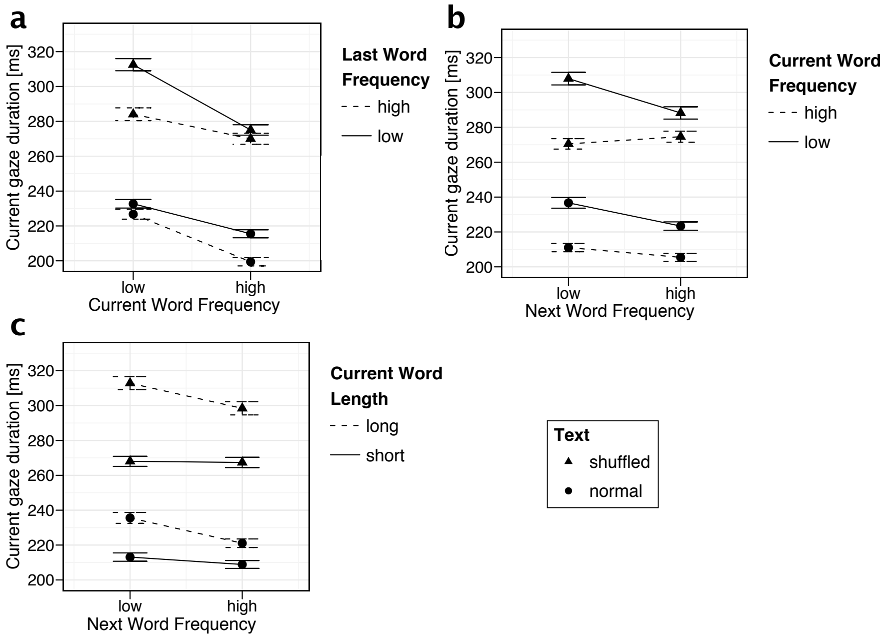


**Figure 2-4**

Modulation of gaze durations on word n due to three interactions for readers of normal (circles) and shuffled (triangles) text: (a) frequency of word *n* frequency of word *n* – 1, (b) frequency of word *n* + 1 frequency of word *n* and (c) frequency of word *n* + 1 length of word *n*. Dependent variable is always gaze duration on word *n*. Short words are five or fewer letters long; DWDS frequency were split on medians. Error bars are within-subject 95% confidence intervals.

For single fixation durations, we also found a strong and highly significant interaction between word *n* and word *n* – 1 frequency for readers of the shuffled PSC (i.e., a foveal load lag effect: *b* = 0.019, *SE* = 0.002, *t* = 12.3). This interaction was significantly stronger (*b* = -0.018, *SE* = 0.002, *t* = -7.2) than the corresponding interaction for normal text.[7] As for gaze durations, the lag-frequency effect was stronger in low- than in high-frequency words *n* (see Figure 2-5a).

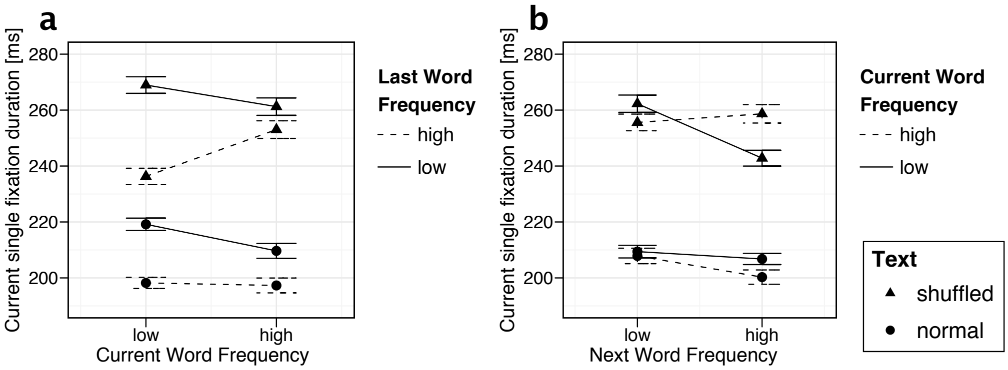


**Figure 2-5**

Modulation of single fixation durations on word n due to two interactions for readers of normal (circles) and shuffled (triangles) text: (a) frequency of word *n* frequency of word *n* – 1 and (b) frequency of word *n* + 1 frequency of word *n*. Dependent variable is always single fixation duration on word *n*. Short words are five or fewer letters long; DWDS frequency were split on medians (calculated across both groups). Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005).

## Successor effects

The interaction between word *n* and word *n* + 1 frequency on gaze durations was not significant for readers of normal sentences, replicating prior research (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Kliegl, et al., 2006; White, et al., 2005). However, we observed a significant interaction in the shuffled-text reading condition (*b* = 0.010, *SE* =

---

[7] The interaction of word n and word *n* – 1 frequency was not significant for the normal PSC reading sample that we used in this study (*t* = 0.49). However, this same interaction has earlier been found to be highly reliable across various samples of participants reading the PSC (Kliegl, 2007; Kliegl, et al., 2006). Therefore, we checked whether the interaction that we found for shuffled text readers was also stronger than the corresponding effect in other samples reading the normal PSC. To do so, we fitted a linear mixed-effects model using the same predictors as the ones reported in Kliegl (2007, Table 1) using non-transformed single fixation durations. We then checked whether the interaction-coefficient for shuffled PSC readers was larger than equivalent coefficients for other samples reading the normal PSC. The comparison with the data reported in Kliegl (2007, Table 1) reveals that the largest coefficient for this interaction in any of the other PSC samples was *b* = 3.0 and was thus more than two standard errors below the coefficient that we found for the shuffled reading group (*b* = 4.5, *SE* = 0.41, *t* = 11.2). Thus, the interaction of word n and word *n* – 1 frequency was stronger in readers of shuffled text compared to many observed samples of participants reading the normal PSC.

0.002, $t$ = 5.4): The frequency effect of word $n$ + 1 on gaze durations was stronger if word $n$ was a low-frequency word (see Figure 2-4b).

Similarly to the gaze duration data, the interaction of frequency of word $n$ and word $n$ + 1 was significant in the shuffled ($b$ = 0.010, $SE$ = 0.002, $t$ = 4.9) but not in the normal text reading condition ($t$ = -0.55; condition-difference: $b$ = -0.010, $SE$ = 0.002, $t$ = -4.0) when analyzing single fixation durations. For the shuffled text readers, the parafovea-on-fovea effect of word $n$ + 1 frequency on single fixation durations was negative (i.e., longer fixation durations next to low-frequent words $n$ + 1) if the foveal word had a low-frequency. Surprisingly this effect numerically turned positive for high-frequent words $n$ (i.e., shorter fixation durations next to low-frequent words $n$ + 1; see Figure 2-5b).

In summary, foveal load effects were much stronger in readers of shuffled text. In particular, the frequency of the last word $n$ – 1 modulated effects of current-word frequency more strongly, and the current-word frequency modulated successor-frequency effects when reading shuffled text.

### 2.3.3 Further tests of relative word-frequency effects

#### *2.3.3.1 Relative Lag-frequency effects – fixation durations*

If the preview of word $n$ + 1 during fixations on word $n$ depends on the interaction of word $n$ length and frequency, then increased preview should show in shorter fixations on the next word $n$ + 1 (i.e., in a reduced spill-over effect). To test this, we refit the primary linear mixed model described above to regress the (log) duration of the first fixation on word $n$ + 1 after having made a single fixation on word $n$ on all the predictors reported above. In addition, we added the lag-frequency times lag-word length interaction to the set of fixed effects (note that these lag effects correspond to the current-word frequency and length effects in the previous models). Cases in which word $n$ + 1 was skipped during first-pass reading were excluded from the analysis, resulting in a total of 16,577 fixations. While there was a highly significant interaction of word $n$ length × frequency on fixation durations on word $n$ + 1 for readers of shuffled text ($b$ = -0.179, $SE$ = 0.033, $t$ = -5.4), this interaction was significantly weaker for readers of normal text ($b$ = 0.119, $SE$ = 0.055, $t$ = 2.2). As can be seen in Figure 2-3b, high-frequency words $n$ lead to shorter fixation durations on word $n + 1$ and this frequency-based preview benefit effect was significantly stronger for short compared to long words n. This was particularly the case for readers of shuffled text.

#### 2.3.3.2 Relative Lag-frequency effects – regression probability

To follow up on the reversed frequency effects for short words, we tested how word $n$ length, word $n$ frequency, and their interaction influenced the probability of regressing back to word $n$ after having fixated word $n + 1$ once. We fitted a generalized (logistic) linear mixed model using regressions from word $n + 1$ to word $n$ (after a first-pass single fixation on word $n$ and one fixation on word $n + 1$) as the binary dependent variable (N = 21,129 fixations). Predictors in the model were word $n$ frequency and length (i.e., 1/ wl), their interaction, frequency of word $n + 1$, as well as interactions of these variables with experimental condition (shuffled vs. normal PSC readers) using crossed random intercepts over subjects and over unique word id.

The effect of word $n$ frequency (i.e., of the regression target) on regression probability significantly differed between shuffled and normal PSC reading ($b$ = -0.35, $SE$ = 0.07, $p$ < .001). Readers of normal text made significantly more regressions to low-frequency compared to high-frequency words (i.e., a negative frequency effect; $b$ = -0.17, $SE$ = 0.05, $p$ < .001). Readers of shuffled text, to the contrary, made significantly more regressions to high-frequency compared to low-frequency words (i.e., a positive frequency effect; $b$ = 0.18, $SE$ = 0.05, $p$ < .001; see Figure 2-3c). We further tested how the word $n$ frequency effect depended on word length in the two reading conditions and found a marginally significant interaction for readers of shuffled text ($b$ = 0.85, $SE$ = 0.48, $p$ = .08). Post hoc tests revealed that readers of shuffled text made more regressions to short high-frequency compared to short low-frequency words (i.e., a positive frequency effect for short words; $b$ = 0.24, $SE$ = 0.06, $p$ < 0.001) while this frequency effect was not significant for long words $n$ ($p$ = .27). In readers of the normal PSC the frequency effect did not depend on word length ($p$ = .26). Note that differences in word $n + 1$ frequency between reading conditions cannot be the source of these effects because this was statistically controlled for in the regression model.

## 2.4 Discussion

Eye movements in reading are affected by both low-level visual and oculomotor factors as well as higher-level cognition related to language processing. With the present work we introduce the shuffled-text reading task as a new paradigm to investigate the interplay of low-level and high-level factors in reading. In the reported experiment, the words of a well-investigated corpus of single sentences (PSC, Kliegl, et al., 2004; Kliegl, et al., 2006) were randomly shuffled to create meaningless word lists. For each shuffled sentence, words from different original sentences were randomly selected. Participants' task was to read the presented text. To ensure that participants complied with the instructions, about a third of the trials were followed by a comprehension question (normal sentences) or a word recognition probe (shuffled word lists).

The eye movements of participants reading these shuffled meaningless sentences were compared with those from participants who read the normal meaningful PSC sentences. A detailed statistical analysis of variables known to modulate fixation times (cf., Kliegl, 2007; Kliegl, et al., 2006) showed various similarities and differences between the two tasks. Overall, our predictions as outlined in the Introduction were supported by the experimental results.

First, there was a considerable degree of similarity in the eye movements between readers of shuffled and normal text. We investigated how seven visuomotor variables influenced single fixation durations: the length of the fixated word n, the length of the last word $n - 1$, and the length of the next word $n + 1$, the amplitudes of the incoming and outgoing saccades, and the slope and location of the fixation-duration inverted-optimal viewing position (IOVP) effect. We found no evidence that these influences on single fixation durations differed between readers of shuffled and normal text, with only one marginal difference for the length of word $n - 1$. This finding is consistent with our assumption that similar visual and oculomotor processes were in place when reading shuffled and normal text.

Second, there was no current-word frequency main effect on fixation times when reading shuffled text. This is surprising, but in line with work by Rayner and colleagues who found no effect of word frequency on eye movements in a task where participants searched for a target word in normal text (Rayner & Fischer, 1996; Rayner & Raney, 1996). The absence of word-frequency effects in visual search suggests that lexical word processing does not influence eye movements in this task. Was lexical processing also irrelevant for eye guidance when reading shuffled text? Although low-frequency words did not receive longer fixations than high-frequency words overall, we nevertheless found several strong and expected effects of word frequency on fixation durations during shuffled-text reading. In particular, effects of distributed processing, i.e., the influence of lag- and successor-word frequency, the quadratic effect of current-word frequency (Kliegl, 2007), and the coefficient for the interaction of current-word frequency with word length were highly reliable and more or less unchanged during reading of shuffled as compared to reading of normal text. Overall, low-frequency words were not looked at longer when reading shuffled text. However, this standard effect of word frequency was present for long words (see Figure 2-3). At the same time, we found reversed effects of current-word frequency on gaze durations for short words (Figure 2-3) and on single fixation durations. In these cases, fixations were longer on high- than on low-frequency words, which is opposite to what is found in normal reading. Taken together, these effects suggest that readers of shuffled text processed words lexically and that lexical word processing influenced their eye movements.

Notably, the probability of making a between-word regression as well as the distributions of leftward-oriented saccades were virtually identical for the two reading conditions [see Appendix A, Figure A-1b]. This striking agreement in distributions is well in line with the notion that most regressive eye movements when reading easy normal sentences like the PSC are triggered by unfinished word recognition (cf., Engbert, et al., 2005; Nuthmann & Engbert, 2009). However, we cannot exclude the possibility that additional post-lexical processes, assuming that they might occur in one way or the other when reading shuffled text, may trigger the same amount and the same distribution of regressive eye movements in both tasks.

Third, we found support for our predictions with regard to slower processing. All measures of fixation durations (single, first, and second fixations as well as gaze durations) were significantly increased when reading shuffled as compared to normal text. Also, we observed a reduced skipping rate along with a strong increase in refixation probability. First and foremost, we attribute these results to the fact that the shuffling procedure removes the predictability of words. In addition, post-lexical integration and memorization of words should be harder in shuffled text, potentially contributing to the slower reading speed. In particular, the observed increase in second- and more-pass reading fixations may reflect active attempts of readers to try and memorize words and/or understand meaningless shuffled text. Another effect hinting towards memorization processes in shuffled-text reading is the stronger effect of word length as compared to normal reading. As longer words take more time to encode phonologically (Baddeley, Thomson, & Buchanan, 1975), the stronger word length effect would be in line with the idea that readers encode words in the phonological loop when reading shuffled text.

Further in-depth analyses revealed very specific processing differences between the two tasks. We argue that the reported pattern of results supports the hypothesis that the perceptual span was more strongly modulated by foveal load among readers of shuffled text compared to readers of normal sentences. In the following, we provide a detailed discussion of the results with respect to distributed processing (Section 2.4.1), the modulation of the perceptual span (Section 2.4.2), alternative explanations for changed frequency effects (Section 2.4.3), and PG vs. SAS models (Section 2.4.4).

### 2.4.1 Replication of effects of distributed processing

Recently, Kliegl et al. (2006) used corpus analyses to investigate the influence of the foveal word *n* as well as of neighboring words *n* – 1 and *n* + 1 on fixation durations on word n. They reported strong and consistent parafovea-on-fovea effects, yet their validity has been questioned (Rayner, Pollatsek, et al., 2007) (but see Kliegl, 2007). Much of the criticism pertained to the correlational nature of the reported lag and successor effects. Here, we counter this argument by reporting robust and highly reliable effects of distributed processing for readers of shuffled text. When creating the shuffled word lists, each word was selected at random from all words in the corpus, and this random selection was done for each participant separately. Thus, observed effects are experimental in nature and allow the conclusion that processing neighboring words *n* – 1 and *n* + 1 causally affected fixation durations on the fixated word n. The effects of neighboring words on fixation durations on word *n* were highly similar in normal and shuffled-text reading. This (a) suggests that these effects generalize to other reading situations, and (b) supports the validity of these effects in normal sentence reading.

### 2.4.2 A stronger modulation of the perceptual span in shuffled-text reading

Our prediction was that readers of shuffled text should primarily focus on the processing of salient low-frequency content words to better remember them for the recognition task. From a perspective of a theoretical framework supporting parallel processing of words in the perceptual span (e.g., SWIFT, Engbert, et al., 2005), such a strategy predicts a stronger modulation of the perceptual span by foveal load (Henderson & Ferreira, 1990) during reading of shuffled text. This prediction was supported by our findings.

#### *2.4.2.1 Relative lag effect*

The primary prediction derived from the foveal load hypothesis (Henderson & Ferreira, 1990) (see also Balota, et al., 1985; Inhoff, et al., 1989; Inhoff & Rayner, 1986; Rayner & Pollatsek, 1987) states that the difficulty of a word *n* – 1 (e.g., its frequency) modulates the amount of preview that is available for the next word *n* during fixation on word *n* – 1. High-frequency words *n* – 1 would allow strong preprocessing of word *n* during the previous fixation. This preview can be measured by the benefit of having seen a correct compared to an incorrect preview during the previous fixation (Henderson & Ferreira, 1990). In corpus analyses, extensive parafoveal preprocessing of word *n* during fixations on word *n* – 1 should attenuate the current-word frequency effect on word n. Previous words *n* – 1 of low-frequency should result in a strong current-word frequency effect, while high-frequency words *n* – 1 should go along with weaker current-word frequency effects (cf., Kliegl, et al., 2006). In the present data, this interaction was stronger for readers of shuffled compared to normal text. We conclude that the modulation of the perceptual span is stronger in readers of shuffled text than in readers of normal text. Readers of shuffled text widen their perceptual span more strongly when fixating a word of high-frequency and focus their attention more strongly when reading a low-frequency word. To follow up on this hypothesis, we derived several qualitative predictions from a parallel model of word processing during reading (assuming that the total amount of processing resources is limited).

#### 2.4.2.2 Current-word frequency effects

The data supported the prediction that current-word frequency effects should be weaker if the modulation is stronger. In fact, when reading shuffled text, the frequency effect completely disappeared (gaze durations) or even turned into a small positive effect (single fixation durations). This is a noteworthy finding, because the negative word-frequency effect for fixation times (longer fixations on low-frequency than on high-frequency words) is one of the cornerstones of research on gaze control in reading (e.g., Altarriba, Kroll, Sholl, & Rayner, 1996; Calvo & Meseguer, 2002; Henderson & Ferreira, 1990, 1993; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennison & Clifton, 1995; Kliegl, 2007; Kliegl, et al., 2004; Kliegl, et al., 2006; Raney & Rayner, 1995; Rayner, 1977; Rayner, et al., 2004; Rayner & Duffy, 1986; Rayner & Fischer, 1996).[8] The effect reflects the longer processing times associated with low-frequency words as compared to high-frequency words.

We will now propose an explanation for the pattern of current-word frequency effects observed in the present data. According to the foveal load hypothesis, a low-frequency word *n* captures more attentional resources than a corresponding word of high-frequency. This should not only modulate the preview for the next word, but also reduce the additional time that is needed to process the low-frequency word. If the allocation of additional processing resources is strong enough (i.e., if the additionally captured resources are equal to the additional processing demands), this mechanism is capable of canceling out any immediacy effects of word frequency on fixation durations. In its most extreme version, a strong dynamical modulation of the perceptual span could even produce reversed, that is positive, effects of current-word frequency on fixation durations.

---

[8] Going beyond fixation durations during reading, word frequency also affects word processing in many other psycho-linguistic tasks. That words, which occur frequently in a given language, are recognized more easily than words that appear less frequently is perhaps the single most robust finding in the whole literature on visual word recognition. The basic result holds across the entire range of laboratory tasks used to investigate reading. For example, frequency effects are seen in lexical decision [. . .], in naming [. . .], semantic classification [. . .], perceptual identification, [. . . and] spoken word recognition [. . .] and therefore appear to be a central feature of word recognition in general (Norris, 2006, p. 327) (also see e.g., Monsell, 1991; Murray & Forster, 2004; Whaley, 1978).

#### 2.4.2.3 Relative successor effect

According to the foveal load hypothesis, the parafovea-on-fovea frequency effect from word *n* + 1 should depend on the frequency of the currently fixated word n. Previous studies did not find such an interaction (e.g., Henderson & Ferreira, 1990; Kliegl, et al., 2006). We replicated this null effect for normal sentence reading. However, we observed a significant interaction for readers of shuffled text (see Figure 2-4b for gaze durations and Figure 2-5b for single fixation durations). This finding lends further support to the interpretation that the dynamical modulation was stronger in readers of shuffled text compared to readers of sentences.

#### 2.4.2.4 Effects of relative current-word frequency

The foveal load hypothesis further predicts that the strength of the current-word frequency effect depends on the length of the fixated word. The frequency effect should be stronger for long words than for short words, which is schematically illustrated in Figure 2-1. For long words *n* of high (Figure 2-1a) or low (Figure 2-1b) frequency, it is more or less only the currently fixated word *n* that falls into the perceptual span. As a consequence, the effects of lexical processing are fully visible in the current-word frequency effect. Indeed, we found a strong frequency effect for long words in both reading conditions (Figure 2-3a). The situation is different for short words. According to the foveal load hypothesis, a short low-frequency word is read with a narrowly focused perceptual span (Figure 2-1d). In this case, all processing resources are focused on word n. If the currently fixated word *n* is not only short but also high-frequent, the perceptual span should be enlarged, such that also the upcoming word *n* + 1 falls into the span (Figure 2-1c). Under the assumption of constant processing resources, this distribution of attention across two words can slow down the processing of the currently fixated word n, modulating the frequency effect observed for short words n. For normal reading, we found a small standard (i.e., negative) frequency effect for short words (Figure 2-3a). For the shuffled text, this effect turned into a positive effect such that low-frequency words were actually fixated shorter than high-frequency words. Thus, the foveal load hypothesis is compatible with our experimental findings.

#### 2.4.2.5 Lag effects of relative word frequency

Another prediction that directly follows from such reasoning is that the preview for the upcoming word *n* + 1 should depend on the interaction of word *n* frequency and length. As noted above, long words *n* fill more or less the whole perceptual span regardless of their frequency. As a consequence, preview for word *n* + 1 will barely differ between conditions of low (Figure 2-1a) and high (Figure 2-1b) foveal load. Accordingly, word *n* frequency should not strongly influence first fixation durations on the next word *n* + 1. Indeed, we found weak effects of word *n* frequency for readers of normal and for those of shuffled text if word *n* was long (Figure 2-3b). Again, the situation is different for short words n. During fixations on short low-frequency words *n* the perceptual span is narrow and does not allow for much preprocessing of the next word (Figure 2-1d). For short and high-frequency words *n* the next word *n* + 1 largely falls into the perceptual span (Figure 2-1c). Strong parafoveal processing in this condition will reduce the processing needs for word *n* + 1 when fixating on it. Thus, foveal load during fixations on short words should strongly influence the amount of parafoveal preprocessing. Empirically, the effect of word *n* frequency on first fixation durations on word *n* + 1 was strong for short words in both reading conditions, but stronger for readers of shuffled text (Figure 2-3b). Thus, fleshing out the foveal load hypothesis within a parallel processing framework makes an interesting double-prediction concerning frequency effects of short words n: Word *n* frequency should weakly influence fixation durations on word *n* (or even show a reversed influence), but should strongly affect fixation durations on word *n* + 1. Thus, there should be a trade-off between the two effects. The data support this prediction, as both effects are stronger for readers of shuffled compared to normal text.

#### 2.4.2.6 Regression probability

We examined how often readers regressed back to word *n* after having fixated word *n* + 1 once (and after having made a first-pass single fixation on word n) (Figure 2-3c). Readers of normal text generated significantly more regressions to low-frequency words. In contrast, when reading shuffled text more regressions were made to short words of high-frequency compared to short words of low-frequency. Thus, in shuffled-text reading, short high-frequency words did not only receive longer fixation times, but also more regressions than short low-frequency words.

Henderson and Ferreira (1990) demonstrated that the perceptual span is modulated by foveal load in normal reading. Here we applied this foveal load hypothesis to derive a qualitative model of our results on shuffled-text reading. It turned out that the foveal load hypothesis, combined with a processing gradient (PG) model of eye-movement control, provides a coherent theoretical for the explanation of a set of complicated and highly interacting effects. When reading difficult (i.e., low-frequency) words, shuffled text readers focus their attention so strong that they process these words even faster than easy (i.e., high-frequency) words. Likewise, processing of the next word is reduced. When fixating easy (high-frequency) words, on the other hand, readers of shuffled text widen their perceptual span such that high-frequency words – in particular if they are of short length – are fixated longer and attract more regressions compared to short words of low-frequency. At the same time parafoveal processing of word $n + 1$ is enhanced and fixation times on this word are reduced.

Why do we observe a stronger dynamical modulation of the perceptual span for readers of shuffled texts? As we speculated in the Introduction, readers of shuffled text may have focused on the processing of salient low-frequency content words when trying to remember the words in the shuffled text. In contrast, they may have widened their perceptual span when encountering high-frequency words because they did not expect to be probed about these words. Such processing would in fact be a good strategy because only content words, but not function words were queried in the memory task. It may be that readers were aware of this fact and adapted their processing to optimize the processing of task-relevant words. In sum, we propose that (a) a strong focus on low-frequency content words coupled with (b) limited processing resources that are spatially distributed via a dynamically modulated attentional gradient can lead to the disappearance or reversal of word-frequency effects during the reading of shuffled text.

### 2.4.3 Alternative explanations for changed frequency effects

As one of our findings, under certain conditions the effect of current-word frequency was strongly attenuated or even reversed when reading shuffled text. We argued that the dynamics of attention modulation in a PG model can qualitatively explain such an effect and the conditions under which it should occur. However, it could still be that frequency effects were reversed not because of the dynamics of attention modulation and eye-movement control but because high-frequency words were more difficult to process when reading shuffled text than low-frequency words. For example, short high-frequency words might slow down reading and attract regressions because they have more high-frequency orthographic neighbors, or because function words (as opposed to content words) are difficult to process when encountered in shuffled text. However, control analyses showed that these specific characteristics of short high-frequency words were not responsible for the observed patterns of results (see Online supplementary material).

Specific memorization processes related to the mirror effect (e.g., Reder et al., 2000) may provide another alternative explanation for why word-frequency effects were reversed. When studying a list of unrelated words, words of low-frequency were shown to be easier to recognize than words of high-frequency (e.g., Reder, et al., 2000). It has repeatedly been shown that the effect is specific to retrieval and does not hold during encoding (e.g., de Zubicaray, McMahon, Eastburn, Finnigan, & Humphreys, 2005; Diana & Reder, 2006; Miozzo & Caramazza, 2003; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Naveh-Benjamin & Guez, 2000; Rao & Proctor, 1984). However, it is possible that subjects are aware of their better recognition performance for low-frequency words. Thus, it could be that even though high-frequency words are more easily identified, readers actually invest more time in memorizing these words. This could potentially lead to reversed effects of word frequency because high-frequency words are fixated longer or because readers make more regressions to these words.

Finally, we hasten to emphasize that high-frequency words were not generally processed longer than low-frequency words. Effects of word frequency were often in the expected direction (see e.g., effects of successor and lag frequency). They were reversed only under very specific circumstances, in particular for short words. In addition, and critically, reduced or reversed effects of word *n* frequency (on fixation durations on word *n* and regression probability) were associated with an enhancement of these effects on fixation durations on word *n* + 1. Thus, a generally increased processing difficulty for high-frequency words cannot be responsible for the specific pattern of results in the present study.

### 2.4.4 PG and SAS models

We have shown that PG models incorporating the principles outlined above can, in principle, explain our results. A model like SWIFT might provide a parsimonious account based on a single mechanism, that is the modulation of the perceptual or attentional span by foveal load (see Engbert, 2007, for an implementation of the foveal load hyopthesis with the SWIFT model). In contrast, given their basic principles, SAS models would not naturally predict the effects reported here. In particular, finding strong, experimental effects of distributed lexical processing and not finding the standard current-word frequency effect and, under some conditions, finding reversed current-word frequency effects, is not readily explained by the E-Z Reader model (Reichle, Warren, et al., 2009).

As such, high-level effects of distributed processing provide a challenge for SAS models (Engbert & Kliegl, 2011; Kliegl, et al., 2006). These effects have been the subject of considerable debate (Rayner, Pollatsek, et al., 2007) (but see Kliegl, 2007). As many as about 50 variables are known to influence word recognition (see Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). The corpus analyses by Kliegl and colleagues, finding pervasive effects of distributed processing, included only a limited number of such variables: They tested the effects of frequency, length, and predictability of words. If any of the remaining, uncontrolled variables (e.g., of the fixated word n) were correlated with the frequency of the next word $n$ + 1, then corpus analyses could show significant successor effects of next-word frequency. However, these effects would, in fact, not stem from the processing of the next word $n$ + 1, but instead from lexical processing of the currently fixated word $n$ (cf., Rayner, Pollatsek, et al., 2007). Rayner and colleagues (2007) implemented this hypothesis to simulate results from Kliegl and colleagues (2006) with the E-Z Reader model. They assumed that the predictability of word $n$ + 1 was correlated with an unobserved variable influencing lexical processing of word n. Introducing this simple correlation was sufficient for the E-Z Reader model to show substantial effects of word $n$ + 1 predictability on fixation durations on word n. Introducing a similar correlation with word $n$ + 1 frequency would enable the E-Z Reader model to show substantial effects of word $n$ + 1 frequency on fixation durations on word n. As noted above (see Section 2.4.1), correlations between neighboring word properties are absent in shuffled text. Each word was randomly selected for each shuffled word list and each reader separately. Therefore, unobserved properties of the fixated word *n* cannot be systematically related to the frequencies of neighboring words. Thus, our results on distributed lexical processing are of experimental nature. They impose boundary conditions for computational models of reading.

According to proponents of the E-Z Reader model, simulations could in principle accommodate lexical influences from neighboring words if these were due to mislocated fixations (e.g., Rayner, Pollatsek, et al., 2007). In reading, due to oculomotor error in saccade programming a significant proportion of fixations are mislocated in that they fall on words to the left or right of the intended target word (Engbert & Nuthmann, 2008). In E-Z Reader, it is the intended rather than the fixated word that will receive lexical processing. However, we believe that numerical simulations are necessary to explore the possibility that mislocated fixations can induce parafoveal-on-foveal effects in SAS models. Some empirical evidence for the mislocation hypothesis has been reported (Drieghe, Rayner, & Pollatsek, 2008; subsequently challenged by Kennedy, 2008), but the results were not substantiated by quantitative estimates. Moreover, mislocated fixations trigger short-latency saccades to produce the fixation-duration IOVP effect (Nuthmann, et al., 2005). Thus, these short fixations are not triggered by lexical word processing, and should be independent of the frequency of the intended (neighboring) word. This, however, is inconsistent with Rayner et al.'s (2007) hypothesis that mislocated fixations cause effects of the neighboring (intended) word frequency on fixation durations on the fixated word *n* (cf., Kennedy, 2008).

In general, drawing conclusions from the shuffled-text reading task about theoretical models of eye-movement control is preliminary. First, numerical simulations of the models need to be carried out. Second, it is unclear at present how different cognitive processes (e.g., related to memory demands) influence eye movements during reading of shuffled text compared to normal text reading. Therefore, further empirical as well as computational research is needed to illuminate these issues.

## 2.5 Conclusion

In the present paper we introduced the shuffled-text reading paradigm as a new paradigm to study the interactive control of eye movements by higher-level cognitive and lower-level visuomotor factors. We found that a number of variables known to influence eye movements in reading showed similar effects when reading shuffled texts. Thus, the basic mechanisms of visuomotor and lexical processing are at work independent of whether meaningful sentences are presented or not. However, shuffled text has an impact on global parameter settings and modulates strategies for information processing in reading. We demonstrated two such influences. First, our findings add to the body of literature suggesting that the predictability of words eases their processing and speeds up reading (e.g., Balota, et al., 1985; Ehrlich & Rayner, 1981), albeit from a novel perspective. In the shuffled-text reading paradigm, word predictability is removed while word frequency remains intact. We showed that this manipulation of word predictability as well as potential differences in the memorization of words slowed down reading. The findings also contribute to the current debate about serial as opposed to parallel processing of words in a sentence (Engbert & Kliegl, 2011; Reichle, Liversedge, et al., 2009). We observed distinct experimental effects of spatially distributed processing (Kliegl, et al., 2006), indicating that several words are simultaneously affecting fixation duration at a time. These effects were more strongly modulated by foveal load in the shuffled reading task as compared to normal reading.

## 2.6 Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft (Grant EN 471/2-1). We would like to thank Reinhold Kliegl for his statistical advice, Sarah Risse for her support during data analysis and the writing process, and Keith Rayner for valuable comments on previous versions of this manuscript. We also thank Denis Drieghe and other anonymous reviewers for their helpful comments.

# 3 The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model

Daniel J. Schad* and, Ralf Engbert

*Department of Psychology, University of Potsdam, Potsdam, Germany*

* Corresponding author. Email: Daniel.Schad@uni-potsdam.de

*Running head*: A zoom-lens model of attention in reading

Published

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20* (4-5, Special Issue on Computational Approaches to Reading and Scene Perception), 391. doi: 10.1080/13506285.2012.670143

## Abstract

Assumptions on the allocation of attention during reading are crucial for theoretical models of eye guidance. The zoom lens model of attention postulates that attentional deployment can vary from a sharp focus to a broad window. The model is closely related to the foveal load hypothesis, i.e., the assumption that the perceptual span is modulated by the difficulty of the fixated word. However, these important theoretical concepts for cognitive research have not been tested quantitatively in eye movement models. Here we show that the zoom lens model, implemented in the SWIFT model of saccade generation, captures many important patterns of eye movements. We compared the model's performance to experimental data from normal and shuffled text reading. Our results demonstrate that the zoom lens of attention might be an important concept for eye movement control in reading.

*Keywords*: Computational modeling; Eye movements; Foveal load hypothesis; Perceptual span; Reading; Zoom lens model of attention

## 3.1 Introduction

How is attention allocated to the text during reading? This is one of the crucial questions driving experimental as well as theoretical research on eye movement control. Two classes of cognitive models can be distinguished based on the theory of attentional deployment that they incorporate. Serial attention shift models (SAS; e.g., E-Z Reader: Reichle, 2011; Reichle, et al., 1998; Reichle, et al., 2003) (see also Engbert & Kliegl, 2001) assume that an attention spotlight (Eriksen & Hoffman, 1972; LaBerge, 1983; Posner, 1980) focuses on a single word at a time (Inhoff, et al., 1989). In SAS models, the attentional spotlight shifts serially from one word to the next to move a reader's eyes through the text (for a recent overview see Reichle, 2011). Processing gradient models (PG; e.g., SWIFT: Engbert, et al., 2002; Engbert, et al., 2005; Glenmore: Reilly & Radach, 2006) propose that attention is allocated to a spatially extended region of the text to support parallel processing of several words at a time. In these models, the attentional gradient continuously drops off towards the visual periphery, where processing of visual stimuli is slowed (cf. Downing & Pinker, 1985; Shulman, Wilson, & Sheehy, 1985).

Both SAS and PG models of attentional deployment in reading can be combined with a prominent concept of selective visual attention formulated in the zoom lens model (Eriksen & St. James, 1986; LaBerge & Brown, 1989; N. G. Müller, Bartelt, Donner, Villringer, & Brandt, 2003). According to this model, the focus of visual attention can change in size, between sharply focusing on a narrow area and being widely distributed over a large part of the visual field. In reading, the zoom lens of attention is supported by the foveal load theory (Henderson & Ferreira, 1990), which postulates that the perceptual span (McConkie & Rayner, 1975; Rayner, 1975) is modulated by foveal processing difficulty. A key motivation for the development of a zoom lens model for reading is related to its prediction on effects of word frequency and word length on fixation durations. A modulation of the attentional span in a computational model can potentially decrease or even reverse these effects, since a broad span during a fixation on a high frequency word should slow foveal processing rate. Interestingly, such decreased and reversed effects of word frequency and word length have been found in a shuffled text reading paradigm (Schad, Nuthmann, & Engbert, 2010).

The perceptual span is the region of effective vision during reading and extends 3-4 letters to the left and about 14-15 letters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1998). It has been studied in the moving window paradigm (McConkie & Rayner, 1975), where only the fixated part of the text is visible to the reader, while the remaining text is covered with a mask that moves with the eyes. The foveal load hypothesis postulates that the size of the perceptual span is modulated by foveal load or the processing difficulty of the fixated word (Henderson & Ferreira, 1990). If foveal load is low, then the perceptual span is wide and text processing during one fixation extends over several neighbouring words. In the case of high foveal load, the perceptual span is small and only the fixated word is processed during a fixation. Support for the foveal load hypothesis comes from studies using the boundary paradigm (Rayner, 1975), where effects of target-word preview were observed only when processing the preboundary word was easy, but not when it was difficult (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; Schroyens, et al., 1999; White, et al., 2005). Foveal load effects can be explained based on zoom lens model of attention (Eriksen & St. James, 1986; LaBerge & Brown, 1989; Schad, et al., 2010). Applying the model to reading, the assumption is that foveal processing controls the focus of the zoom lens.

First, we will review the main results from the recent study on shuffled text reading (Schad, et al., 2010). In particular, Schad et al. (2010) discussed specific hypotheses about how a zoom lens model could account for differences in eye movement control between reading of shuffled and normal text. Second, we developed an advanced version of the SWIFT model (Engbert, et al., 2002; Engbert, et al., 2005) incorporating a dynamically-modulated processing span (SWIFT 3). Third, the model is applied to experimental data during reading of shuffled and normal text. Finally, we will carry out further explorative simulations of the model to investigate its predictions on experimental data.

### 3.1.1 Shuffled versus normal text reading

Schad et al. (2010) investigated eye movements during reading of normal and of shuffled text. To create shuffled text, words from the German Potsdam Sentence Corpus (PSC, Kliegl, et al., 2004; Kliegl, et al., 2006) were randomly shuffled. For each word list, words were drawn from the PSC without replacement such that different words in a list would normally stem from different original sentences in the PSC. This procedure was designed to reduce all local relations between words to chance level, e.g.,

Affen Vorschlag Armen schmale Giebel Kanzler dem besser.
*Monkeys suggestion poor/arms narrow gable chancellor the better.*

Jede ihrer Förster im Jahr Hunde meisten Gräfin Bauern.
*Each [of her/their] foresters [in the] year dogs most countess countrymen.*

To ensure that participants would read the words in the lists, they were occasionally given recognition probes for the words that had been contained in the last list.

Statistical analyses of eye movements revealed several interesting similarities and differences between normal and shuffled text reading. First, Schad et al. (2010) found reliable effects of spatially distributed word processing during reading of both normal and shuffled text. Specifically, word frequency and length of the upcoming word $N$+1 as well as of the preceding word $N$−1 affected fixation durations on the fixated word $N$, replicating *successor*- (word $N$+1) and *lag*- (word $N$−1) effects from normal text reading (Kliegl, 2007; Kliegl, et al., 2006). Different from corpus analyses of normal text reading, in shuffled texts word neighbourhood is randomized. Therefore, effects of distributed word processing in shuffled text are of experimental nature and are not confounded with characteristics of the fixated region (Rayner, Pollatsek, et al., 2007; Schad, et al., 2010).

Eye movements during shuffled and normal text reading also showed pronounced differences. As mentioned earlier, standard effects of current-word frequency and length were reversed during reading of shuffled text. During normal sentence reading, fixations are longer on long words than on short words (Just & Carpenter, 1980; Rayner, et al., 1996). Likewise, readers usually look longer at low frequency than at high frequency words (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner & Duffy, 1986). For shuffled text, however, both of these standard effects were absent and even reversed. Surprisingly, fixation durations were longer for short words as compared to long words, and readers looked longer at high frequency words than at low frequency words. These effects are intriguing, as effects of word frequency belong to the most reliable and widely found effects in psycholinguistic and eye movement research (Rayner, 1998, 2009). Schad et al. (2010) did not have a good explanation for reversed effects of word length. Concerning word frequency effects, we discussed the hypothesis that word frequency might be reduced during shuffled text reading, because the signal to move the eyes is less affected by lexical word processing (cf. Rayner & Fischer, 1996). However, lexical influences were reliable as we found expected effects of word frequency, e.g., for the previous word *N*-1 (lag-frequency effects) and the upcoming word *N*+1 (successor-frequency effects), which suggested that lexical processing of these words affected eye movements. Moreover, word frequency effects were reversed in some conditions. Alternatively, Schad et al. argued that the new effects in shuffled text reading could be explained parsimoniously by a foveal load or zoom lens model: Based on analyses of statistical models, we derived the hypothesis that the perceptual span is more strongly dynamically modulated by foveal load for readers of shuffled text than for readers of normal text.

## 3.2 SWIFT 3: The zoom lens of attention in the SWIFT model

In the SWIFT model (Engbert, et al., 2002; Engbert, et al., 2005), a set of word-based activations controls saccade target selection, and commands to program saccades are generated by a random process. To adjust the processing time to the difficulty of the fixated word, an inhibitory control process, called foveal inhibition, was implemented (see also Engbert & Kliegl, 2011; Richter, et al., 2006).

A key motivation to develop an activation-based model for the control of eye movements in reading was to derive an integrative framework for all types of saccades (i.e., forward, skipping, refixation saccades, and regressions). In close analogy to the dynamic field theory (Erlhagen & Schöner, 2002), the activation field determines probabilities for target selection at any point in time. This concept guarantees the existence of movement targets independent of the timing of upcoming saccade programs. Such a framework is essential for building models that implement the partial independence of spatial ("where" to move the eyes) and temporal ("when" to move the eyes) decisions on saccadic eye movements, conceptually required from models of the oculomotor physiology (Findlay & Walker, 1999).

For the simulation studies on the zoom lens of attention, we modified the processing span of the model. We assume that a letter-based processing rate is an inverse-parabolic function with two parameters that determine the extension of the processing span to the left and to the right. The processing span extends to $-\delta_L$ on the left and to $\delta_R$ on the right of the fixation point at the origin. Moreover, we assume that the asymmetry of the processing span is generated by a dynamical adjustment of the extension to the right, i.e.,

$$\begin{aligned} \delta_R &= \delta_0 \left( 1 + \delta_1 \left( 1 - \frac{a_k(t)}{A} \right) \right), \\ \delta_L &= \delta_0 \end{aligned} \tag{1}$$

where $a_k(t)$ denotes the time-dependent activation of word $k$ at time $t$; $A$ is the maximum of the activation reflecting the maximum possible word difficulty in the model. For the simulations, it turned out that an inverse-parabolic form of the processing span was necessary to constrain its spatial extent by experimental data during the simulations. Using such a functional form, the letter-based processing rate at an eccentricity $\varepsilon$ was given by

$$\lambda(\varepsilon) = \lambda_0 \begin{cases} 0 & : \quad \varepsilon < -\delta_L \\ 1 - \varepsilon^2 / \delta_L^2 & : \quad -\delta_L \le \varepsilon < 0 \\ 1 - \varepsilon^2 / \delta_R^2 & : \quad 0 \le \varepsilon < \delta_R \\ 0 & : \quad \delta_R < \varepsilon \end{cases}, \tag{2}$$

where a normalization constant $\lambda_0 = 3/(2(\delta_L + \delta_R))$ is necessary to scale the total processing rate to one (independent of the values of $\delta_L$ and $\delta_R$).

For the simulations, we implemented a fully stochastic framework proposed recently by Trukenbrod and Engbert (submitted). In this framework, all dynamical variables are realized by independent, parallel discrete random walk processes (Figure 3-1).

A typical numerical output of a single reading trajectory of the SWIFT model is displayed in Figure 3-1, by plotting the time evolution of several model states and processes along the vertical axis. In the main panel of Figure 3-1, vertical lines below each word represent the set of lexical activations $\{a_n(t)\}$ and the thick dashed vertical line shows the fixation location $k(t)$. The sequence of words fixated in this example is

{1, 3, 3, 2, 4, 5, 6, 7, 9}.

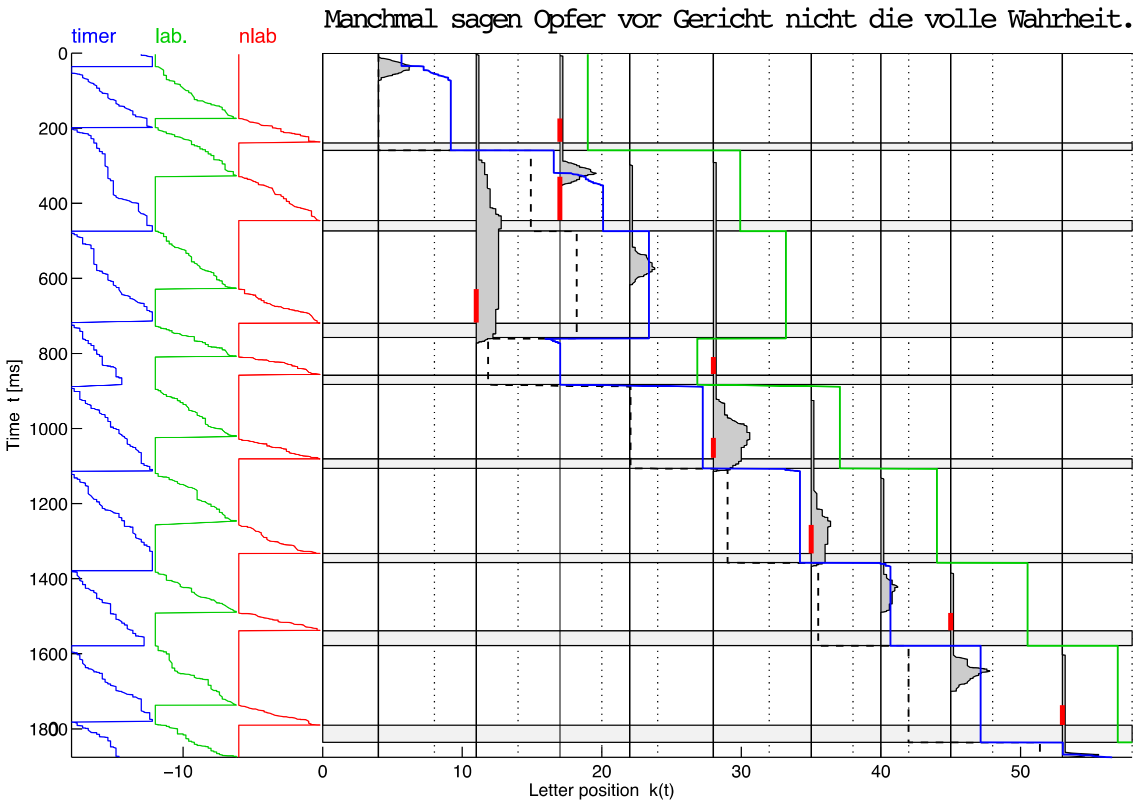


**Figure 3-1**

Simulated trajectory of the SWIFT model with attentional zoom lens.

The blue and the green lines indicate the extension of the perceptual span to the right of fixation. The green line marks the extension of the perceptual span for nonlexical preprocessing of words, which has been estimated as extending 15 letters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1998). The blue line represents the rightward extension of the lexical word processing span. During preprocessing of foveal words, that is, in the increasing phase of the lexical word activation, the word processing span is at a fixed minimum. After preprocessing of the foveal word is completed, however, then the lexical processing span is dynamically modulated by the lexical activation of the fixated word (Equation 1). Highly activated foveal words cause the processing span to be narrow in size. If foveal lexical word activation is reduced, however, then the span size dynamically increases up to its estimated maximum size.

The three lines in the left panel of Figure 3-1 display the states of sequentially coupled, directed random walk processes, which evolve over time. From the left to the right, the first random walk process displays the evolution of the random saccade timer and shows how evidence for a new saccade program accumulates over time. Note that the random oculomotor timer is subject to inhibition from foveal lexical word activations. Foveal activation temporarily inhibits the progression of the random walk and delays the onset of the next saccade program. Second, when the random oculomotor saccade timer reaches its threshold, a labile saccade program is triggered. At the end of the labile saccade program, a saccade target is determined and saccade programming enters into its stabile phase (red bars in the main panel of Figure 3-1 indicate the selected saccade target and their length represents the duration of the labile programming stage). Finally, a saccade is executed, during which visual input from the retina is suppressed (see Figure 3-1, the horizontal grey bars).

Additional new parameters were related to (1) a global inhibition (ppf) that slowed processing of words to the right as long as nonvanishing activations were to the left of the word considered (iota), (2) a partial reset of activation during the increasing part of processing during a saccade, and (3) a reduction of the processing rate by a constant factor (f) during postlexical processing, i.e., the decreasing part of the activation.

### 3.2.1 Predictions for shuffled text reading

Schad et al. (2010) proposed specific hypotheses about how eye movement control differs between shuffled and normal text reading. Here, we will test these qualitative predictions on a fully quantitative basis by estimating parameters of the SWIFT 3 model separately for normal and for shuffled text reading. Schad et al. (2010) hypothesized that the control of eye movements may be less affected by ongoing lexical processing when reading shuffled text. In the SWIFT model, the $\beta$ parameter determines how strongly lexical processing (i.e., word frequency) influences word activations. We therefore predict that the $\beta$ parameter should be reduced in the shuffled-SWIFT model as compared to the SWIFT model for normal text reading. Moreover, in the SWIFT model processing of foveal words influences eye movements via foveal inhibition of the autonomous saccade timer and we predict that this influence (captured in model parameter h) is reduced for shuffled-SWIFT. Moreover, we suggested that the perceptual span is more strongly modulated by foveal load during reading of shuffled text as compared to reading of normal text. In the SWIFT 3 model, the δ1 parameter determines how strongly the processing span is modulated. We predicted that the $\delta_1$ parameter should be larger for shuffled-SWIFT than for normal-SWIFT.

To test hypotheses, we defined a procedure and a set of criteria designed to avoid potential pitfalls associated with model fitting (see Appendix D), including a split-half procedure to guard against overfitting, where independent data sets are used to (1) optimize model parameters (on a training set) and to (2) evaluate model predictions (on a test set).

### 3.2.2 Results from parameter estimation

When fitting the shuffled-SWIFT model to the training set, we defined measures of fixation times and probabilities separately for each subject (see Engbert et al., 2005, for the key principles of our procedure). Experimentally, the text had been randomly shuffled separately for each of the 30 subjects (Schad, et al., 2010). For each subject, we thus computed word-based measures of fixation durations and probabilities for 850 words of the subject-specific version of the shuffled corpus (all words except for the first and the last word per list). This procedure represents eye movements at the level of individual fixations and saccades. The model produced averages over 20 model runs of single, first, second, and total fixation durations as well as probabilities for skipping, two fixations, three or more fixations, and the number of regressions for each word and for each subject separately. These simulations demonstrate that it is possible to fit a cognitive model of eye movement control (SWIFT 3) to data at the level of individual eye movements. This is an advantage compared to earlier simulation studies (e.g., Engbert, et al., 2005).

Estimated parameter values from the training set for the normal-SWIFT and the shuffled-SWIFT model (see Figure 3-2 and Table 3-1) corresponded to our qualitative predictions. The lexical parameter $\beta$ was smaller for shuffled-SWIFT than for normal-SWIFT. The $\beta$ parameter approached zero for shuffled-SWIFT, indicating that lexical influences on word activities were strongly reduced for shuffled text readers. In addition, foveal inhibition was reduced for shuffled-SWIFT (smaller $h$ parameter). Taken together, these results are compatible with the view that the influence of cognition on eye movements is reduced in shuffled text reading. Second, the $\delta_1$ parameter was larger for shuffled-SWIFT than for the normal-SWIFT model. This indicates that the dynamical modulation of the processing span was stronger for shuffled than for normal text reading. In its current formulation (see Equation 1), the dynamical modulation of the processing span depends on the size of both, the $\delta_0$ parameter and the $\delta_1$ parameter. To get an estimate of how strongly the span differs between its focused and its defocused state, independent of the overall size of the span, we derived a new parameter $\delta_{1\text{additive}}$. This parameter was calculated from the estimated values for the span-parameters $\delta_1$ and $\delta_0$, via

$$\delta_{1\text{additive}} = \delta_1 \, \delta_0 \, . \tag{3}$$

Substituting $\delta_{1\text{additive}}$ for $\delta_1$ in Equation 1 yields,

$$\delta_R = \delta_0 + \delta_1 \left(1 - \frac{a_k(t)}{A}\right). \tag{4}$$

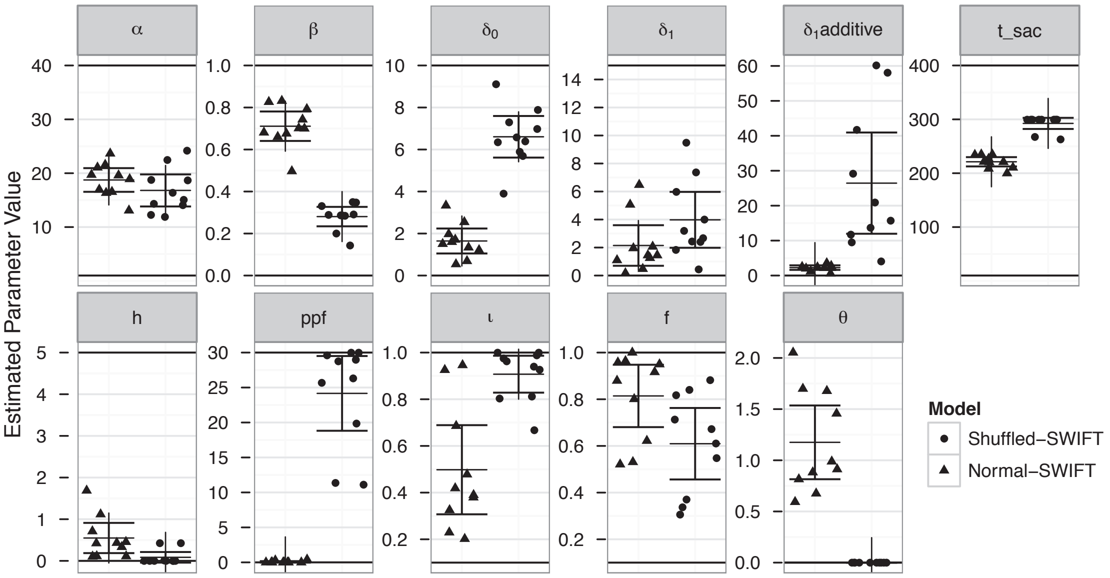


**Figure 3-2**

Results from the estimation of model parameters for the normal-SWIFT (triangles) and the shuffled-SWIFT (points) models. A genetic algorithm running for 13,000 generations was used to estimate individual sets of model parameters. This was repeated 10 times with random starting values for normal-SWIFT and for shuffled-SWIFT. Points/triangles show results from individual estimation runs; midlines indicate the average of the parameter estimates across 10 estimation runs; error bars indicate 95% confidence intervals.

**Table 3-1**

SWIFT parameters for shuffled and normal text reading.

| | Parameter | Symbol | Shuffled Text | | Normal Text | | t-test (Welch) | Range[b] | Equa tion |
|---|---|---|---|---|---|---|---|---|---|
| | | | *M*[a] | *SE*[a] | *M*[a] | *SE*[a] | p-value | | |
| Lexical parameters | Frequency, intercept | $\alpha$ | 16.8 | 1.3 | 18.7 | 0.97 | .25 | 1–40 | |
| | Frequency, slope | $\beta$ | 0.28 | 0.02 | 0.71 | 0.03 | .000 | 0–1 | |
| | Predictability | $\theta$ | 0 | - | 1.18 | 0.16 | | 0–4.6 | |
| | Global inibition | $ppf$ | 24.16 | 2.36 | 0.069 | 0.047 | .000 | 0–30 | |
| Visual processing | Visual span, constant | $\delta_0$ | 6.61 | 0.44 | 1.64 | 0.26 | .000 | 0–10 | 1, 4 |
| | Visual span, dynamic | $\delta_1$ | 3.98 | 0.88 | 2.15 | 0.64 | .112 | 0–15 | 1, 4 |
| | Visual span, dynamic - additive[c] | $\delta_{1additive}$ | 26.45 | 6.40 | 2.27 | 0.29 | .004 | | 3 |
| | Visual span, preprocessing | $pspan$ | 15.0 | - | 15.0 | - | - | - | |
| | Word length exponent | $\eta$ | 0.3 | - | 0.3 | - | - | - | |
| | Preprocessing factor | $f$ | 0.609 | 0.068 | 0.814 | 0.059 | .04 | 0.1–1 | |
| | Global decay | $\omega$ | 0.01 | - | 0.01 | - | - | - | |
| | Transfer across Saccades | $\iota$ | 0.91 | 0.03 | 0.50 | 0.08 | .001 | 0.1–1 | |
| | Eye-Mind lag | $pcd$ | 30.0 | - | 30.0 | - | - | - | |
| Saccade timing | Random timing (ms) | $t_{sac}$ | 292.5 | 4.6 | 221.2 | 3.8 | .000 | 10–400 | |
| | Random timing – starting value | $t_{sac0}$ | 0.8 | - | 0.8 | - | - | - | |
| | Inhibition factor | $h$ | 0.085 | 0.056 | 0.549 | 0.160 | .02 | 0–5 | |
| | Target selection weight | $\gamma$ | 1.0 | - | 1.0 | - | - | - | |
| Saccade programs | Labile stage (ms) | $\tau_{lab}$ | 100.0 | - | 100.0 | - | - | - | |
| | Refixation factor | $refix$ | 0.7 | - | 0.7 | - | - | - | |
| | Mislocated Fixation factor | $misfac$ | 0.75 | - | 0.75 | - | - | - | |
| | Nonlabile stage (ms) | $\tau^0_{nl}$ | 50.0 | - | 50.0 | - | - | - | |
| | Latency Modulation | $\kappa_0$ | 2.5 | - | 2.5 | - | - | - | |
| | Latency Modulation | $\kappa_1$ | 0.3 | - | 0.3 | - | - | - | |
| | Saccade execution (ms) | $\tau_{ex}$ | 30.0 | - | 30.0 | - | - | - | |

*Notes.* [a] Means and standard errors over 10 independent runs of the genetic algorithm (for each run using the optimal parameter set from the last 500 generations). Other parameters were fixed for previous model versions or for theoretical reasons, and no standard errors are given for these parameters. [b] Parameter boundaries used in fitting. [c] Parameter was calculated from fitted values.

The results showed that the zoom-lens response was much stronger in shuffled-SWIFT than in normal-SWIFT, even when controlling for task-differences in the (focused) size of the perceptual span. We also found a larger global inhibition in the shuffled-SWIFT model (larger *ppf* parameter). This result is highly plausible because words in a list are unrelated, and this should cause strong interference when multiple words are simultaneously processed. This stronger global inhibition may also cause stronger foveal load effects in shuffled text reading (Schad et al., 2010), as foveal processing difficulties inhibit processing of upcoming words.

We also obtained the following parameter differences between the shuffled-SWIFT and the normal-SWIFT model: For shuffled-SWIFT, the processing span was estimated to have an overall larger size, as reflected in a larger $\delta_0$ parameter. Also, the average rate of the autonomous oculomotor timer, $t_{sac}$, was estimated to be larger in the shuffled-SWIFT model compared to normal-SWIFT. This effect is clearly related to the slower speed at which shuffled text is read, either due to a mindless eye movement control (Reichle, et al., 2010; Vitu, et al., 1995), because words cannot be predicted from the preceding context, or due to postlexical processing (e.g., memory encoding) of shuffled text. The latter interpretation is also supported by a smaller $f$ parameter in shuffled-SWIFT, indicating that postlexical processing is slowed relative to lexical word processing. Thus, estimates for SWIFT parameters indicated that readers did engage in postlexical word processing when reading shuffled texts. Lastly, the $\iota$ parameter was increased during shuffled text reading, suggesting that early visual representations were better transferred across saccades. It may be that this higher stability in visual input for shuffled texts results from the stronger global inhibition in this task. If processing of upcoming words succeeds against competing representations from other words, then the resulting representations may be more stable compared to normal reading, where global inhibition is small. Finally, the $\theta$ parameter determines influences of word predictability on eye movement control. We set the $\theta$ parameter to zero for the shuffled-SWIFT model because words cannot be predicted from their preceding context in shuffled text. The estimated value of $\theta$ for the normal-SWIFT model was consistent with estimates based on previous model versions.

### 3.3 Simulation results

To evaluate model performance, we compared model predictions to empirical data from the test set. The summary results are computed from 300 runs of the models.

For normal text, we simulated 300 runs of the model for the Potsdam Sentence Corpus. The shuffled text corpus was randomly shuffled for each subject separately, such that each subject read a different corpus of shuffled text. The corpus of each single subject in the test sample was simulated with 20 runs of the model, yielding a total of 300 model simulations for 15 different versions of the shuffled corpus.

First, we investigated predictions of the SWIFT 3 model for normal and for shuffled text with respect to distributions of (1) fixation durations, (2) saccade lengths, and (3) within-word landing positions (effects on the preferred viewing location, PVL), and effects of within-word landing position on (4) refixation probabilities (optimal viewing position effect, OVP), and (5) fixation durations (inverted optimal viewing position effect, IOVP). Details of analyses and results are provided as *Supplementary Information* (available at http://read.psych.uni-potsdam.de/pmr2/). Overall, predictions of the SWIFT and the shuffled-SWIFT models were successful in reproducing standard effects on eye movements in normal and shuffled text reading, respectively. For distributions of fixation durations, model simulations were in good agreement with experimental results. The shuffled-SWIFT model captured the increase in mean and variance of fixation duration distributions during shuffled text reading. Likewise, distributions of forward- and backward-oriented saccade lengths were well reproduced by the SWIFT 3 model. For shuffled text reading, forward-directed saccades were clearly shortened for the experimental data and this effect was reproduced qualitatively by the shuffled-SWIFT model. The SWIFT 3 model also reproduced landing position distributions (including a leftward-shift in the preferred viewing location, PVL, for shuffled text), and the OVP effect on refixations. Moreover, SWIFT 3 successfully predicted stronger IOVP effects in shuffled text reading for single and first of multiple fixation durations, and these predictions are parameter free and arise from the model architecture.

### 3.3.1 Word-based measures: Effects of word length and word frequency

We focused on summary statistics of how current word length and frequency affect diverse eye movement measures during normal and shuffled text reading (see Figure 3-3). We were interested to investigate whether simulations of the SWIFT model reproduce reversed effects of word length and frequency on fixation durations during shuffled text reading. Figure 3-3 demonstrates that the model simulations qualitatively reproduced task differences in measures of fixation durations and probabilities, as well as in effects of word length and frequency.

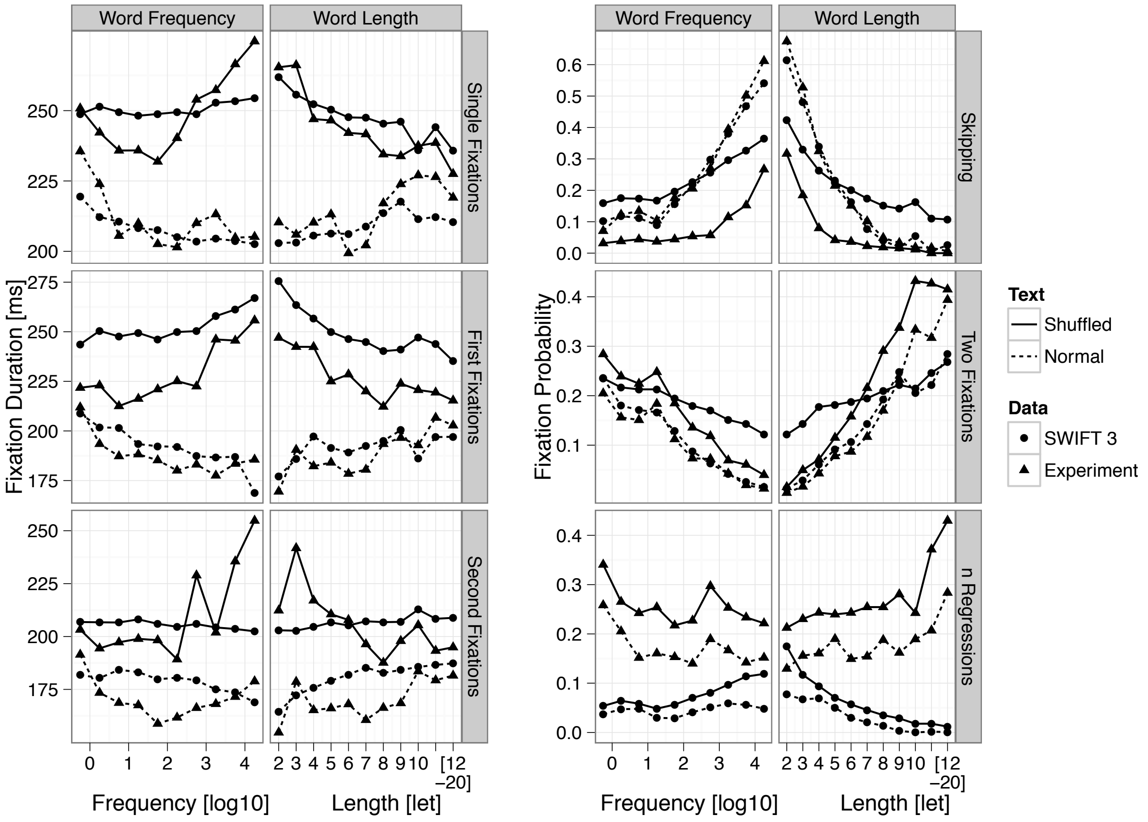


**Figure 3-3**

Effects of word length and frequency on different measures of fixation durations and probabilities for model simulations (points) and experimental data (triangles) of shuffled (solid lines) and normal (dashed lines) text reading. Left panel: Mean durations of single, first, and second fixations. Right panel: Mean probabilities for skipping and two fixations, and the mean number of between-word regressions.

Readers of shuffled text exhibit prolonged fixation durations on all measures, including single fixation durations, first of multiple fixation durations, and second fixation durations. The simulations of the model reproduced all of these differences. Moreover, SWIFT captured the influences of word frequency and length on fixation durations during normal text reading. As is usually found in reading studies, fixation durations in normal text were longer on long compared to short words and they were longer on low frequency words compared to high frequency words. These standard results were also present in the model simulations for all fixation duration measures. In experimental data on shuffled text reading, effects of word length and frequency were reversed for all measures of fixation durations. Readers looked longer at high frequency words than at low frequency words, and similarly, readers spent more time fixating short than long words in shuffled text. These effects were reproduced by simulations of the shuffled-SWIFT model: Simulated fixation durations showed reversed effects of word frequency and length for single and first fixation durations, but no effect for second fixation duration.

For measures of fixation probabilities, the model qualitatively reproduced experimental results. Experimentally, word skipping is at a very low rate for long and for low frequency words, but strongly increases for short and/or high frequency words (Brysbaert & Vitu, 1998; Rayner, 1998). This skipping pattern was present for both shuffled and normal text reading, and was also present in the simulation results. Skipping probability was strongly reduced for readers of shuffled text, which was basically driven by a strong reduction in skipping of short and high frequency words. Simulations of the shuffled-SWIFT model captured this task-effect: In the simulated eye movements, word skipping was also considerably reduced. Empirically, readers make more refixations on long and on low frequency words as compared to short and high frequency words, and the SWIFT models for both, normal and shuffled text reading, reproduced these effects. The refixation rate was also overall higher in shuffled text reading, and the shuffled-SWIFT model reproduced this effect. However, the model underestimated the amount of refixations on long and low frequency words, but overestimated refixations on short and high frequency words for shuffled text. Mismatches between model predictions and experimental data in skippings and refixations may have been caused by the large perceptual span in shuffled-SWIFT. The SWIFT model also generated regressive between-word saccades. The model, however, did not adequately capture the overall number of regressions and the effects of word length and frequency, suggesting that postlexical processes that are currently not implemented in the SWIFT model may contribute to regression behaviour. We conclude that the SWIFT 3 model qualitatively reproduced benchmark results on eye movements during first-pass reading of normal and of shuffled text, including reversed length and frequency effects for shuffled text.

### 3.3.2 Distributed processing effects

Much research has been carried out under the *immediacy assumption* that primarily current word processing affects fixation durations during reading (Morrison, 1984; Rayner, 1998). However, several recent studies have found effects of *spatially distributed word processing* (Inhoff, et al., 2005; Kennedy & Pynte, 2005; Kliegl, 2007; Kliegl, et al., 2006; Kliegl, et al., 2007; Risse & Kliegl, 2011) and these effects and their interpretation have been subject to considerable debate (e.g., Pollatsek, et al., 2006a; Rayner, Pollatsek, et al., 2007). Corpus analyses of normal text reading have found reliable effects of the upcoming word *N*+1 (successor effects) and the previous word *N*-1 (lag effects) on fixation durations on the fixated word *N* (Kliegl, et al., 2006). The validity of these findings has been called into question by Rayner et al. (2007). In corpora of normal text, word neighbourhood is not under experimental control, making it difficult to control for potential confounds associated with neighbouring words. Different from normal text, word neighbourhood is under experimental (random) control in shuffled text. In this more highly controlled context, we have replicated effects of distributed word processing from normal reading (Schad et al., 2010), supporting the *distributed processing assumption*. Here, we investigate predictions of the SWIFT model for effects of distributed processing during shuffled and during normal text reading (Figure 3-4). Overall, the qualitative pattern of effects is well replicated by the model.

### Lag effects

Empirically, the length of word *N*-1 exerts a very strong influence on single fixation durations on word *N*, such that single fixation durations are longer if word *N*-1 was long. Likewise, frequency of word *N*-1 strongly affects single fixation durations on word *N*, with longer fixations after low frequency words *N*-1. Both of these effects are consistent across tasks and similar for normal and for shuffled text reading. These strong lag effects are also present in data simulated by the SWIFT model for both reading tasks. Several mechanisms are responsible for producing the effects. First, a fixation on a long word *N*-1 will generate less preview for word *N*, and, consequently, prolong fixations on word *N*. Second, the processing span will be smaller on average during previous fixations if word *N*-1 is a low frequency word compared to the case when it is a high frequency word. This also reduces the amount of preview that is available and prolongs fixation durations on word *N*. Third, foveal inhibition slows the progress of the random saccade timer. Depending on whether a saccade program is already running, foveal inhibition can either affect the saccade timer for the current, or for the next saccade. If no saccade program is active, then slowing the autonomous saccade timer will prolong the current fixation duration. If a (labile or nonlabile) saccade program has already been started, then foveal inhibition will prolong the duration of the next fixation. Thus, lexical processing from word *N*-1 can spill over into longer fixation durations on word *N*.

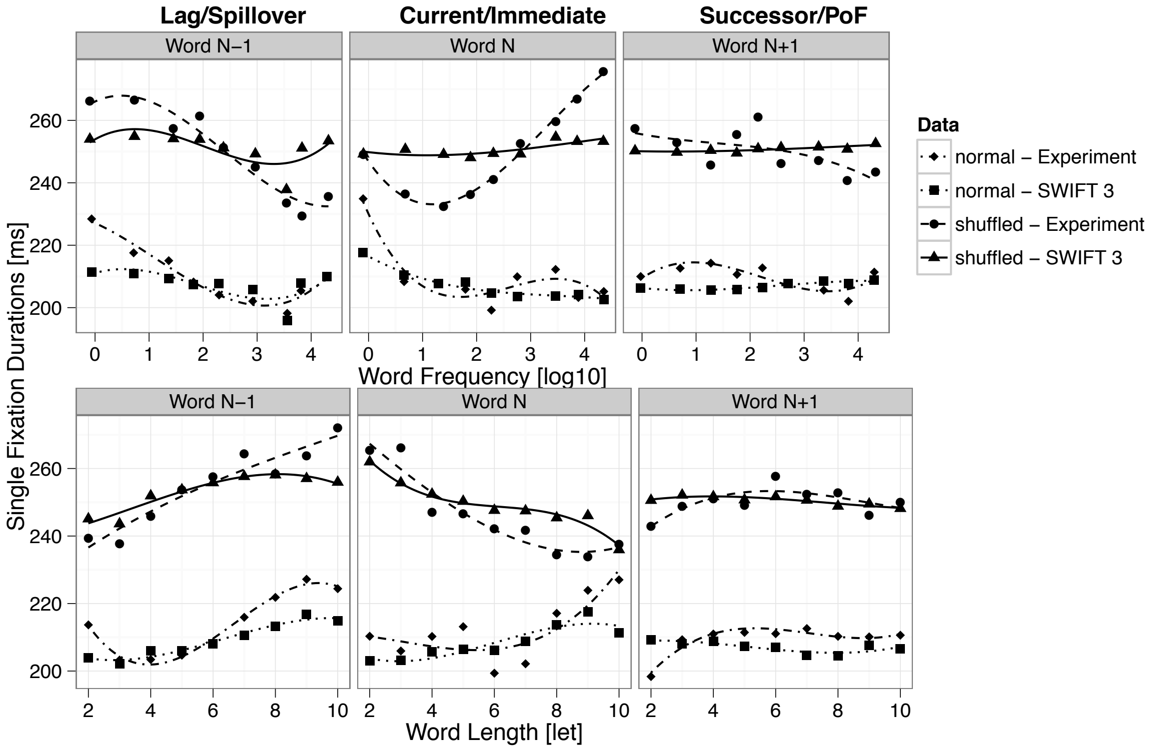


**Figure 3-4**

Analysis of distributed processing effects for model simulations of shuffled (triangles & solid lines) and of normal (squares & dotted lines) text and experimental data on shuffled (points & dashed lines) and normal (diamonds & dot-dashed lines) text reading. Top row: Average single fixation durations on word *N* as a function of word frequency of the previous word (word *N*–1, left column), the current word (word *N*, middle column), and the next word (word *N*+1, right column). Predictions from separate regression analyses involving cubic effects on averaged data for each condition are shown. Bottom row: Corresponding plots as a function of word length.

### Successor effects

The SWIFT model contains no explicit mechanism for modulating fixation durations as a function of processing upcoming words *N*+1. Interestingly, the model nevertheless shows effects of the upcoming word *N*+1, due to selection effects. Specifically, the likelihood for a refixation depends on lexical activation of the next word *N*+1. As the lexical activation of word *N*+1 is a function of the fixation duration on word

*N* and lexical processing of word *N*+1, the durations of single fixations and of the first of multiple fixations can exhibit selection effects from word *N*+1 processing. In addition, the intended saccade length could generate small effects of parafoveal processing by influencing saccade programming time. For long words *N*+1, the intended saccade length may on average be larger, and saccade programming will be faster. This effect can cause longer fixation durations before short words (and before high frequency words, due to the correlation between word length and word frequency).

#### Current word effects

Interestingly, for shuffled text reading effects of distributed processing are dissociated from immediacy effects. Lag and successor effects are in the same direction as in normal text reading, while current word effects are reversed for shuffled text reading. In Figure 3-4, this is visible as the effects for words *N*-1 (Figure 3-4, left panel) and *N*+1 (Figure 3-4, right panel) are highly similar between normal and shuffled text reading. Effects for the current word, to the contrary, strongly differ between shuffled and normal text reading (see Figure 3-4, central panel).

### 3.3.3 Model prediction: Fixation durations before skipping

In this section, we investigate model predictions for fixation durations before word skipping. In SWIFT 3, we presented a mechanism to explain the pattern of skipping costs and benefits observed in reading studies. It is a theoretically interesting question whether average fixation durations before word skippings are longer (skipping costs) or shorter (skipping benefits) compared to fixation durations before normal forward saccades to the next word *N*+1 (Drieghe, et al., 2004; Hogaboam, 1983; Kliegl, 2007; McConkie, et al., 1994; Pollatsek, et al., 1986; Pynte, et al., 2004; Radach & Heller, 2000; Reichle, et al., 1998; Risse & Kliegl, 2011). Kliegl and Engbert (2005) investigated this question and found reliable skipping benefits for short and for high frequency words in a highly controlled statistical analysis. Their results show that skipping costs are typical for long and low frequency words, whereas skipping benefits are reliable for short and high frequent words. Our present analyses for normal text reading are based on a subset of the data used by Kliegl and Engbert (2005) and we here replicate their basic findings (Figure 3-5).

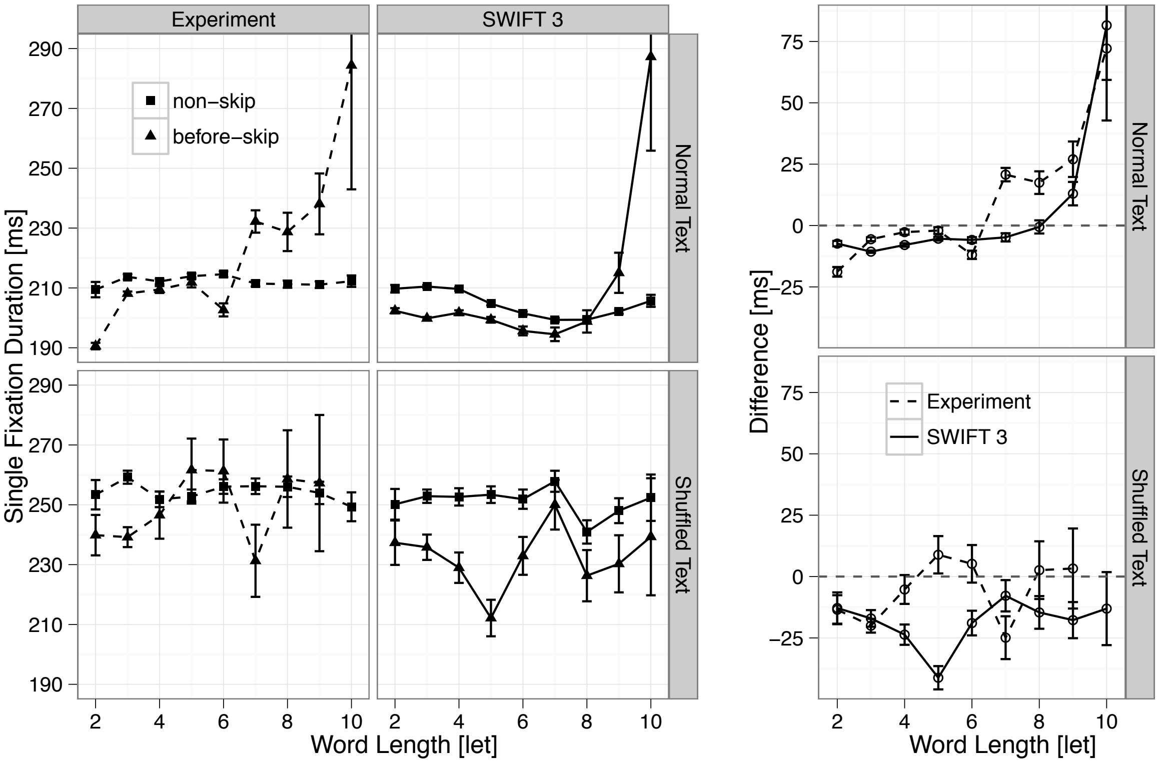


**Figure 3-5**

Single fixation durations before skipping (solid triangles) and nonskipping (solid squares) saccades as a function of word length of the skipped word for observed (left panel, dashed lines) and simulated (central panel solid lines) data during normal (upper panel) and shuffled (lower panel) text reading. The right panel displays the skipping difference in single fixation durations [SFD before skipping – SFD before nonskipping] for experimental (dashed lines) and simulated (solid lines) data, where positive difference values indicate skipping costs, and negative difference values indicate skipping benefits. Error bars are cell-based SEM.

Mathematical models of eye movement control have predicted skipping costs (e.g., E-Z Reader: Reichle, et al., 1998), and this was also the case for previous versions of the SWIFT model (SWIFT 2: Engbert, et al., 2005). Figure 3-5 shows that the SWIFT 3 model successfully produces skipping benefits for short words during normal text reading. Moreover, the model predicts skipping costs for long words, which is well in line with the observed data. For shuffled text, effects of word skipping on fixation durations were less stable due to the smaller amount of available data. To get a reliable estimate of skipping costs and benefits in shuffled text reading, we combined data from both sub-samples of the experimental data (training set and test set) for our analysis. Figure 3-5 shows that during shuffled text reading, skipping benefits are present for short words as has been observed for normal text reading. However, the effects for long words differed from those during normal text reading. For long words, we did not observe reliable skipping costs for readers of shuffled text.

The SWIFT 3 model also makes predictions about differences in skipping costs and benefits between tasks. For shuffled text reading, Figure 3-5 shows that the SWIFT 3 model successfully predicts the skipping benefits observed for short words. For long words in shuffled text, moreover, the model correctly predicts the absence of skipping costs. This prediction is quite surprising, given that we had no theoretical reason a priori to expect the effect and given that skipping costs and benefits were not explicitly included in fitting of model parameters. That the model simulations nevertheless predict the effect lends strong support to the mechanisms generating skipping benefits and costs in SWIFT 3. Next, we will investigate these model mechanisms in more detail.

### 3.3.4 How specific are model predictions?

The previous analyses demonstrated that the SWIFT 3 model successfully reproduced key patterns of eye movements in shuffled and normal text reading. Based on the split-half procedure, we now investigate predictions for experimentally observed eye movements in a given test set by computing correlations between predicted and observed data. Predictions are based on (1) the SWIFT 3 model for the respective task, (2) experimental data observed in the other task, and (3) predictions from the SWIFT 3 model for the other task. As a minimal criterion for model validity, predictions based on the SWIFT model for the respective task (a) should be as good or better than predictions based on experimental data from the other task (b) or model predictions for the other task.

First, we used experimental data in normal text reading to predict data observed during shuffled text reading. For all measures of fixation probabilities correlations between predicted and observed values were very high (*r*s ≥ .85). These high correlations do not uncover clearly task-specific eye movement effects, and we therefore focus our analyses on fixation durations. Figure 3-6 and Table 3-2 show correlations between predicted and observed data. Eye movements during normal text reading were

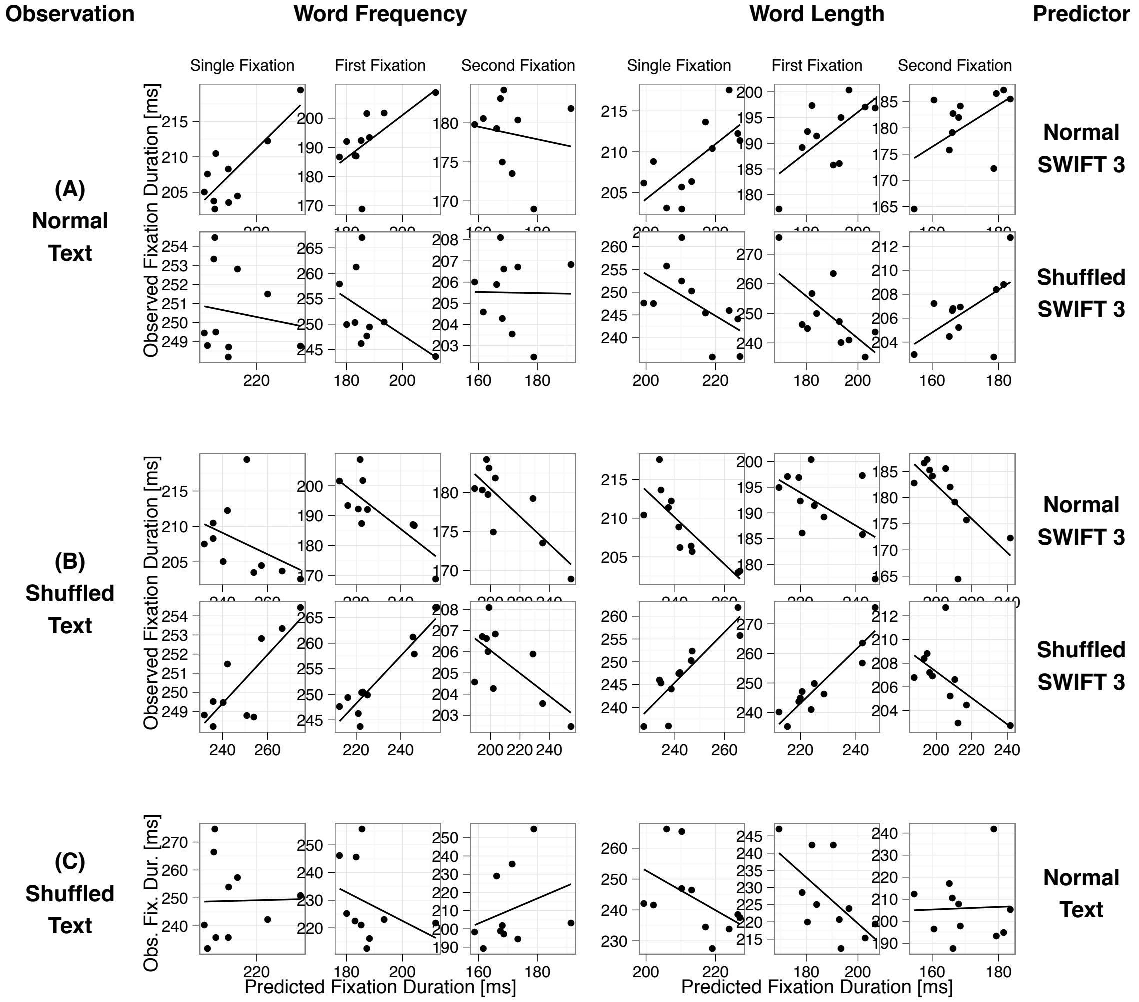


**Figure 3-6**

Shown are correlations between predicted and experimentally observed eye movement measures in the test data sets for normal and for shuffled text. Plotted at the ordinates are observed fixation durations from two tasks (indicated in the left-most column): Normal text reading (A), and shuffled text reading (B, C). Plotted at the abscissae are predicted fixation durations, where predictions are based on different sources (which are indicated in the right-most column): Predictions are based on simulations of the normal-SWIFT model [(A)-upper panel and (B)-upper panel], simulations of the shuffled-SWIFT model [(A)-lower panel and (B)-lower panel], and observed data from normal text reading (C). To compute correlations between observed and predicted data, word frequency (left panel) and length (right panel) were split into bins (the same bins used in Figure 3-3) and correlations were computed over average fixation durations per bin. Analyses were repeated for different measures of fixation durations, including the durations of single fixations (Panels 1–4, counted from left to right), first of multiple fixations (Panels 2–5) and second fixations (Panels 3–6).

**Table 3-2**

Correlations between predicted and observed fixation durations in the test data sets for normal and shuffled text

| | *Single fixation duration* | | *First of multiple fixation duration* | | *Second fixation duration* | |
|---|---|---|---|---|---|---|
| *Split by word* | *length* | *frequency* | *length* | *frequency* | *length* | *frequency* |
| Experimental Data (Test Set) | | | | | | |
| (a) Prediction by normal-SWIFT | | | | | | |
| Normal | .67 | .87 | . 68 | .65 | .60 | -.12 |
| Shuffled | -.69 | -.36 | -.84 | -.77 | -.63 | -.73 |
| (b) Prediction by shuffled-SWIFT | | | | | | |
| Shuffled | .88 | .80 | .92 | .93 | -.59 | -.67 |
| Normal | -.56 | -.14 | -.68 | -.48 | .58 | -.01 |
| (c) Prediction by experimental data (normal text) | | | | | | |
| Exp. data (shuffled text) | -.50 | .02 | -.63 | -.34 | .04 | .29 |

*Notes.* To compute correlations, word length and frequency were split into bins (cf. Figure 3-3) and correlations were computed over average fixation durations per bin.

best predicted by simulations of the normal-SWIFT model (Figure 3-6A, upper panel). Correlations between predicted and observed values were generally positive and high (*r*s ≥ .60, except for one slightly negative correlation). However, predictions for fixation durations during normal text reading failed when based on the shuffled-SWIFT model (Figure 3-6A, lower panel) or on experimental data observed during shuffled text reading (Figure 3-6C). For these cases, correlations between predicted and observed data were low (all *r*s ≤ .29, one exception: *r* = .58) or negative (8 out of 12 correlations). Likewise, fixation durations during shuffled text reading were best predicted by simulations of the shuffled-SWIFT model (Figure 3-6B, lower panel). Correlations between predicted and observed values were very high and positive for single and for the first of multiple fixation durations (*r*s ≥ .80). Only effects in second fixation durations were not well captured by the shuffled-SWIFT model (*r*s ≈ -.60). Again, predictions based on the normal-SWIFT model (Figure 3-6B, upper panel) or on experimental data observed for normal text reading (Figure 3-6C) were not successful. Correlations with model predictions were all negative, and correlations with experimental data were negative or low (rs ≤ .29).

We conclude that parameter estimates for both models, the normal-SWIFT model and the shuffled-SWIFT model, captured task-specific effects. Critically, they did not only fit eye movements in the two tasks. Additionally, models highly successfully predicted eye movements in the test sets from the split half-validation procedure.

## 3.4 Simulation experiments: How does the dynamic processing span affect eye movements?

Next, we investigated the consequences of the zoom lens model for eye movements during reading. We had hypothesized that reversed effects of word length and frequency stem from a higher dynamic modulation of the processing span by foveal word activation, i.e., we had predicted that a larger $\delta_1$ parameter reduces or reverses the influence of word frequency on first-pass fixation durations. To test this prediction in the SWIFT 3 model, we manually decreased the dynamic modulation of the processing span: The $\delta_{1\text{additive}}$ parameter for shuffled-SWIFT (estimated as 26.45) was set to the value estimated for the normal-SWIFT model (2.27) and 300 model simulations with this reduced modulation of the processing span were performed. We thus disenabled the stronger span-modulation in shuffled-SWIFT.

The results from these model simulations are displayed in Figure 3-7. As expected, the reversed effects of current-word frequency on single fixation durations were absent in the simulations. Moreover, the reversed effects of current-word length also disappeared, suggesting that a strong zoom-lens response can also explain reversed effects of current-word length. We conclude that the zoom lens model of attention, implemented as a dynamic processing span in the SWIFT model, can explain (1) variations in the effects of current-word frequency and length on fixation durations and (2) dissociations between immediacy effects and effects of distributed processing.

Next, we investigated the consequences of the zoom lens dynamic for skipping costs and benefits. We analysed fixation durations before skipping when the strong dynamic modulation of the processing span was disabled in shuffled-SWIFT. As a result, the model did not show the observed skipping benefits any more. Instead, it produced skipping costs across nearly all word lengths (see Figure 3-8). This result suggests that the dynamic modulation of the processing span is a key mechanism generating skipping benefits in the SWIFT 3 model.

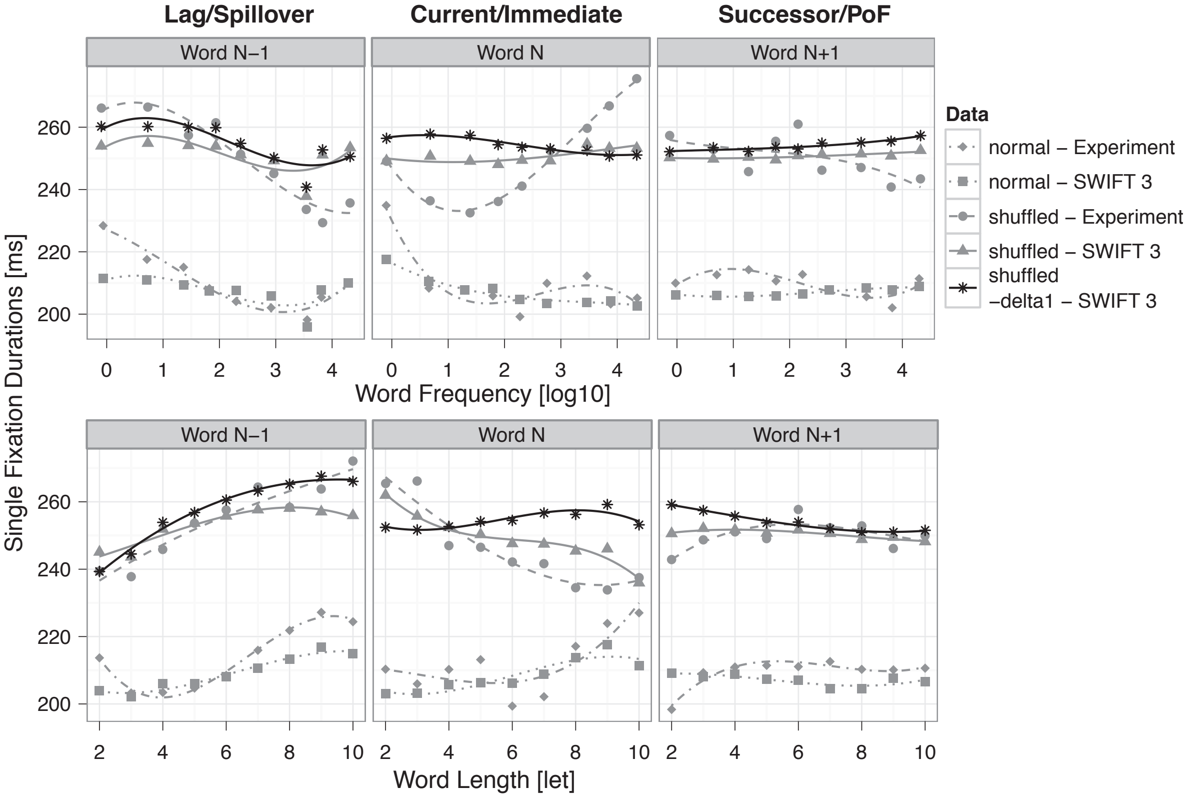


**Figure 3-7**

Effects of the dynamical processing span on spatially distributed word processing. Shown are the same results as in Figure 3-4 (grey), including model simulations of shuffled (triangles & solid lines) and of normal (squares & dotted lines) text and experimental data on shuffled (points & dashed lines) and normal (diamonds & dot-dashed lines) text reading. In addition, simulations of the shuffled-SWIFT model are presented, where the strong dynamic modulation of the processing span was disabled (black stars & solid lines, "shuffled – delta1"). Top row: Average single fixation durations as a function of word frequency of the previous word (word *N*–1, left column), the current word (word *N*, middle column), and the next word (word *N*+1, right column). Predictions from separate regression analyses involving cubic effects on averaged data for each condition are shown. Bottom row: Corresponding plots as a function of word length.

A possible mechanistic analysis of the origin of skipping benefits in the zoom lens version of the SWIFT 3 model is beyond the current study and will be published elsewhere.

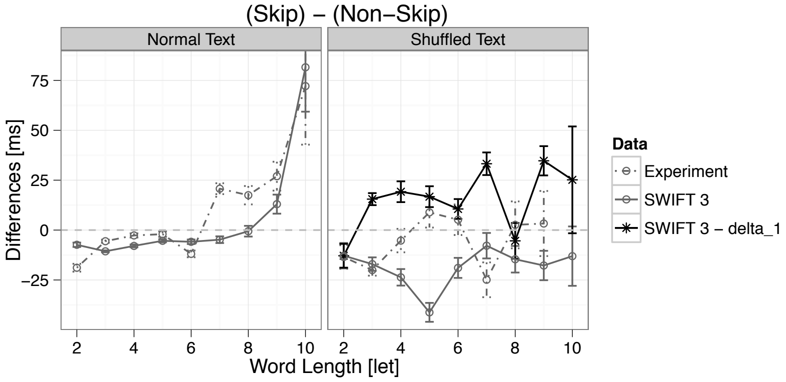


Figure 3-8

Effects of the dynamic processing span on skipping costs and benefits. Shown are the same results as in the right panel of Figure 3-5 (grey): Skipping-differences in single fixation durations (SFD before skipping – SFD before nonskipping) as a function of word length of the skipped word for normal (left panel) and shuffled (right panel) text reading for observed (dot-dashed lines & circles) and simulated (solid lines & circles). In addition, simulations of the shuffled-SWIFT model are presented, where the strong dynamic modulation of the processing span was disabled (black lines & stars, “SWIFT 3 – delta_1”, right panel).

## 3.5 General Discussion

In this paper, we developed and analysed a zoom lens version of the SWIFT model for eye movement control based on data from normal and shuffled text reading. We challenged the model with strong experimental eye-movement effects, like reversed effects of word length and frequency (Schad, et al., 2010). Both models, the normal-SWIFT and the shuffled-SWIFT variants (differing in parameter values only), were in good agreement with data related to standard effects of eye guidance in reading. The models reproduced distributions of (1) fixation durations, (2) saccade lengths, and (3) within-word landing positions (including effects on the preferred viewing location, PVL; Rayner, 1979) as well as (4) effects of within-word landing positions on refixation probabilities (optimal viewing position effect, OVP; O'Regan & Lévy-Schoen, 1987) and (5) on fixation durations (inverted optimal viewing position effect, IOVP; Nuthmann, et al., 2005, 2007; Vitu, et al., 2001) (see Supplementary Information, available at http://read.psych.uni-potsdam.de/pmr2/). Critically, in a split half-procedure model predictions were evaluated on a data set that was independent from the one used for parameter fitting to guard against overfitting (see the Appendix D for a procedure for model validation).

We found effects of word length and word frequency on fixation durations to be reversed in shuffled text reading, whereas the effects were in the standard direction in normal text reading, and these strong effects were well reproduced by the model simulations. For shuffled text, readers surprisingly looked longer at short words compared to long words, and they also looked longer at high frequency words than at low frequency words (Schad, et al., 2010). The model simulations qualitatively reproduced these reversed effects, and a simulation experiment showed that a strong zoom-lens response in shuffled-SWIFT was responsible for the success. This finding supports our previous hypothesis (Schad, et al., 2010) that SWIFT, as a parallel graded attention model, equipped with a zoom lens mechanism provides a theoretical framework that can explain reversed effects of word frequency. Moreover, it also uncovers a clear and strong, but previously unnoticed (Schad, et al., 2010) influence of attention modulation on effects of word length, a result that may inspire future tests of the dynamic processing span in the SWIFT 3 model.

Effects of spatially distributed processing in the model were in agreement with the observed data. Earlier work by Kliegl et al. (2006) and Schad et al. (2010) reported spatially distributed effects of word frequency and length in experiments on normal and on shuffled text reading. It is important to note that these effects are of experimental nature for shuffled text because word neighbourhood is under experimental (random) control. Distributed processing effects were highly similar between shuffled and normal text reading, and at the same time immediacy effects of word length and frequency qualitatively differed between tasks. The SWIFT 3 model successfully reproduced this empirical dissociation of distributed processing effects from immediacy effects.

The SWIFT 3 model was successful in reproducing experimentally observed fixation probabilities. In shuffled text word skipping was reduced and refixations were increased compared to normal text and the shuffled-SWIFT model reproduced these findings at a qualitative level. Moreover, parameter variations between the normal-SWIFT and the shuffled-SWIFT models reproduced standard effects of word length and frequency on word skipping and refixations.

### 3.5.1 Model predictions

Kliegl and Engbert (2005) analysed fixation durations before word skipping using an advanced statistical bootstrapping approach and discovered the systematic effect that skipping costs occur for long and for low frequency (target) words, whereas skipping of short and high frequency words produces highly reliable skipping benefits. Our simulations demonstrated that SWIFT 3 is the first model that can explain skipping benefits; in particular, SWIFT 3 predicted experimental skipping benefits for short words and predicted skipping costs for long words in normal text reading. For shuffled text, we also found reliable benefits for skipping of short words. Skipping of long words, however, was not associated with the costs that had been observed in normal text reading. This finding is very interesting because it is novel, because there was no theoretical reason to predict such an effect a priori, and because skipping costs and benefits were not explicitly included in the procedure for parameter fitting. Nevertheless, the SWIFT 3 model reproduced the absence of skipping costs in shuffled text reading.

### 3.5.2 What do we learn about shuffled text reading?

Randomly shuffling words in a corpus of text is a strong manipulation that may affect many different aspects of eye movement control during reading, including attentional, linguistic (lexical, syntactic, semantic), visual, and oculomotor processes. Here, we simultaneously investigated different control processes in a mathematical eye movement model. First, we tested the hypothesis (Schad, et al., 2010) that readers' eye movements are less strongly coupled to ongoing lexical word processing (see also Nuthmann, et al., 2007; Rayner & Fischer, 1996; Reichle, et al., 2010; Vitu, et al., 1995) when reading shuffled text, and this hypothesis was supported by the simulation results. For shuffled-SWIFT, the influence of lexical processing on word activations was reduced, together with a reduced foveal inhibition of the autonomous saccade timer. We conclude that eye movements are less coupled to ongoing lexical processing during shuffled text reading, leading to a more autonomous or "mindless" control of eye movements.

What factors may cause this processing difference between reading tasks? First, readers may scan over the (boring) shuffled word lists in the first pass at a rather superficial level, accepting the risk that some long or low frequency words are not completely processed. This strategy may indeed be efficient for shuffled text, where words need to be encoded for later recognition: Low frequency words have a benefit in recognition memory (Reder, et al., 2000), and processing low frequency words at a superficial level may therefore suffice to remember these words for the recognition probes. Alternatively, during reading of normal sentences, contextual (e.g., syntactic, semantic, or purely statistical) constraints ease the processing of individual words. This facilitation is not available in random lists of unrelated words. Therefore, lexical information may become available too late to reliably inform eye movement control.

Second, Schad et al. (2010) suggested that the perceptual span could be more strongly modulated by foveal load in readers of shuffled text as compared to readers of normal text, and our simulation results provided support for this prediction. If adaptive control of eye movements is *reduced* during reading of shuffled text (i.e., more autonomous control), we considered it surprising to find an *increased* adaptive control of the attentional focus (i.e., increased zoom-lens response). This result is interesting given that both mechanisms, attentional and behavioural control, share a common function during reading: They both provide means to adapt limited cognitive resources to local processing difficulties. Cognitive-saccadic coupling during normal reading allows for optimal control because reading proceeds fast for easy words, and difficult words are fixated long enough for sufficient processing. Similarly, focusing attention on low frequency words and defocusing attention for easy words also adapts processing to local needs. Based on this analysis, a strong dynamical modulation of the zoom lens during shuffled text reading may compensate for the mindless control of eye movements.

As an alternative, the strong modulation of the processing span in shuffled text may result from the serial nature of the shuffled text reading task. Shuffled text enforces a rather serial processing of words because none of the words can be predicted from the context. Accordingly, word skippings are strongly reduced and even very short and high frequency words are often fixated. When readers of shuffled text fixate on such words, which are processed easily, then it would be an optimal strategy to strongly widen the processing span to maximize preview of parafoveal words. The changed fixation patterns in shuffled text may therefore cause a stronger dynamical modulation of the processing span and a global increase in the size of the perceptual span, both of which were supported by our model simulations.

As a complementary finding, global inhibition was increased in the shuffled-SWIFT compared to the normal-SWIFT model, suggesting that inhibition is larger for unrelated words in a randomly shuffled list. This finding introduces a new and previously overlooked mechanism that may explain and contribute to foveal load effects when reading shuffled or normal text.

Third, despite the reduced cognitive-saccadic coupling in shuffled text reading, lexical and even postlexical processes seem to be intact, as was indicated by overall high word activations and a slowed deactivation of words in shuffled-SWIFT (see Lamme, 2003, for dissociations between awareness and attention). These results may indicate that readers attempt to memorize words for later recognition probes. Additionally, our simulation results suggest that visual and oculomotor processes in shuffled text reading may differ from normal text reading, as transfer of visual information across saccades was enhanced and the speed of the autonomous saccade timer was reduced in shuffled-SWIFT.

We developed a numerical simulation of eye movements during shuffled text reading based on the SWIFT model to capture important cognitive processes of eye guidance in this task. We take a parsimonious approach by using an existing model (SWIFT 3) to explain strong effects in a novel task (shuffled text reading) without adding post hoc assumptions about task-specific processes. An alternative strategy may be to introduce new task-specific assumptions to explain experimental results. For example, low frequency words have a benefit in recognition memory (the mirror effect, Reder, et al., 2000) and readers of shuffled text may use this fact to save encoding time on low frequency words. Note, however, (1) that previous research has found mirror effects for retrieval but not for encoding (e.g., Diana & Reder, 2006) and (2) that it may be difficult to reconcile a mirror-effect account with specific aspects of our findings, like the strong standard lag- and successor-effects. It would be interesting to implement and test this and other alternative accounts in the future, of course. To support these investigations, we provide all data, analysis scripts, and the computer code of SWIFT 3 via an online repository (see link later). From an experimental perspective, our simulations make specific predictions that need to be investigated in future experimental work, for example, testing attention allocation using the *boundary paradigm* (Rayner, 1975) or the *moving window paradigm* (McConkie & Rayner, 1975).

### 3.5.3 The zoom lens model of selective visual attention

The SWIFT 3 model demonstrates that the zoom lens model of selective visual attention (Eriksen & St. James, 1986; LaBerge & Brown, 1989) can add to the understanding of eye movement control in reading. It combines the concept of the zoom lens with the idea of a processing gradient. The zoom lens in SWIFT 3 has been inspired as an account for the foveal load hypothesis, which states that parafoveal preview depends on the difficulty of the fixated word (Henderson & Ferreira, 1990). As one of our key results, we demonstrated with the development of SWIFT 3 that a zoom lens-type modulation of the processing span by foveal load could reduce and even reverse effects of foveal processing difficulty (Schad, et al., 2010). Moreover, we showed that the zoom lens mechanism contributed to a mathematical explanation of systematic variations of skipping benefits and costs (Kliegl & Engbert, 2005).

### 3.5.4 Summary

In the present research, we studied eye movement control during reading of normal and shuffled text using an advanced version of the SWIFT model (Engbert, et al., 2005). Based on statistical analysis of eye movements, we previously (Schad, et al., 2010) derived hypotheses on differences in eye guidance between both reading tasks. Here, we quantitatively investigated these hypotheses. Our results demonstrate that the SWIFT 3 model generalizes to explain specific aspects of eye movements during shuffled text reading. They further support our hypothesis that during shuffled text reading, readers reduce adaptive control of eye movements, but increase their adaptive control of attention. Thus, the implementation of a new mechanism, the dynamic modulation of the processing span, in the SWIFT model turned out to be a powerful mechanism to explain effects in experimental data.

## 3.6 Acknowledgments

We thank Albrecht W. Inhoff, Keith Rayner and Greg Zelinsky for valuable comments on the manuscript. This work was supported by Deutsche Forschungsgemeinschaft (Grant EN 471/ 2 1 and FOR 868). Supplementary information as well as data, analysis scripts, and source code for SWIFT 3 simulations are available upon request and at the Potsdam Mind Research Repository, at http://read.psych.uni-potsdam.de/pmr2/.

# 4 Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels

Daniel J. Schad[a,*], Antje Nuthmann[b], Ralf Engbert[a]

[a] *Department of Psychology, University of Potsdam, Potsdam, Germany*

[b] *Department of Psychology, University of Edinburgh, Edinburgh, UK*

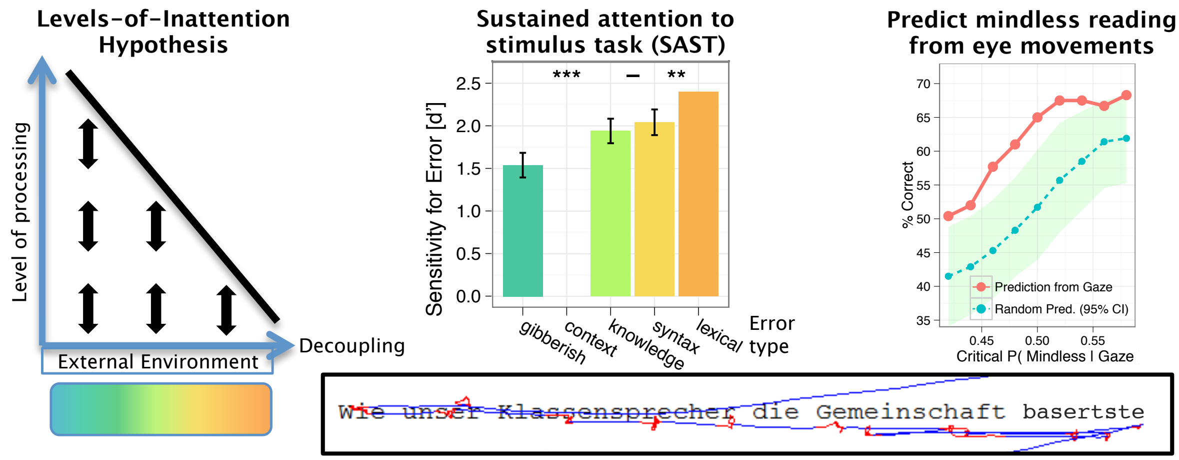
[*] Corresponding author. Email: Daniel.Schad@uni-potsdam.de

*Running head:* Objective measures reveal levels of inattention

Published

Schad, D. J., Nuthman, A., Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*(2), 179-194. doi: 10.1016/j.cognition.2012.07.004

## Graphical Abstract



## Highlights

- It is possible to predict states of mindless reading from eye movements
- As mind wandering reduces external attention cognitive processing is decoupled
- We propose the levels-of-inattention hypothesis postulating graded decoupling
- We developed the SAST to measure decoupling using psychophysics of error detection
- Results on different levels of mindless reading support levels of inattention

## Abstract

When the mind wanders, attention turns away from the external environment and cognitive processing is decoupled from perceptual information. Mind wandering is usually treated as a dichotomy (dichotomy-hypothesis), and is often measured using self-reports. Here, we propose the levels of inattention hypothesis, which postulates attentional decoupling to graded degrees at different hierarchical levels of cognitive processing. To measure graded levels of attentional decoupling during reading we introduce the sustained attention to stimulus task (SAST), which is based on psychophysics of error detection. Under experimental conditions likely to induce mind wandering, we found that subjects were less likely to notice errors that required high-level processing for their detection as opposed to errors that only required low-level processing. Eye tracking revealed that before errors were overlooked influences of high- and low-level linguistic variables on eye fixations were reduced in a graded fashion, indicating episodes of mindless reading at weak and deep levels. Individual fixation durations predicted overlooking of lexical errors five seconds before they occurred. Our findings support the levels of inattention hypothesis and suggest that different levels of mindless reading can be measured behaviorally in the SAST. Using eye tracking to detect mind wandering online represents a promising approach for the development of new techniques to study mind wandering and to ameliorate its negative consequences.

*Keywords*: Reading; Eye movements; Mind wandering; Signal detection theory; Levels of processing; Sustained attention

## 4.1 Introduction

Most people experience mental states in which they are no longer attending to the task at hand and are instead thinking about something else (Schooler, et al., 2011; Smallwood & Schooler, 2006). This ubiquitous phenomenon of mind wandering, which was long ignored in the cognitive sciences, has recently received considerable attention (Christoff, et al., 2009; Killingsworth & Gilbert, 2010; Levinson, Smallwood, & Davidson, 2012; McVay & Kane, 2010; Reichle, et al., 2010) and is thought to be tightly related to the brain's default mode of operation (Buckner, et al., 2008; M. F. Mason, et al., 2007). Mind wandering and task focus are typically treated as a dichotomy (Schooler, et al., 2011; Smallwood, 2010b; Smallwood, et al., 2011), where people are either mind wandering or focused on a given task. To investigate dichotomous aspects of mind wandering many previous studies have relied on subjective self-reports (Giambra, 1995; Smallwood & Schooler, 2006). Our main goal with the present work is to propose the levels of inattention hypothesis, which assumes that different hierarchical levels of cognitive processing are decoupled from external input in a graded fashion, reflecting states of deep and weak attentional decoupling. To measure different levels of decoupling during reading, we introduce a new paradigm, the sustained attention to stimulus task (SAST), which is based on signal detection analyses of readers' sensitivity for errors in the text. Analyzes of a large dataset of eye movements during mindless reading support the levels of inattention hypothesis and show that eye tracking technology can be utilized to predict states of mindless reading online.

The phenomenon of mind wandering involves two specific alterations in cognitive processing (Schooler, et al., 2011; Smallwood & Schooler, 2006). First, during mind wandering attention is directed away from the external environment (i.e., attention lapses), which reduces cognitive processing of perceptual information (Kam, et al., 2011; Smallwood, Beach, et al., 2008). This process of attentional (or perceptual) decoupling can lead to failures in the performance of external tasks (Christoff, et al., 2009; McVay, Kane, & Kwapil, 2009; Robertson, et al., 1997; Smallwood, et al., 2006). Second, mind wandering often involves stimulus independent thoughts (SIT) where attention is directed towards internal information derived from memory (Smallwood & Schooler, 2006; Stawarczyk, Majerus, Maquet, & D'Argembeau, 2011).

The cognitive sciences have described the mind as consisting of a multitude of different cognitive processes (Gazzaniga, 2009). As one important principle these processes are organized at different hierarchical levels, ranging from early low-level perceptual-motor processes towards increasingly abstract representations at higher levels (Cohen, 2000; Craik & Lockhart, 1972; Gazzaniga, 2009). For reading, various models – including models of eye-movement control (Engbert, et al., 2005; Reichle, Warren, et al., 2009) and theories of language processing (Graesser, et al., 2002; Kintsch, 1998; Malmkjaer, 2002) – have postulated hierarchical processing at visuomotor, lexical, syntactic, semantic, and discourse levels. How (in)attention affects different lower and higher levels of stimulus processing was long discussed in the debate about early (Broadbent, 1958) versus late (Deutsch & Deutsch, 1963; Treisman, 1960) attentional selection, and there is evidence that attentional selection can attenuate processing at early or late stages (Chun, et al., 2011; Lavie, 2005).

Mind wandering reduces external attention and can attenuate stimulus processing at all levels of the cognitive hierarchy (Smallwood, 2011b). This was demonstrated in studies investigating high-level episodic memory encoding (Riby, Smallwood, & Gunn, 2008; Smallwood, Baracaia, Lowe, & Obonsawin, 2003; Smallwood, McSpadden, & Schooler, 2008; Smallwood, et al., 2006), intermediate task-relevant stimulus processing (Barron, et al., 2011; O'Connell, et al., 2009; Smallwood, Beach, et al., 2008), early low-level multimodal perceptual processing (Kam, et al., 2011; Weissman, et al., 2006), and sensory input processes (Smilek, et al., 2010b). The present work concerns how these diverse findings can be integrated into a coherent theoretical framework.

The cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) proposes a mechanism to explain decoupling in a hierarchical cognitive system. According to the model, mind wandering reduces cognitive processing of incoming information at a very early perceptual level and across multiple sensory modalities. The consequences of such low-level decoupling then “cascade downward through the cognitive system” (Smallwood, Fishman, et al., 2007, p. 233) and cause decoupling at higher levels. Based on this mechanism, the model parsimoniously explains why decoupling impairs performance in “as wide a range of tasks as perception, encoding and reading” (Smallwood, 2011b, p. 68).

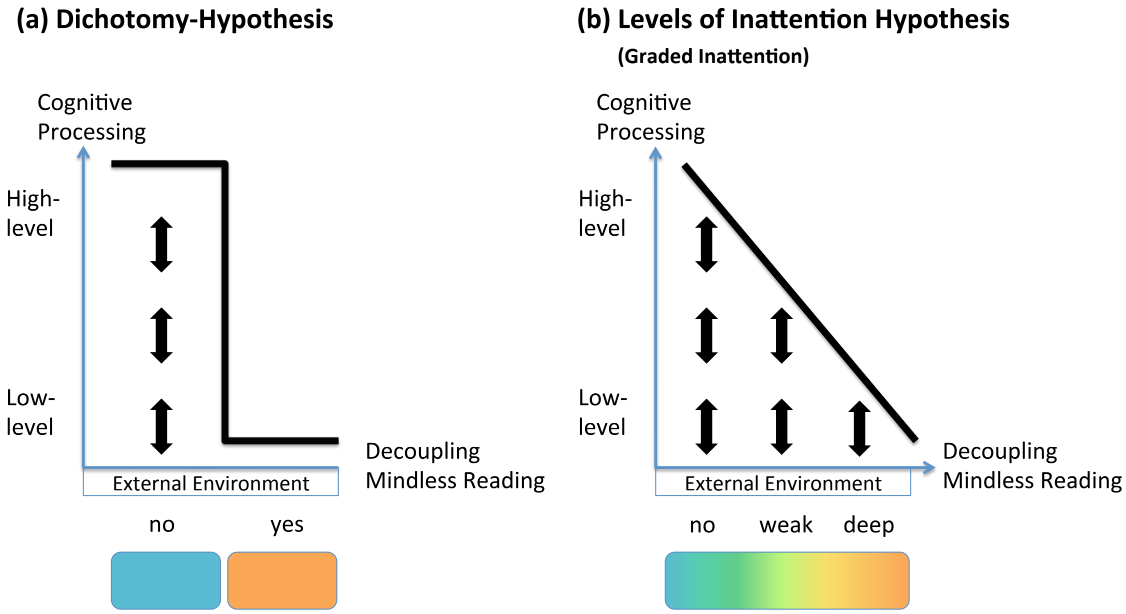


**Figure 4-1**

Schematic illustration of two theoretical hypotheses about how different levels of cognitive processing are decoupled from the external environment during inattention. It is illustrated how high-level and low-level cognitive processing is coupled (below the black line, black arrows) or decoupled (above the black line) from the external environment. (a) The dichotomy-hypothesis proposes that attentional decoupling occurs in an all-or-none fashion, where cognitive processing is either coupled (left, blue) or decoupled (right, red) from external input. (b) The levels of inattention hypothesis proposes graded degrees of decoupling, including fully coupled (left, blue), weakly decoupled (middle, green/yellow/orange), and deeply decoupled (right, red) processing.

Stimulus-independent thought and stimulus-dependent thought are usually treated as a dichotomy (Smallwood, et al., 2011), and this view has dominated previous research (e.g., Christoff, 2012; Fox, et al., 2005; Killingsworth & Gilbert, 2010; Levinson, et al., 2012; McVay & Kane, 2012b; Reichle, et al., 2010; Smallwood, 2010b). Here, we investigate attentional decoupling and whether it is of a dichotomous or a hierarchically graded nature. First, the dichotomy-hypothesis proposes that different levels of cognitive processing are decoupled from external input in an all-or-none fashion (see Figure 4-1a): during task focus all hierarchical levels of cognitive processing are coupled to the external environment, but when the mind wanders this coupling breaks down at all levels. As a potential mechanism, attentional decoupling may always attenuate early perceptual processing stages across modalities (reflecting early attentional selection, Broadbent, 1958) and the consequences of this low-level decoupling may cascade into the system to impair analysis at higher levels (Smallwood, 2011b; Smallwood, Fishman, et al., 2007). For the phenomenon of mindless reading, the dichotomy-hypothesis predicts that impaired visual representations of the text prevent a successful analysis at the lexical, syntactic, semantic, and the discourse level.

As an extension of the dichotomous view, we propose the levels of inattention hypothesis (Figure 4-1b): We postulate that cognitive processing of external input does not always fail at an early perceptual level, but fails at different hierarchical levels, resulting in different graded degrees of weak and deep attentional decoupling. During occasional episodes of deep decoupling, cognitive processing of external input ceases at an early perceptual level (early attentional selection), and the consequences of this low-level decoupling cascade into the system to cause decoupling at higher levels (Smallwood, 2011b; Smallwood, Fishman, et al., 2007). As a new contribution, we postulate states of weak decoupling, where high-level cognitive processing is decoupled from the external environment (i.e., late attentional selection, Deutsch & Deutsch, 1963) but low-level processing is fully intact. Lastly, during states of full attentional coupling external information is processed at all levels. Combining the levels of inattention hypothesis with the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) predicts that decoupling at different levels is hierarchical because reduced cognitive processing at one specific level will cause decoupling at higher levels in the hierarchy.

Previous studies on attentional decoupling have typically focused on dichotomous aspects of the decoupling process: many studies investigated decoupling in the sustained attention to response task (SART) via failures to inhibit the response to rare target stimuli (Manly, et al., 1999; Robertson, et al., 1997; Smallwood, et al., 2004; Smallwood, et al., 2006), and/or via dichotomous measures of SIT (Kam, et al., 2011; Reichle, et al., 2010; Smallwood, Beach, et al., 2008). However, some previous studies suggest that the underlying phenomenon may not be dichotomous. A recent model (Cheyne, et al., 2009) has proposed three discrete states of task engagement/disengagement – occurrent task inattention (to dynamically changing "moment-to-moment stimulus meaning"), generic task inattention (to the "general task environment"), and response disengagement (i.e., inattention to "motor behavior") – and found support for these states in analyses of the SART (see also Cheyne, et al., 2011; Seli, Cheyne, & Smilek, 2012). Moreover, based on principle component analyses, Smallwood and colleagues (Smallwood, 2010a; Smallwood, McSpadden, Luus, et al., 2008) (also see McVay & Kane, 2012a) showed that performance errors were preceded by a gradual shift in response times from slow to fast responses, which may lend support to a graded nature of decoupling.

With the present work we test theoretical hypotheses by studying attentional decoupling during reading. Mind wandering has long been thought to be elusive to vigorous scientific investigation because it is difficult to induce and control in the laboratory. For example, mindless reading was considered to "be very difficult to study experimentally" (Rayner & Fischer, 1996, p. 746). Previous research has approximated mindless reading via scanning of z-strings, where each letter in a text is replaced by the letter 'z' and subjects are asked to move their eyes across the z-strings 'as if they were reading' (Nuthmann & Engbert, 2009; Nuthmann, et al., 2007; Rayner & Fischer, 1996; Vitu, et al., 1995). Other studies have approached mindless reading by studying old readers (Christiane Wotschack & Kliegl, 2011) or via reading of randomly shuffled text, where the order of words in a text is randomly shuffled and subjects have the task to read the meaningless word lists (Schad & Engbert, 2012; Schad, et al., 2010)[1]. To catch spontaneous episodes of mind wandering during normal reading, research has focused on thought sampling methods, where subjects are asked to report about their inner experiences of mind wandering (Giambra, 1995; Reichle, et al., 2010; Schooler, et al., 2004; Smallwood & Schooler, 2006).

Both approaches have their limitations. Approximating mindless reading via paradigms like 'z'-string scanning may not capture the phenomenon of mind wandering. Studying mind wandering using the thought sampling method is subject to the limitations associated with subjective self-report on cognitive processes, i.e., introspection (Nisbett & Wilson, 1977), and continuously monitoring one's conscious thought may change behavior. As a complementary approach, indicators for mind wandering have been derived from behavioral measures of attentional decoupling. Previous behavioral approaches include failures to inhibit the response in the sustained

---

[1] Based on the levels of inattention hypothesis, we suggest that z-string scanning (Vitu, et al., 1995) may be regarded as approximating a state of deep mindless reading, where no language processing is present. Shuffled text reading (i.e., reading random word lists), to the contrary, may approximate weak mindless reading, where processing of higher-level text meaning is absent, but some processing of individual words is intact (Schad & Engbert, 2012; Schad, et al., 2010).

attention to response task (SART: Bellgrove, Hawi, Gill, & Robertson, 2006; Christoff, et al., 2009; Johnson et al., 2007; Manly, et al., 1999; Molenberghs et al., 2009; Robertson, et al., 1997; Seli, Cheyne, Barton, & Smilek, 2012; Smallwood, et al., 2006; Smilek, Carriere, & Cheyne, 2010a), and reaction times in a word-by-word reading paradigm (Franklin, et al., 2011). However, there is currently a lack of objective measures that catch mind wandering in natural and complex tasks like normal reading.

### 4.1.1 Present experiment

To fill this gap in current experimental approaches, we introduce the sustained attention to stimulus task (SAST), which is based on psychophysics of error detection in a reading experiment. Our analyses use recordings of eye movements to derive measures for attentional decoupling. Methodologically, a corpus of normal text was manipulated by inserting specific meaningless error sentences containing different kinds of errors. A control condition was added where error sentences contained no error. Readers were asked to indicate whenever they noticed that the text turned meaningless. Mindless reading was operationally defined as (a) overlooking an error passage (single-trial level), and (b) low sensitivity for errors (aggregated level). In this new paradigm, we utilize classical psychophysical methods from signal detection theory (Wickens, 2002) to distinguish between sensitivity for errors (i.e., the propensity for mindless reading) and a general tendency of readers to respond in a certain fashion. The approach does not require instructions about mind wandering, and may be less intrusive and more objective than self-report measures used in previous studies. However, we cannot exclude the possibility that instructions about errors may affect reading behavior as readers may pay increased attention to detect the errors in the text. To counteract such effects we (a) optimized the experimental setting to increase the chance of observing mindless reading in the eye tracker (see Methods section for details) and (b) included high-level errors such that text comprehension was necessary to detect the errors and relatively normal reading can be expected.

To avoid detecting mindlessness when readers were in fact paying attention to the task several measures were taken: first, very easy texts were selected to ensure that readers would have no comprehension difficulties (cf. Smallwood, Fishman, et al., 2007). Second, readers received instructions and examples explaining the different error types. Third, readers were encouraged to respond also when unsure about the presence of an error.

**Table 4-1**

Types of errors used for error sentences.

| Type of Error | Construction / Description | Example sentence |
|---|---|---|
| **Control** | • no error, meaningful text | (1) The wall was made from big worked stones.<br>(2) On all birthdays, he congratulates his classmates and the teacher. |
| **Lexical** | • one word is replaced by a morphologically & phonologically legal pseudo-word<br>• detectable via lexical, but not via pure orthographic or phonological processing<br>• does NOT resemble any real word that could fit into the text | (1) The wall was begrothed from big worked stones. |
| **Syntactic** | • one word in the sentence is moved to a different location, causing a syntactic error | (1) The wall worked was made from big stones. |
| **Semantic** | • statements in the sentence contradict world knowledge | (2) He always thinks of buying new hamsters for the bathroom at school. |
| **Discourse** | • neighboring sentences are inconsistent with each other (e.g., direct contradictions of statements)<br>• each single sentence is correct (no lexical, syntactic, or semantic error) | (2) <u>He welcomes the guests</u> on behalf of the class. He congratulates his classmates on their birthdays <u>but never welcomes the guests</u>. |
| **Gibberish text** | • changed order of nouns or pronouns within a sentence<br>• correct syntax<br>• (Smallwood, Fishman, et al., 2007) | (2) On all classmates, he congratulates his birthdays and the teacher. |

To generate measures for low-level and high-level decoupling, we constructed errors at different levels of the text (Table 4-1). (*i*) We replaced one word in an error sentence by a pseudo-word, causing a *lexical* error. If low-level lexical processing is decoupled from the text, then readers cannot detect lexical errors. Second, (*ii*) we included *syntactic* errors as a measure for syntactic processing. (*iii*) Statements that are incompatible with the readers' world knowledge were included to construct *semantic* errors. If medium-level sentence meaning is not processed, then readers cannot detect semantic errors. (*iv*) We included sentences that clearly contradicted their context to construct *discourse* errors. These can be detected only when readers integrate the meanings from neighboring sentences into a single representation, and thus tested for high-level discourse processing. Lastly (*v*), we reordered nouns and pronouns from the meaningful control sentences to construct *gibberish text* for comparability with previous research (Smallwood, Fishman, et al., 2007). Readers may automatically construct meaning from meaningless gibberish text without noticing by reordering words (Ferreira, Bailey, & Ferraro, 2002), and we therefore expect gibberish text to reflect high-level construction or repair processes. All errors were constructed to (a) lack an overall meaning and (b) show no similarities to any possible meaningful sentence. For example, pseudo-words were not implemented as spelling-errors, but constructed to have no similarities to any existing word. This was done to ensure that overlooking errors would indicate mind wandering and would not occur because readers constructed meaning from meaningless text.

Based on dichotomous versus graded conceptions of decoupling, we derived predictions for readers' sensitivity for different error types. The levels of inattention hypothesis predicts that sensitivity should differ between error types: readers should be very sensitive to low-level errors (e.g., lexical errors) as these should be overlooked only during deep decoupling. To the contrary, readers should be less sensitive to errors assessing high-level text processing (discourse errors, gibberish text) because already weak decoupling prevents detection of these errors. Based on the dichotomy-hypothesis, attentional decoupling should either cause no differences in sensitivity between error types, or any differences in sensitivity should be due to different durations (rather than depths) of mind wandering.

We recorded eye movements in the SAST to derive measures for different levels of cognitive text processing during reading (Rayner, 1998). Readers usually look longer at phrase- and sentence-final words compared to non-final words, and this wrap-up effect is related to the high-level process of integrating words and constructing a text meaning (Just & Carpenter, 1980; Warren, et al., 2009). Moreover, readers look longer at low-frequency compared to high-frequency words (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner, 1998; Rayner & Duffy, 1986) reflecting low-level lexical processing. Reichle et al. (2010) were the first to study mind wandering during reading using eye tracking. They argued that lexical and linguistic influences on eye movements are reduced during mindless reading, indicating a decoupling of cognitive processing from the text (see also Rabovsky, Alvarez, Hohlfeld, & Sommer, 2008; Rayner & Fischer, 1996; Rayner & Raney, 1996; Schad & Engbert, 2012; Schad, et al., 2010). The levels of inattention hypothesis predicts that during states of weak decoupling high-level (wrap-up) processes should be reduced, but low-level (lexical) influences should be intact. During deep decoupling, however, high- and low-level processes should be reduced.

*Predicting mindless reading from eye movements*. A major current challenge and chance for mind wandering research is to identify objective and reliable online-markers that allow detecting episodes of mind wandering (including their onset and offset) without relying on subjective self-reports or interfering with task performance (Franklin, et al., 2011). Previous findings (Reichle, et al., 2010; Smilek, et al., 2010b;

Uzzaman & Joordens, 2011) suggest that eye movements may be ideally suited for this purpose because they (a) provide a good measure of moment-to-moment cognitive processing and attention (Rayner, 1998, 2009), (b) occur with high frequency in virtually all tasks (Liversedge, Gilchrist, & Everling, 2011), and (c) are relatively easy to record and analyze.

At the same time, it may be difficult to predict mind wandering from eye movements. First, finding that mindless reading predicts measures of eye movements [i.e., a high probability *P*(eye | mindless)] is not the same as finding that eye movements predict mindless reading [i.e., a high probability *P*(mindless | eye)], and these two probabilities can be very different[2]. Here, we use a Bayesian analysis to determine the *posterior* probability, *P*(mindless | eye), that a reader is currently in a state of mindless reading given a recorded eye movement. Second, when observing mean differences between mindful and mindless reading at the level of groups (averaged over participants, trials, and/or individual eye movements) it remains unclear whether mindlessness can be inferred from the eyes at the level of individual eye movements or trials. Such predictions might be difficult to derive, because eye-movement measures exhibit considerable variance (Kliegl, et al., 2006; Rayner, 1998). Notably, reading fixations crucially depend on the words and sentences being read. However, the design of the present study allows investigating mindless and mindful reading on exactly the same text material, including specific target words.

## 4.2 Materials and methods

### 4.2.1 Participants and materials

Thirty German high school students, aged between 17 and 20 years, were paid 45 € each to participate in the study. Informed consent was obtained from all participants. All participants had normal or corrected to normal vision. Participants read 50 stories, taken from elementary school textbooks and slightly modified for the experiment (henceforth *Potsdam Mindless Reading Corpus*, PMC). The text corpus comprised about 17,500 words distributed across 216 pages of text.

---

[2] For example, the probability for professors to have a high-school degree, *P*( high-school | professor ), likely approaches one, while the probability for high-school graduates to become a professor, *P*( professor | high-school ), is much lower. Treating these probabilities as equal reflects the fallacy of the transposed conditional Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology, 100*(3), 426-432..

### 4.2.2 Apparatus

In an attempt to create a situation where participants were likely to encounter episodes of mindless reading, readers were seated in a comfortable, laid-back easy chair where they could rest their head on a headrest and their legs on a footstool. An arm mount was positioned for the recording of eye movements, holding an EyeLink 1000 eye tracker (SR Research) in the remote setup and a 17-inch flat panel LCD screen. The monitor was positioned slightly above the eye level of the reader and then tilted downward, so that a reader's line of gaze would be perpendicular to the vertical plane of the monitor. The viewing angle of the monitor and monitor tilt were occasionally adjusted to achieve maximum comfort for each reader. Viewing distance was approximately 50 cm, at which each letter of text horizontally subtended approximately .37 degrees of visual angle. The eye tracker sampled left eye position at a rate of 500 Hz. Readers could move their head freely, but for the most part chose to rest it on the head rest of the chair. The EyeLink remote system tracked possible head movements and corrected measured eye position for these movements. The stories were presented in black against a brown-grey background. A rectangular dark brown-grey frame was drawn around the text to create the impression of reading from a sheet of paper. Monitor brightness was reduced to the minimum.

### 4.2.3 Design and errors in the text

Two experiments were conducted in succession. Each experiment required participants to read 25 stories. Sixty-two error sentences were defined at quasi-random locations in the PMC, with each story containing one or two error sentences. For each error sentence, several different versions of similar length were constructed. They contained six different kinds of linguistic errors (including an error-free control condition) and were designed to probe for five different levels of mindless reading (Table 4-1). Which error type was presented at a given location in the text was varied between readers. This was done within experiments 1 and 2 separately. Importantly, the design allowed us to test the effects of different levels of mindless reading on the same text material. Across both experiments, errors were presented in 48 out of 62 target locations per participant, resulting in a relatively low average presentation rate of one error per 354 words (equivalent to 4.5 text pages). In the remaining 14 target locations meaningful sentences were presented as a control condition.

### 4.2.4 Procedure

The experiment was advertised as “relaxed reading”. Upon arrival, readers were instructed to relax on the chair and to find a comfortable position to sit in. Readers’ task was to read the stories for comprehension, and it was emphasized that they should read in a relaxed manner. Participants were told that the text would sometimes be more or less incoherent. They were informed about the various kinds of errors that might occur and this was illustrated by example sentences. Participants were instructed to press the space bar on the computer keyboard whenever they noticed an error in the text. At the beginning of the experiment, participants read three pages of text for practice, each containing one error. They then read the 50 stories of Exp. 1 and Exp. 2. Between experiments, participants were allowed to take a short break where they could stand up and stretch. Within a given experiment, story order was randomized for each subject. Readers could move forwards and backwards in the text by pressing arrow keys on the keyboard. We allowed readers to move backwards in the text to ease transitions into a relaxed reading mode. Presentation of each page of text was preceded by a fixation check to ensure calibration quality. Successful error detection was defined as pressing the space bar on the keyboard after reading an error sentence and before moving on to the next text page. After reading all texts, participants completed two memory tests, the details of which are not reported here.

### 4.2.5 Data processing and analysis

The cognitive parsing algorithm of the SR Research EyeLink software was used to determine the positions and durations of readers’ individual fixations. Fixations were then assigned to pages and lines of text, individual words, and letters (Supplementary Information). (Generalized) Linear mixed effects models ([G]LMMs, Baayen, et al., 2008; Kliegl, et al., 2010; Pinheiro & Bates, 2000) were used to test differences between mindless and mindful reading, and the control condition (Supplementary Information). (G)LMMs can be viewed as a generalization of linear regression and allow estimation of random effects (i.e., effects of factor levels that are randomly sampled from a population; here: participants, words, and text pages) in addition to fixed effects [i.e., effects that are repeatable across experiments and can be either discrete (e.g., experiment number) or continuous (e.g., word frequency)]. For large sample sizes the t-statistic effectively corresponds to the z-statistic. Therefore, for the LMMs (two-tailed testing), we took absolute $t$ values larger than 1.645 to indicate marginal significant effects ($p < .10$), values larger than 1.96 to indicate significant effects ($p < .05$), and $t$ values larger than 2.576 ($p < .01$) or 3.291 ($p < .001$) to indicate highly significant effects (cf. Kliegl, Ping, Dambacher, Yan, & Zhou, 2011).

### 4.2.6 Data selection

For analyses of eye movements, errors with an overall detection rate of less than 30% were excluded (12.2%). Eye movements from false alarm trials were discarded. To unconfound mindless reading and skimming we excluded trials in which less than 50% of the words in the error sentence were fixated (4.0% of trials; Supplementary Information), leaving a total of 1,793 trials for analyses. Under the assumption that readers were already on/off task on the words prior to the error (Reichle, et al., 2010; Smallwood, Fishman, et al., 2007), we first analyzed eye movements in an interval of 14 words preceding each error sentence. Next, we generalized the analyses to different interval sizes using the same selection criteria. Only words on the same page of text as the error sentence and only eye movements made during the first viewing of each page of text were analyzed. Also, in each trial we only analyzed eye movements that were made prior to fixating any of the words from the error sentence so that the analyses did not include data from the error sentence, nor data that was collected after subjects had read the error sentence. (For the measure of the "number of reading passes" we made an exception to this selection criterion and also included fixations made after reading the error sentence.) For the 14-words interval, the selection resulted in a total of 24,528 fixations on 20,498 words and 19,313 first-pass fixations on 15,539 words. (Firstpass fixations include all fixations on a word before the reader makes a regression back to this word or previous words in the text.) To select valid word-based fixation time measures like gaze duration (the cumulative duration of all first-pass fixations per word), standard criteria used in reading research were applied (e.g., removing calibration problems, blinks, irregular fixation behavior [lines with less than 50% fixated words], first and last fixation per line, long and short fixations and saccades; see Supplementary Information). This procedure resulted in valid gaze durations for 9,435 words, including 11,106 first-pass fixations. Overall, there were slightly more words with invalid first-pass fixations during mindless reading (40.9%) than during the control condition (38.9%), and mindful reading (38.7%), mainly because there were more lines with irregular fixation behavior and more calibration problems during mindless reading (Supplementary Information).

## 4.3 Results

It took readers an average of 2 hours and 45 minutes (range: 1:40 h to 4:20 h) to read all texts.

### 4.3.1 Error detection

Readers overlooked 39% of the errors in Exp. 1, and 44% of the errors in Exp. 2. False alarm rate, reflecting responses in the control condition without errors, was 7% in Exp. 1 and 3% in Exp. 2. We used signal detection theory (Wickens, 2002) to assess readers' ability to detect errors (i.e., sensitivity for errors, *d'*, reflecting the propensity for mindful reading) and response bias (*c*). When studying mindless reading we inevitably observe highly imbalanced data. These are adequately handled by (generalized) linear mixed effects models [(G)LMMs], which we used to implement the signal detection analyses (Wright, et al., 2009, Supplementary Information). Experiments did not significantly differ in sensitivity (Exp. 1: $d' = 1.90$; Exp. 2: $d' = 2.15$; $p > .10$) and there was a marginal effect in response bias (Exp. 1: $c = -1.55$; Exp. 2: $c = -1.98$, $z = 1.68$, $p < .10$) reflecting slightly fewer responses in Exp. 2. Sensitivity for errors decreased over the course of Exp. 2 [$\Delta d'$(per page of text) = -0.014, $z = -1.94$, $p = .052$], but not across Exp. 1 ($p = .50$; slope-difference between experiments: $\Delta d' = 0.017$, $z = 2.01$, $p < .05$).

Figure 4-2 depicts how sensitivity differed between error types (for both experiments: $p < .001$). Planned contrasts revealed that these differences followed the predictions: in Exp. 1, readers were most sensitive to (*i*) semantic errors, followed by (*ii*) discourse errors (difference to semantic errors: $\Delta d' = -0.80$, $z = 5.73$, $p < .001$), and (*iii*) gibberish text (difference to discourse errors: $\Delta d' = -0.31$, $z = 2.40$, $p < .05$). In Exp. 2, readers were most sensitive to (*i*) lexical errors, followed by (*ii*) syntactic errors ($\Delta d' = -0.46$, $z = 3.02$, $p < .01$), (*iii*) semantic errors ($\Delta d' = -0.04$, $z = 0.26$, $p = .79$), and (*iv*) gibberish text ($\Delta d' = -0.54$, $z = 3.63$, $p < .001$). Thus, readers more easily noticed low-level errors, and were less sensitive to high-level errors. This finding is compatible with the idea that different levels of attentional decoupling led to overlooking of different kinds of errors.

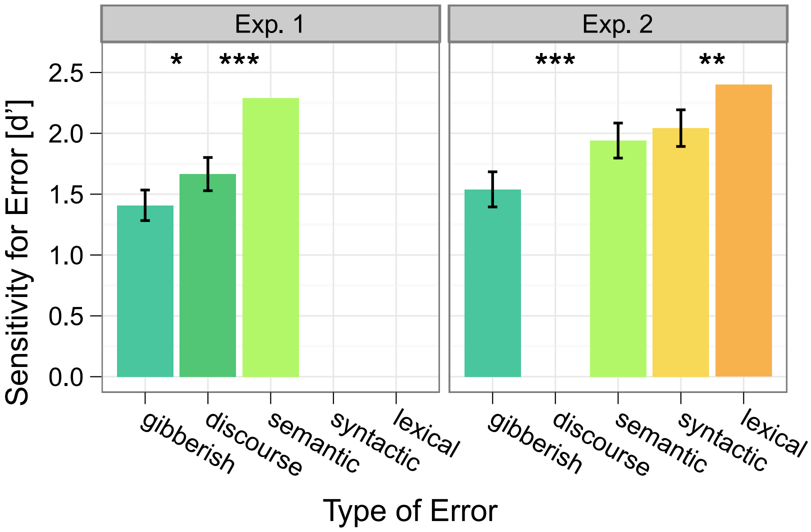


**Figure 4-2**

Sensitivities (d') from a mixed effects signal detection analysis for different types of errors in Exp. 1 and 2. Conditions *lexical* and *syntactic* were not tested in Exp. 1 to discourage strategies other than understanding the text. *Discourse* was not tested in Exp. 2 to focus on levels of deep mindless reading. Conditions are color-coded, ranging from high-level errors (dark green; left) testing weak mindless reading to low-level errors (orange; right) testing deep mindless reading. Error bars are SEM from a GLMM testing (sliding) differences in sensitivity between neighboring error types (Venables & Ripley, 2002). (**p* < .05, ***p* < .01, ****p* < .001, —*p* > .10)

### 4.3.2 Analyses of eye-movements

We hypothesized that overlooking an error indicates an episode of mindless reading. Assuming that most of the time readers were already off task on the words before the error (Reichle, et al., 2010; Smallwood, Fishman, et al., 2007), we first analyzed eye movements made in an interval of 14 words preceding each error sentence. To test the generality of the findings, follow-up analyses considered different interval lengths. Unless otherwise noted, data from different types of errors and from the two experiments were pooled for analyses.

Global analyses focused on common measures of eye movements used in reading research (Rayner, 1998). Nine word-based measures of fixation durations and saccade probabilities were computed (Supplementary Information). For the analyses we used (G)LMMs to investigate how fixed effects like mindless reading affect measures of eye fixations, and determined regression coefficients, *b*, to estimate the size of these influences. Unless otherwise noted we used unstandardized regression coefficients, where *b* estimates the change in the dependent variable given a one-unit change in the independent variable. Out of the nine measures, only one measure significantly differed between mindless and mindful reading: readers read words with fewer passes during mindless reading as compared to mindful reading ($b = 0.10$; $t = 5.0$, $p < .001$). Differences in any of the other eight variables were not significant (Supplementary Information).

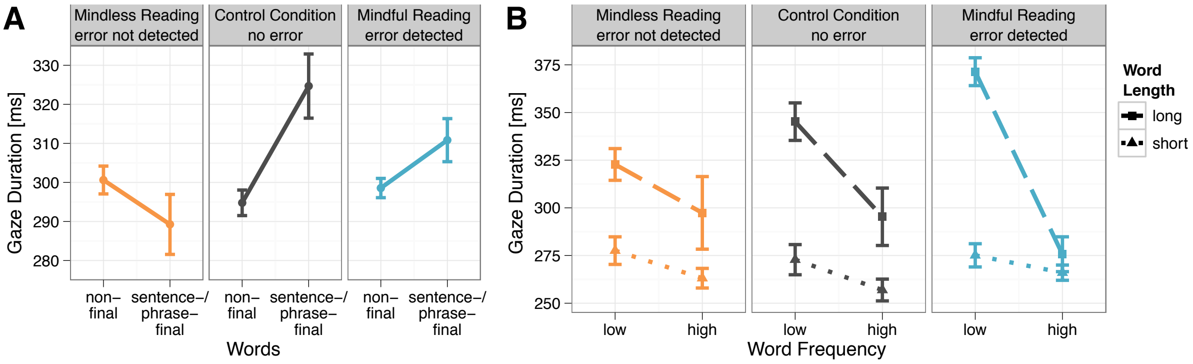


**Figure 4-3**

Influences of high-level and low-level linguistic variables on gaze durations in three reading conditions: mindless reading, a control condition, and mindful reading. Gaze durations were measured on 14 words prior to the error sentence and were residualized for random effects from the LMM; error bars are SEM. **(*A*)** High-level wrap-up effect: gaze durations on *sentence- or phrase-final* words versus *non-final* words. Data is from both experiments. **(*B*)** Low-level lexical effect: modulation of gaze durations due to word length, word frequency, and their interaction in Exp. 2. Short words are six or fewer letters long; word frequencies were median split.

Next, we performed local analyses to test whether the influence of lexical and linguistic variables on gaze durations is reduced during mindless reading as compared to mindful reading or the control condition. As can be seen in Figure 4-3*A* sentence- and clause-final words were fixated longer than other words, replicating the wrap-up effect found in many reading studies (Just & Carpenter, 1980; Warren, et al., 2009). The average wrap-up effect across all reading conditions (mindful, mindless, and control) was not present in the LMM ($t = -1.47$, $p > .10$; Supplementary Information) after statistically controlling for word length, word frequency, and random between-word variance. Notably, the wrap-up effect was strongly reduced in the mindless reading condition (Figure 4-3*A*), and this difference was significant (wrap-up in mindless versus mindful condition: $b = -18.0$, $t = -1.90$, $p < .10$; control versus mindless: $b = 23.7$, $t = 2.15$, *p* < .05), and did not differ between experiments (|*ts*| < 0.8, *ps* > .10).[3]

Prior to overlooking errors in Exp. 2, the effects of lexical variables on gaze durations were reduced. As can be seen in Figure 4-3*B*, readers overall looked longer at long words than at short words. They also looked longer at words of low frequency than at words of high frequency, and the effect of word frequency was stronger for long than for short words (all |*ts*| >= 2.9, *ps* < .01), replicating key findings in reading research (Kliegl, et al., 2006; Rayner, 1998). However, these effects were considerably reduced during episodes of mindless reading. Word length (1/wl) had a weaker effect on gaze durations during mindless reading than during mindful reading (*b* = 155, *t* = 1.93, *p* < .10) or the control condition (*b* = -195, *t* = -1.93, *p* < .10). The main effect of word frequency (log10 freq) did not significantly differ between mindless reading and mindful reading or the control condition (|*ts*| <= 1.4, *ps* > .10). The word frequency effect, however, was hardly modulated by word length during states of mindless reading (Figure 4-3B, left panel). Statistically, this modulation was much weaker than during mindful reading (*b* = -175, *t* = -4.6, *p* < .001) or the control condition (*b* = 92, *t* = 1.95, *p* < .10). As is visible in Figure 4-3*B*, for long words the frequency effect was strongly reduced during mindless reading (*b* = 15, *t* = 2.8, *p* < .01; for post-hoc tests see Supplementary Information). For short words, the frequency effect was not significant during mindful reading, but marginally significant during mindless reading, and the slope-difference was significant. It may be that lexical processing of short words is more automatic and does not require the kind of higher-level attention measured in our paradigm. In summary, lexical processing effects were reduced before errors were overlooked in Exp. 2, indicating episodes of deep mindless reading.

Next, we (a) extended our local analyses presented in Figure 4-3 to intervals ranging from 10 to 20 words prior to the error and (b) performed more explicit tests for differences between experiments. When participants were in the initial phase of the reading task in Exp. 1 we expected that during mindless reading cognitive processing might be weakly decoupled from the text. Accordingly, high-level influences on gaze durations should be reduced but low-level influences may be intact. In contrast, after having spent much time in the lab reading boring texts readers may pay less attention to

---

[3] Post-hoc tests showed that the difference in the wrap-up effect between the mindful condition and the control did not exceed the level of chance (*t* = 0.63, *p* >.10).

the reading task in Exp. 2, and cognitive processing may be deeply decoupled during mindless reading. Hence, text processing should fail at all levels of processing and both high-level as well as low-level influences should be decoupled.

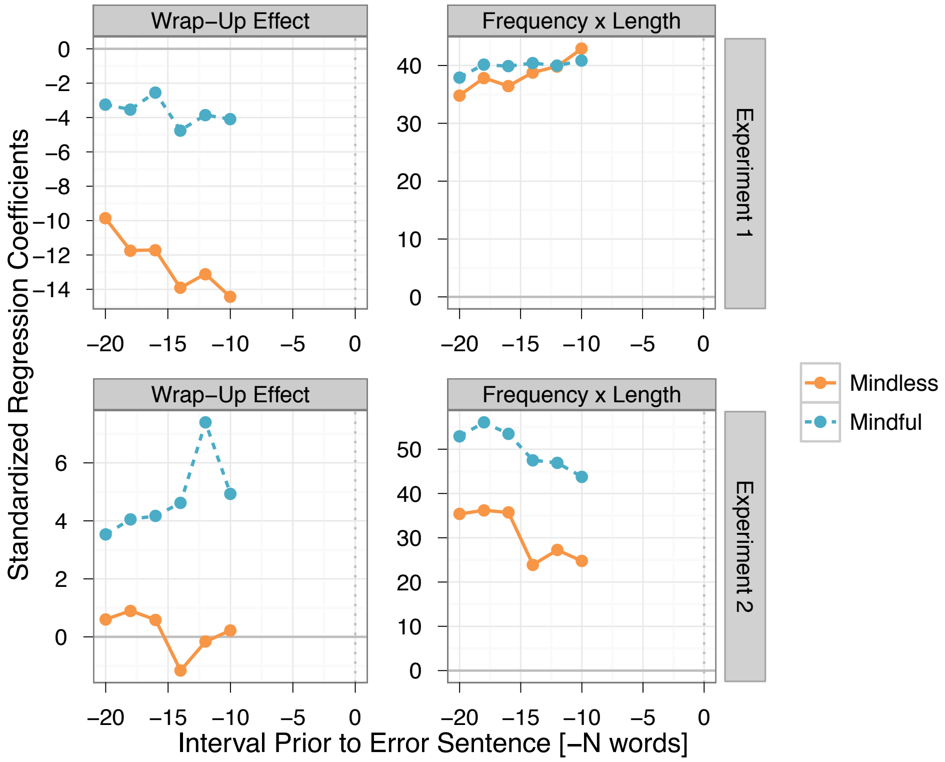


**Figure 4-4**

Effects of high- and low-level linguistic variables on gaze durations during mindless (orange, solid line) and mindful (blue, dashed line) reading in Exp. 1 (upper panels) and Exp. 2 (lower panels) for different intervals (*N* of words) prior to the error sentence. Graphs depict standardized regression coefficients from LMM analyses. For each interval, a separate LMM analysis was conducted and tested whether a given effect differed between mindless and mindful reading. For the high-level wrap-up effect (left panels), positive regression coefficients represent the standard wrap-up effect of longer fixations on final compared to non-final words. For the low-level interaction between word frequency and length (right panels), positive coefficients indicate a stronger frequency effect for long words than for short words.

Figure 4-4 displays standardized regression coefficients representing the relative influences of high-level wrap-up and low-level lexical (word frequency × length interaction) variables on gaze durations during mindless and mindful reading. The results show that wrap-up effects were reduced during mindless reading (Figure 4-4, left panel) for all intervals [marginal (*t*s > 1.7, *p*s < .10) to significant (*t*s < 2.1, *p*s > .01) reduction; 20-words: *t* = 1.54, *p* > .10] and this effect did not significantly differ between experiments (|*t*s| < 0.95, *p*s > .10)[4]. In our previous analyses (Figure 4-3*B*) we had found that in Exp. 2 the word frequency effect was reduced during mindless reading for long words (but not for short words). Figure 4-4 (lower right panel) shows that this effect was highly reliable for all intervals (*t*s > 2.80, *p*s < .01; mindful versus mindless reading). However, it was absent in Exp. 1 (|*t*s| < 0.63, *p*s > .10), and the difference between experiments was significant for all intervals (*t*s < -1.97, *p*s < .05). Taken together, for Exp. 1 we observed a dissociation between reduced high-level wrap-up effects and intact low-level lexical effects (Figure 4-4, upper panels), which provides support for our expectation that cognitive processing was weakly decoupled when mindless reading occurred in the initial part of the study. For Exp. 2, however, the results indicate states of deep decoupling as both high-level wrap-up effects as well as low-level lexical influences on gaze durations were reduced during mindless reading (Figure 4-4, lower panel).

A central prediction that emerges from the proposed levels of inattention hypothesis is that overlooking different kinds of errors reflects different levels of attentional decoupling. To further test this prediction we analyzed eye movements for different error types. For the analyses we defined three broad categories of error types: (a) high-level errors (gibberish text and discourse errors), (b) medium-level errors (semantic and syntactic errors), and (c) low-level errors (lexical errors). This aggregation helped to reduce complexity and to improve the stability and reliability of the LMM analyses. We then generated a statistical measure for attentional decoupling: for the high-level wrap-up and the low-level lexical variable, we determined the influence of this variable on gaze durations (by computing the standardized regression coefficient in an LMM). Next, we determined how this influence differs between mindless and mindful reading. The resulting difference-value (coded as an interaction between lexical/linguistic influences and mindless reading) represents a direct statistical measure for attentional decoupling: Negative difference-values indicate that linguistic influences on eye movements are reduced when errors are overlooked. Based on the levels of inattention hypothesis we predict that for low-level errors decoupling should be observed for low-level (lexical) and for high-level (wrap-up) influences,

---

[4] Note that the coefficients for the wrap-up effect during mindless reading were negative for Exp. 1. The reason for this is unclear. The difference between mindful and mindless reading, however, was as expected.

whereas for high-level errors high-level wrap-up effects should be reduced, but low-level lexical effects should be relatively less affected.

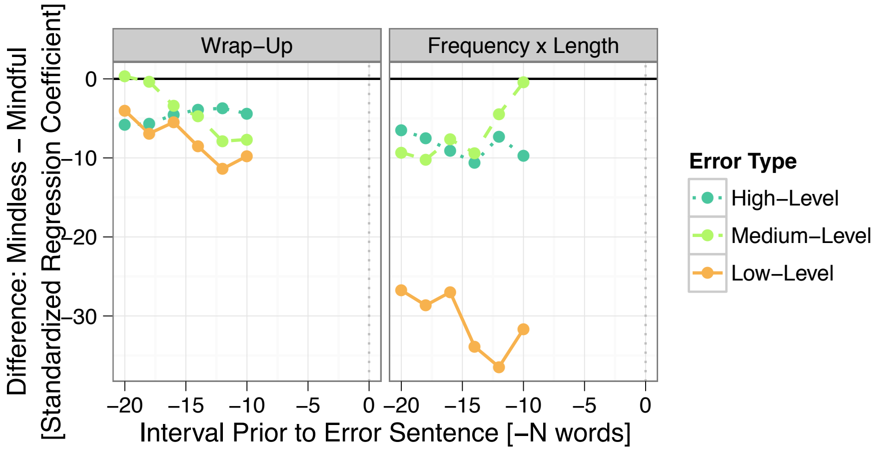


**Figure 4-5**

Differential effects (mindless versus mindful reading) of high-level and low-level variables on gaze durations for different categories of errors. The graphs show how standardized regression coefficients representing the influences of high-level wrap-up (left panel) or low-level lexical (right panel; word frequency × length interaction) variables differ between trials where errors were overlooked (mindless reading) versus detected (mindful reading). Negative difference-values indicate that the influence of linguistic variables on gaze durations is reduced during mindless reading, reflecting attentional decoupling. High-level errors are gibberish text and discourse errors, medium-level errors are semantic and syntactic errors, and low-level errors are lexical errors. Regression coefficients are from LMMs for different intervals of words prior to the error sentence.

For the high-level wrap-up effect (Figure 4-5, left panel) the results suggest that decoupling was present for overlooking of all error types (negative difference-values), and the effect did not significantly differ between error categories (for all intervals: $\chi^2$s(2) < 1.6, *ps* > .47). This finding suggests that when any type of error is overlooked high-level processing is decoupled from the text. (Note that the wrap-up effect is overall smaller in size compared to the word frequency × length interaction.) For the low-level lexical influences (Figure 4-5, right panel) the results show that the influence of word frequency and length was strongly reduced when low-level errors were overlooked, but were only slightly affected when high-level errors were overlooked. The difference between error categories was significant for all intervals larger than 12 words ($\chi^2$s(2) >

5.7, *ps* < .06). These findings support our hypothesis that overlooking different types of errors in the SAST reflects graded levels of attentional decoupling: overlooking low-level errors indicated a state of deep decoupling as both high-level and low-level influences on eye movements were reduced. Overlooking high-level errors, to the contrary, indicated a state of weak decoupling as eye movement markers for high-level integration processes were reduced, but low-level lexical processes were intact.

### 4.3.3 Predicting mindless reading from eye movements

Is it possible to infer from the ongoing eye movements whether readers are currently paying attention to the text? To investigate this question, we selected a subset of the data where we expected the strongest effects of mindless reading. Our results suggest that effects of mindless reading on eye movements are most pronounced for lexical processing of long words (Figure 4-3*B*). For the analyses we used gaze durations on very long target words (>= 10 letters), which were located an average of 13.4 words prior to the error in the text. In addition, we focused our analysis on lexical errors because these should best capture reduced lexical processing (cf. Figure 4-5). As is visible in Figure 4-6*A+B*, distributions of gaze durations on target words considerably differed between deep mindless as opposed to mindful reading, and the direction of the effect was consistent with the general findings reported above. During mindful reading we observed a standard word frequency effect, as gaze durations on low-frequency words were considerably prolonged and gaze durations on high-frequency words were shortened. To the contrary, when lexical errors were overlooked during deep mindless reading target word frequency did not clearly modulate the distribution of gaze durations.

Based on these clear-cut results, we performed a Bayesian analysis to predict mindless reading from the gaze durations readers made on specific target words. Based on the graded nature of decoupling, we estimated the prior probability for mindlessness, *P*(mindless), from the overall rate with which errors were overlooked in Exp. 2. The posterior probability for mindless reading given a certain eye fixation, *P*(mindless | gaze), was determined via Bayesian logistic regression (Gelman, et al., 2008). We found that the posterior probability for mindless reading was low when readers' eyes responded to the lexical difficulty of the target word: mindless reading was least likely when readers made long gaze durations on low frequency target words [*P*(mindless | low freq, gaze >= 500 ms) = .33, for continuous predictions see Figure 4-6*C+D*] or when they made relatively short fixations on high frequency target words [*P*(mindless | high freq, gaze < 500 ms) = .42]. To the contrary, the probability for mindless reading was high when readers' eyes did not respond to the lexical difficulty of the target word: failing to slow down the eyes on difficult low-frequency words predicted mindless reading [*P*(mindless | low freq, gaze < 500 ms) = .60]; likewise, failing to speed up on easy high frequency words was an indicator for an absent mind [*P*(mindless | high freq, gaze >= 500 ms) = .63].

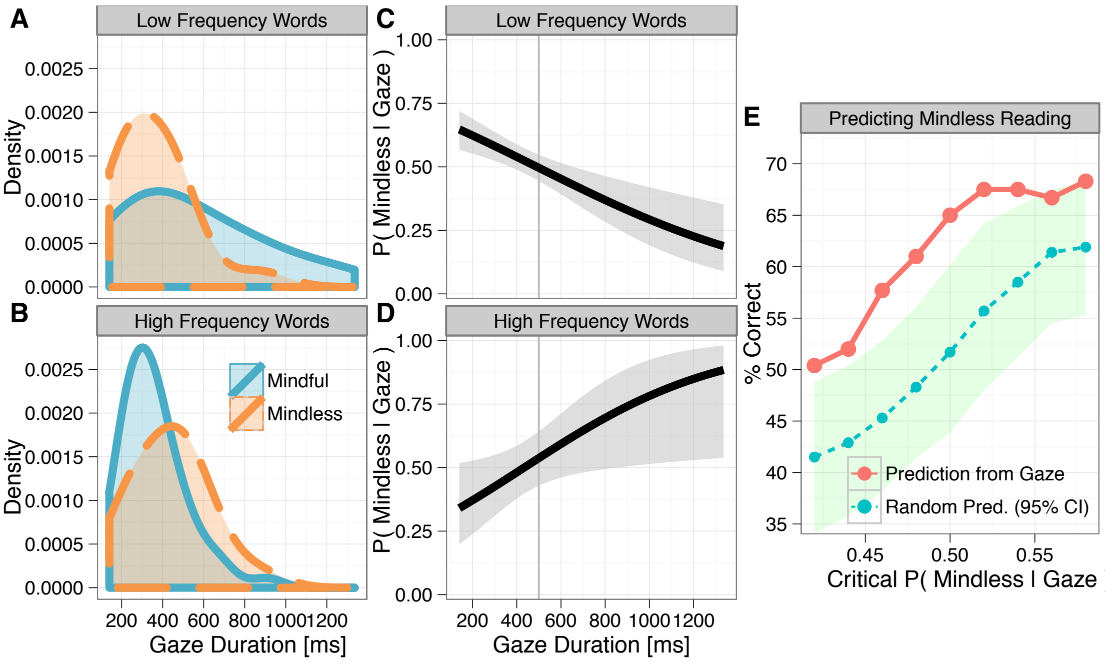


**Figure 4-6**

Predicting states of mindless reading from gaze durations on low-frequency words (panels A and C) and on high-frequency words (panels B and D). Mindless reading is defined as overlooking a lexical error in the upcoming error sentence. Gaze durations are on long words (>= 10 letters) from the 14 words prior to the error sentence. **A and B:** Distributions of gaze durations (estimated densities) during mindful (solid blue line) and mindless (dashed orange line) reading. **C and D:** Posterior probability for mindless reading [point estimates (posterior modes) and SEM] as a function of gaze duration, estimated via Bayesian logistic regression with an informed intercept-prior derived from the overall probability to detect any kind of error in Exp. 2, and Cauchy priors for all other parameters. **A to D:** Log10 word frequencies per million were split at the value zero. **E:** Percent correctly predicted states of mindless reading (red dots, connected by solid line) for different prediction thresholds, P( Mindless | Gaze ), corresponding to different predicted levels of mindless reading. Random predictions (*N* = 1,000) provide a statistical baseline with 95% confidence intervals (dashed green line and light green ribbon).

From the posterior probability for mindless reading (Figure 4-6$C$+$D$) we predicted error detection in the error sentence: We predicted mindless reading when the posterior probability for mindless reading exceeded a critical threshold, and predicted mindful reading when the posterior probability fell below the critical threshold. We used different prediction thresholds, corresponding to different prior expectations for the occurrence of mindless reading, to predict different levels of decoupling. Predictions were successful and significant for a wide range of decision thresholds and reached up to 68.3% correct predictions for deep mindless reading (see Figure 4-6$E$). This finding demonstrates that an individual fixation duration measured on a specific target word in real time can be highly informative about whether a reader's attention is currently focused on the text, or whether it is wandering.

Notably, given the average total reading time of 356 ms and the average target word-error distance of 13.4 words, we predicted overlooking of lexical errors an average of 4.8 seconds before they occurred in the text. This finding suggests that the actual accuracy with which eye movements measure states of mindless reading should be higher than the current estimate of 68.3%. Moreover, predictions were based on information from individual gaze durations readers made on individual target words, and predictions may be further improved by combining information from several words in a trial and from multiple eye movement measures.

## 4.4 Discussion

In the current study, we investigated episodes of mind wandering during reading, where cognitive processing is decoupled from the text as external attention is reduced. Coupled and decoupled processing are often treated as a dichotomy. The central aim of the present work was to introduce the levels of inattention hypothesis, which proposes graded attentional decoupling at hierarchical levels of cognitive processing. To measure levels of attentional decoupling we developed the sustained attention to stimulus task (SAST), a behavioral measure for mindless reading, which is based on readers' sensitivity for errors in the text. We tested predictions from the levels of inattention hypothesis and the cascade model of inattention by performing detailed and reliable analyses of a large corpus of eye-movement data during mindless reading. We found that eye movements were decoupled from low-level and high-level linguistic variables in a hierarchically graded fashion before errors were overlooked. In a Bayesian analysis, we demonstrated that it is possible to use eye movements to predict overlooking of errors five seconds before they occur, and this suggests that eye movements provide an unobtrusive online-indicator for mind wandering. Our findings support the levels of inattention hypothesis and validate the SAST as a behavioral measure of mindless reading.

*Attentional decoupling in the SAST*. As a main result, we found that readers overlooked errors about 40 percent of the time. What factors caused readers to overlook these errors? First, the percentage of overlooked errors is compatible with the estimated amount of time people spend mind wandering in everyday life (Kane, et al., 2007; Killingsworth & Gilbert, 2010), suggesting that we were successful in creating task conditions to investigate mindless reading in the eye-tracking laboratory. Second, mind wandering is known to become more frequent with increasing time on task (Schnitzer & Kowler, 2006; Smallwood, Fishman, et al., 2007; Smallwood & Schooler, 2006) and we replicated this finding in our data. Third, we controlled for skimming as an alternative explanation, and found global eye movement measures to be unaffected when errors were overlooked. Indeed, during mindless reading fixations were sometimes longer (Reichle, et al., 2010) and sometimes shorter (Franklin, et al., 2011) compared to mindful reading depending on whether high or low frequency target words were fixated. These findings indicate that errors may have been overlooked during episodes of mindless reading because cognitive processing is decoupled from the text.

We included different types of errors in the text to measure different levels of mindless reading. The levels of inattention hypothesis predicts that readers should be very sensitive to low-level errors and less sensitive to high-level errors. This prediction was supported by the experimental findings. Readers quite often overlooked high-level errors, like discourse errors and gibberish text. In these cases, high-level text processing may have ceased during episodes of weak mindless reading. Supporting evidence for this interpretation comes from the observation that low-level errors, like lexical and syntactic errors, were rarely overlooked. This finding is compatible with the interpretation that low-level linguistic processes like word recognition or syntactic parsing may be disrupted when low-level errors are overlooked, indicating episodes of deep mindless reading. Collectively, these results are compatible with the levels of inattention hypothesis. However, the alternative dichotomy-hypothesis can explain differences in sensitivity between error types by assuming differences in the durations of mind wandering episodes. The present eye movement analyses help distinguishing between these explanations.

*Decoupling of eye movements*. To investigate more closely how text processing changes when errors are overlooked, we performed local eye movement analyses. During mindful reading, readers slowed down to integrate words toward the end of phrases and sentences. Interestingly, this wrap-up effect was absent before errors were overlooked. This finding suggests that during mindless reading readers overlooked errors in the text because they did not integrate words to construct sentence meaning and to comprehend the text. Moreover, during mindful reading fixation durations were modulated by variables word length and frequency, which constitute empirical markers for word recognition processes. In contrast, before overlooking of errors (Exp. 2) these effects were clearly reduced (Figure 4-3 - Figure 4-5), and sometimes completely absent (Figure 4-6). This finding suggests that errors were overlooked during deep mindless reading because processes of word recognition were incomplete. Importantly, mindless reading affected eye movements on up to 20 words preceding an error sentence (Figure 4-4 and Figure 4-5). Thus, overlooking of errors did not occur because text processing was locally reduced when reading a single sentence or word. Instead, readers' minds were drifting off task over an extended period of time prior to encountering an error. In sum, the present findings suggest that overlooking errors in the SAST indicates episodes of attentional decoupling during mindless reading, where errors are overlooked because text processing is reduced.

While the present results suggest that overlooking errors in the SAST indicates episodes of mindless reading, there may be other specific factors that also contribute to overlooking of errors. Some of these may result from an absent mind; for example, monitoring of text comprehension (Palincsar & Brown, 1984; Smallwood, Fishman, et al., 2007) or memory for task instructions (McVay & Kane, 2009) may be reduced during mindless reading, and may cause readers to overlook errors in the text. Moreover, factors unrelated to mind wandering may lead to overlooking of errors, and may inflate our estimates for the occurrence of mind wandering. Also, decoupling of eye movements from the text may partially result from differences in reading ability or strategy between subjects. It should be noted, however, that we controlled for such effects in the LMM analyses. Importantly, the present eye movement results demonstrate that overlooking an error was preceded by a period of reduced cognitive text processing, indicating an episode of attentional decoupling.

*Hypotheses on the nature of attentional decoupling*. We derived several predictions from hypotheses of attentional decoupling (Figure 4-1) and tested these by analyzing eye-movement data. Critically, the levels of inattention hypothesis predicts states of weak attentional decoupling, where high-level processes are decoupled from the external environment, but low-level processes are still intact. We found eye-movement evidence for weak decoupling in Exp. 1. Here, wrap-up effects, as a measure for high-level integration processes, were reduced when errors were overlooked, but low-level lexical processes (i.e., the frequency × length interaction) remained unaffected. Deep mindless reading, to the contrary, was observed in Exp. 2, when readers had already spent much time in the lab reading boring texts. Here, not only high-level wrap-up, but even low-level lexical effects were reduced before errors were overlooked. As predicted by the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007), the consequences of the low-level decoupling in Exp. 2 may have cascaded into the cognitive system to impair higher-level wrap-up processing. These results demonstrate that graded states of weak (Exp. 1) and deep (Exp. 2) attentional decoupling can be distinguished. This finding is incompatible with a dichotomous view on attentional decoupling and provides support for the levels of inattention hypothesis.

A central prediction from the levels of inattention hypothesis is that overlooking different types of errors reflects different levels of attentional decoupling. The eye-movement data lend support to this prediction. When low-level (lexical) errors were overlooked, eye movements were decoupled from low-level (lexical) variables, and – as predicted by the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) – also high-level (wrap-up) influences were reduced (Figure 4-5). When high-level errors (discourse errors and gibberish text) were overlooked, however, then decoupling was present only for high-level integration processes (reduced wrap-up effect), but low-level lexical processing was barely affected. These eye movement results suggest that overlooking of low-level errors may indicate states of deep attentional decoupling, whereas overlooking high-level errors may indicate states of weak decoupling. These findings support the levels of inattention hypothesis and the cascade model of inattention, but are incompatible with the dichotomy-hypothesis.

As noted above, the dichotomy-hypothesis of mind wandering may explain differences in sensitivity between error types by assuming variable durations rather than variable degrees of attentional decoupling. For example, task focus during the reading of a single pseudo-word is sufficient to detect the lexical error, and the error can be detected even if attention switches quickly between mindless and mindful reading. Thus, overlooking low-level errors may reflect short-lived episodes of decoupling. To the contrary, to detect high-level discourse errors, attention must be devoted to the text during reading of at least two adjacent sentences, and overlooking high-level errors may thus indicate longer episodes of decoupling. These predictions from the dichotomy-hypothesis were not supported by the present eye movement results: fixation durations were decoupled from cognitive processing up to 20 words before encountering an error sentence, and this interval was similar (or even longer) for low-level errors (see Figure 4-5). The eye movement findings therefore suggest that overlooking low-level errors was not only associated with deeper decoupling, but potentially also with longer episodes of attentional decoupling compared to high-level errors. Both of these findings are incompatible with the dichotomous view of attentional decoupling, and are consistent with the levels of inattention hypothesis.

*Conclusions*. Cognitive science has generated theoretical models that describe different aspects of reading (Engbert, et al., 2005; Graesser, et al., 2002; Reichle, Warren, et al., 2009; Staub, 2011) and cognition in general (Cohen, 2000; Craik & Lockhart, 1972; Gazzaniga, 2009) as hierarchically organized processes, where information is represented and processed at various lower and higher levels. A long research tradition has investigated how attention affects processing at such early and late levels (Broadbent, 1958; Deutsch & Deutsch, 1963; Driver, 2001), and the field seems to agree on a continuously graded rather than a dichotomous view of attentional selection (Chun, et al., 2011; Mangun & Hillyard, 1995; Treisman, 1960). Here, we investigated how cognitive processing at different levels becomes decoupled from external information when the mind wanders away from an external reading task. Our results indicate that attentional processes during reading may be of a hierarchically graded nature. Low-level processes turned out to be quite robust against lapses in external attention and seemed to fail only when the mind was deeply absent from the current task. High-level text integration processes, to the contrary, seemed to be far more fragile and drifted off the reading task with high frequency. This result supports hierarchical models of reading and cognition. The levels of inattention hypothesis together with the cascade model of inattention provide a framework to understand and describe graded attentional decoupling at such different levels. Importantly, our findings suggest that the level of inattention may strongly vary between experiments, between experimental conditions, or measures of mind wandering, and what level of inattention is assessed in a specific study may strongly influence experimental results. Therefore, to understand and avoid potential inconsistencies, we suggest that it may be helpful to explicitly measure the depth or degree of decoupling in future studies.

*Questions for future research*. Our findings raise a new, important and open theoretical question: What factors cause decoupling at a specific weak or deep level? Based on previous theorizing, we speculate about possible causes. First, executive control processes may fail (McVay & Kane, 2009, 2010) to varying degrees and controlled high-level processes may be reduced more readily than more automatic low-level processes (Shiffrin & Schneider, 1977). Second, one question is how stimulus independent thought (SIT) is related to the graded levels of attentional decoupling. One possibility is that similar to attentional decoupling, SITs are graded in nature. Another is that SIT emerge only at a particularly deep level of decoupling. Third, the adaptive gain theory of norepinephrine function (Aston-Jones, Rajkowski, & Cohen, 1999) has been proposed as a neurophysiological basis for mind wandering (Smallwood et al., 2012; Smallwood, et al., 2011), and different levels of inattention may result from different degrees of drowsiness and inactivity ("off" state of low locus coeruleus [LC] activity) versus increased vigilance and labile attention ("tonic" mode with high baseline LC activity). Fourth, people may become aware of their wandering mind (Schooler, 2002; Schooler, et al., 2011) more easily when their cognitive processing is deeply decoupled from the external environment (as opposed to when it is only weakly decoupled), and they may therefore direct their minds back on task more often.

Another important question for future research concerns the relation of behavioral measures of attentional decoupling (like the SAST) to more subjective aspects of mind wandering. For example, our findings may trigger research to vigorously test the view that SIT is a dichotomous (versus graded) process, and to learn about how graded decoupling is related to (graded or dichotomous) aspects of SIT. Likewise, in self-report studies of mind wandering it is possible to assess whether participants are meta-aware about their mind wandering (Schooler, 2002; Schooler, et al., 2011). In fact, a recent fMRI study (Christoff, et al., 2009) found that deeper levels of mind wandering [measured as increased activity in the default network and in the executive system, (also see Christoff, 2012)] may be associated with lack of meta-awareness, and this suggests that our paradigm may have the potential to capture subjective awareness of mindless reading in an objective behavioral measure.

*Predicting mindless reading from eye movements*. As a novel contribution, we demonstrated that gaze durations predicted overlooking of lexical errors about five seconds before the error occurred in the text. Thus, recordings of individual eye movements can predict in real time whether a reader is currently in a state of mindless reading at the level of an individual trial. Such a measure may prove highly useful in diverse applications. Objective measures are useful to investigate mindlessness in populations unable to report about their wandering mind, like children or psychiatric patient groups. They could potentially be used to identify and overcome mind wandering in educational or professional settings. They could serve to diagnose individual differences in mind wandering, to objectively evaluate the quality of different texts, or to detect mindlessness in cognitive experiments or crucial real-world tasks like driving (D'Orazio, Leo, Guaragnella, & Distante, 2007) or closed-circuit television (CCTV) monitoring. In research on reading, detecting mindlessness online allows to apply sophisticated eye tracking techniques, like gaze-contingent display changes (McConkie & Rayner, 1975; Rayner, 1975, 1998), during mindless reading to investigate in detail how text processing changes when readers' minds are off task. Finally, objective measures are highly valuable tools for studying mind wandering – when investigating factors influencing the propensity to mind wandering (Sayette, et al., 2009; Sayette, et al., 2010), the consequences of off-task thought (Killingsworth & Gilbert, 2010; Smallwood, McSpadden, & Schooler, 2007), the neural structures (Buckner, et al., 2008; Christoff, et al., 2009; M. F. Mason, et al., 2007) and cognitive processes (Levinson, et al., 2012; McVay & Kane, 2010; Smallwood, 2010b; Smallwood & Schooler, 2006) that initiate, terminate, and support mind wandering and the default mode.

## 4.5 Acknowledgements

We thank Simon P. Liversedge, Shravan Vasishth, and Milena Rabovsky for valuable discussions and Frank Rösler and Jochen Laubrock for comments on the manuscript. We also thank Petra Schienmann, Sandra Beyermann, Linda Weiß, Max

Geisthardt, and Eva Hasselmann for support during preparation of the text material, data collection, and data preprocessing. We are thankful to Jonathan Smallwood, Erik Reichle and one anonymous reviewer for their helpful comments. This work was supported by DFG Grant EN 471/2.

# 5 Mental Effort during Mindless Reading? Pupil Fluctuations Indicate Internal Processing during Levels of Inattention

*Running head:* Pupil size indicates levels of inattention

## Abstract

The theoretical understanding of the cognitive processes involved in mind wandering has recently made important progress, and today the experience is known to involve decoupled external attention (and reduced task-related thought, TRT) and increased internal attention (to task-unrelated thought, TUT). TUT and TRT are often treated as an all-or-none phenomenon. The decoupling hypothesis postulates that both are distinct cognitive and neural states, which compete for domain general cognitive resources (Smallwood, et al., 2011). However, recent results from the sustained attention to stimulus task (SAST) support the levels-of-inattention hypothesis that TRT is reduced in a graded fashion, reflecting different levels of weak versus deep attentional decoupling (Schad, Nuthmann, & Engbert, 2012). Here, I analyze pupil size as a measure for cognitive workload in the SAST to test whether TUTs are graded, or whether they represent a distinct cognitive process. I found that during states of weakly decoupled TRT pupil size was reduced. However, during strongly reduced TRT in states of deep decoupling pupil size was unaffected. These findings suggest that weak decoupling reduces internal resource-demanding processing, but that deep levels of decoupling may be associated with states of internal attention and resource-demanding TUT. They further support the levels-of-inattention hypothesis that TRT is decoupled in a graded fashion, and the decoupling hypothesis that TUT is a distinct mode of deeply decoupled processing.

## 5.1 Introduction

During mind wandering attention turns away from the external environment (Schooler, et al., 2011; Smallwood & Schooler, 2006). This process of attentional decoupling reduces task-related thought (TRT) (Kam, et al., 2011; McVay & Kane, 2012b; Reichle, et al., 2010; Smallwood, Beach, et al., 2008; Smallwood, Fishman, et al., 2007; Smallwood, McSpadden, & Schooler, 2008) and causes errors in the performance of external tasks (Robertson, et al., 1997; Schad, Nuthmann, & Engbert, 2012; Smallwood, et al., 2006). During mind wandering attention also turns inwards to task-unrelated thought (TUT) (Schooler, et al., 2011; Smallwood & Schooler, 2006) and elicits neural activity in the brain's default mode network and in prefrontal executive regions (Christoff, 2012; Christoff, et al., 2009; M. F. Mason, et al., 2007; Weissman, et al., 2006). TRT and TUT are often treated as all-or-none phenomena, and the decoupling hypothesis suggests that both are distinct cognitive and neural states (Fox, et al., 2005; Smallwood, et al., 2011). However, recent evidence suggests that TRT is decoupled in a graded fashion (Schad, et al., 2012), which may question a dichotomy between internal and external thought. In the present work, I study pupil size to investigate whether TUT is graded (as TRT), or whether it occurs in an all-or-none fashion.

Based on the levels-of-inattention hypothesis (Schad, et al., 2012), attentional decoupling is a graded phenomenon that occurs at different levels of cognitive processing: during weak decoupling external attention is reduced at a late processing stage, and higher-level, but not lower-level processes are decoupled (Deutsch & Deutsch, 1963). During deep decoupling, however, higher and lower processing levels are reduced, and the cascading consequences of early attentional selection (Broadbent, 1958) may cause decoupling at higher levels (Smallwood, 2011b; Smallwood, Fishman, et al., 2007) (also see Lavie, 2005). Support for levels of inattention comes from the sustained attention to stimulus task (SAST) (Schad, et al., 2012), which catches different levels of decoupling during reading via overlooking of errors at different levels of the text. Moreover, graded decoupling is consistent with results from the sustained attention to response task (SART) (Cheyne, et al., 2011; Cheyne, et al., 2009; McVay & Kane, 2012a; Smallwood, 2010a; Smallwood, McSpadden, Luus, et al., 2008). Based on the levels-of-inattention hypothesis the question arises how different levels of decoupled TRT are related to internal cognitive activity like TUT, and whether TUT itself is a graded or a dichotomous process.

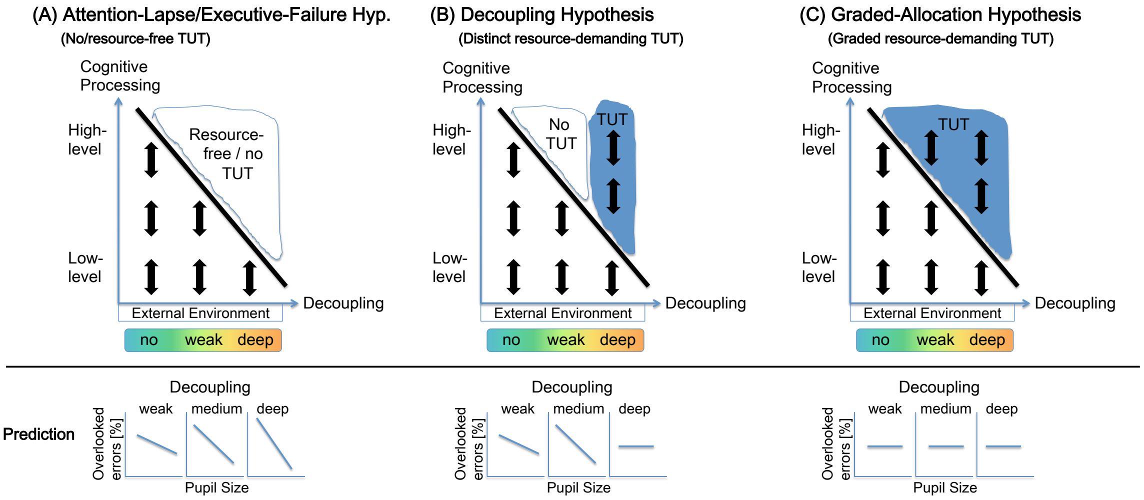


**Figure 5-1**

Illustration of theoretical hypotheses about resource-demanding internal cognitive activity during levels of inattention. Upper panels: The y-axis represents different higher and lower levels of cognitive processing; the x-axis represents different weak and deep levels of decoupling. Cognitive processing that is coupled to the external environment (below the thick black line) demands cognitive resources (black bidirectional arrows); decoupled internal thought (above the black line) may (black arrows, blue area) or may not (no arrows, white area) demand resources. (A) The attention-lapse hypothesis and the executive-failure hypothesis assume that mindless reading does not involve internal resources. (B) The decoupling hypothesis assumes that resource demanding TUT causes deep decoupling, and that weak decoupling lacks TUT. (C) The graded-allocation hypothesis assumes graded division of cognitive resources between internal and external information. Lower panels: Predictions for how pupil size relates to different levels of decoupling (measured as % overlooked SAST errors) during a demanding external task like reading.

How reduced external attention and TRT is linked to internal TUT has remained unclear (cf. McVay & Kane, 2009; Smallwood, in press) and competing hypotheses exist about the role of internal processing in attentional decoupling. The decoupling hypothesis (Smallwood, 2010b, in press; Smallwood & Schooler, 2006) assumes that online and offline thought are distinct modes of cognitive functioning with discrete switches between modes (Smallwood, et al., 2011), and that these modes are supported by distinct and anti-correlated neural networks (Fox, et al., 2005). Indeed, evidence suggests that internal attention during mind wandering draws on domain general processes (Baird, Smallwood, & Schooler, 2011; Christoff, 2012; Christoff, et al., 2009;

Levinson, et al., 2012; M. F. Mason, et al., 2007; Smallwood, et al., 2012; Smallwood, et al., 2011; Stawarczyk, et al., 2011). Based on the decoupling hypothesis, TUT necessitates deep levels of attentional decoupling because it competes with TRT for domain-general cognitive resources (see Figure 5-1B).

As an alternative view, cognitive resources may be gradually engaged with TUT, and I here suggest the graded-allocation hypothesis that cognitive resources are continuously divided between processing of internal and external information (Figure 5-1C). As a third possibility, the executive-failure hypothesis suggests that TUT is resource-free and results from executive failures to keep processing resources on task (McVay & Kane, 2009, 2010). Moreover, reduced external processing may result from lapses of sustained attention or reduced vigilance (Aston-Jones & Cohen, 2005; Oken, Salinsky, & Elsas, 2006; Robertson, et al., 1997) and may lack internal thought processes such as TUT (Figure 5-1A).

Large pupil size provides a direct neurophysiological indicator of mental effort (Beatty, 1982a, 1982b; Hampson, Opris, & Deadwyler, 2010; Janisse, 1977; Karatekin, Marcus, & Couperus, 2007), cognitive control (Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003), and specifically cognitive load during language processing (Beatty, 1982b; Hyönä, et al., 1995; Just & Carpenter, 1993). It is also often used as a proxy for activity in the norepinephrine (NE) brain system (Einhäuser, Stout, Koch, & Carter, 2008), where (i) small pupils indicate drowsy or inattentive cognitive states with low tonic NE activity, (ii) medium-sized baseline pupils and task-evoked responses indicate states of focused attention involving moderate tonic NE activity and transient task-evoked NE bursts, and (iii) large baseline pupils index states of labile attention with high tonic NE activity (Aston-Jones & Cohen, 2005).

Smallwood and colleagues (Smallwood, et al., 2012; Smallwood, et al., 2011) recently suggested that high tonic NE activity may be related to mind wandering, which seems consistent with their finding that in relatively undemanding tasks perceptual decoupling and with TUT were associated with increased baseline pupil size and reduced evoked responses. Moreover, Smallwood and colleagues (Smallwood, et al., 2011) found support for the decoupling hypothesis that off-line and on-line thought are distinct modes of cognitive functioning as baseline pupil size, as a measure for internal processing, showed a highly non-linear increase for trials with very slow reaction times reflecting decoupled TRT.

In the present work, I measure pupil size in the sustained attention to stimulus task (SAST, Schad, et al., 2012) to study how internal cognitive workload is related to different levels of decoupled TRT. The SAST catches episodes of decoupling at different levels behaviorally based on psychophysics of error detection for errors at different levels of the text. The attention-lapse hypothesis (Aston-Jones & Cohen, 2005; Oken, et al., 2006) and the executive-failure hypothesis (McVay & Kane, 2009, 2010) predict that low cognitive load or effort (i.e., small pupil size) is associated with attentional decoupling, and that this association is stronger for deeper levels of decoupling (see Figure 5-1A, lower panel). The graded-allocation hypothesis predicts that overall cognitive load (i.e., pupil size) is not closely related to perceptual decoupling in demanding external tasks because reduced external attention goes along with increased internal cognitive activity (see Figure 5-1C, lower panel). Third, the decoupling hypothesis (Smallwood, 2010b; Smallwood, et al., 2012; Smallwood, et al., 2011; Smallwood & Schooler, 2006) predicts that deep levels of decoupling should be unrelated to cognitive load (or even show larger pupils during deep decoupling), but that medium or weak decoupling should show reduced cognitive load (i.e., pupil size; see Figure 5-1B, lower panel).

## 5.2 Method

Thirty subjects each participated in two experiments employing the SAST (Schad, et al., 2012), which were conducted in direct succession. Each experiment involved reading of 25 boring and easy short stories from the Potsdam Mindless Reading Corpus (PMC), and the 50 stories comprised a total of about 17,500 words. Different error types (lexical, syntactic, semantic, discourse errors, and gibberish text) were included into the text. Subjects were instructed to read the text in a relaxed way and to press the space bar whenever they notice an error in the text. Overlooking errors in the SAST indicates attentional decoupling during episodes of mindless reading, and detecting errors indicates attentional coupling during normal reading. Errors were included at 48 out of 62 error locations. The remaining error locations contained no error, yielding a control condition where the state of mind is not known. Pupil area and gaze of the left eye were tracked using an EyeLink 1000 Remote system (SR Research).

Standard procedures were used to remove invalid items, trials in which readers skimmed over the text, false alarm trials, and invalid fixations from the data (cf., Schad, et al., 2012). Average pupil sizes during the first fixations in first-pass reading on the 24 words preceding an error sentence were analyzed. Importantly, pupil size during mindless and mindful reading and in the control condition, as well as for different error types was measured on the exact same text material, keeping visual features of the display constant across conditions. Testing was done in a completely dark room with minimal monitor brightness and a dark grey-brown text background to minimize visual influences on pupil measurement. Due to incorrect tracker settings data for three participants was removed for analysis. (Generalized) Linear mixed effects models (GLME) as implemented in the *lme4* package (Bates & Sakar, 2008) in the R System for Statistical Computing (R Development Core Team, 2012) were used for statistical analyses, including random effects for subjects and text pages. For the LMEs, $t$-values were treated as $z$-values to compute approximate $p$-values (cf. Kliegl, et al., 2011).

## 5.3 Results

Across both experiments, readers overlooked 32,3% of the errors, indicating episodes of mindless reading. Correspondingly, 67,7% of the errors were detected, indicating that readers were on task. Analysis of pupil size revealed striking overall effects of perceptual decoupling (see Figure 5-2A): averaged over all error types and both experiments pupils at the onset of the error sentence were significantly smaller when errors were overlooked than when they were detected ($b = 18.8$, $SE = 2.4$, $t = 7.9$, $p < .001$) and than in the control condition ($b = 11.5$, $SE = 2.7$, $t = 4.3$, $p < .001$) (this was estimated by centering the linear predictor at zero words prior to the error sentence). The effect of mindless versus mindful reading became stronger as readers approached the error sentence ($b = 1.2$, $SE = 0.34$, $t = 3.6$, $p < .001$), indicating that pupil size was temporally closely associated with zoning out. Additional analyses revealed that these effects were robust when controlling for higher-order polynomials of time on task, line number, word number, fixation duration, saccade amplitude, word length, and word frequency.

I also tested whether average pupil size on the five words prior to an error sentence can be used to predict overlooking the upcoming error. As is displayed in Figure 5-2B, small pupils predicted mindless reading and large pupils predicted mindful reading (binomial GLME: $b = -0.43$, $SE = 0.12$, $p < .001$). A mixed effects signal detection analysis (Wright, et al., 2009) of error detection (see Figure 5-2C) revealed that pupil size in a region ten words prior to the target sentence predicted sensitivity for errors ($\Delta d' = 0.26$, $SE = 0.26$, $p < .06$), but not response bias ($\Delta c = 0.06$, $SE = 0.14$, $p = .6$).

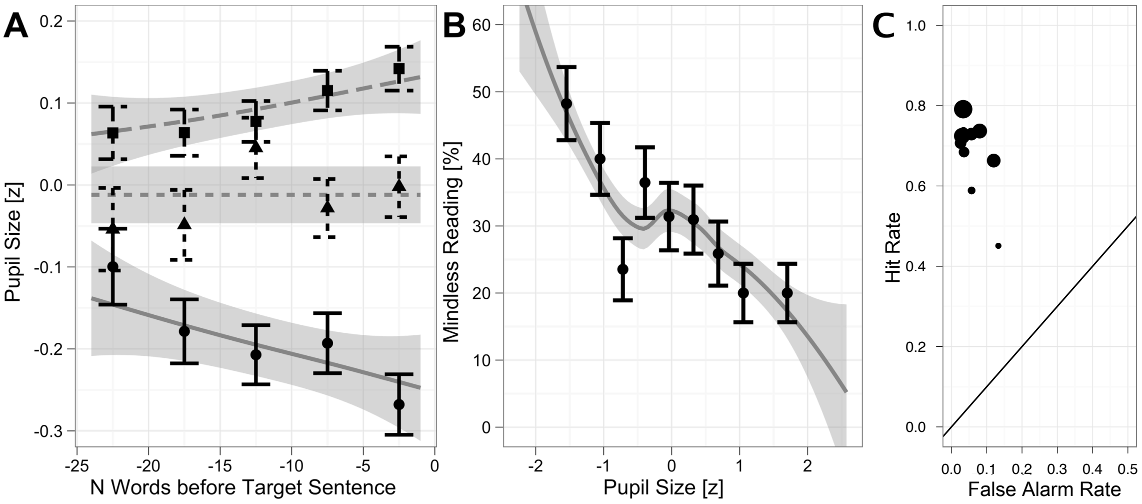


**Figure 5-2**

(A) Pupil size during mindless reading (error not detected, solid bottom line), mindful reading (error detected, dashed upper line), and in a control condition (no error, dotted middle line) for different number of words prior to the error sentence. (B) Likelihood of mindless reading (percent overlooked errors) as a function of average pupil size on the five words prior to the target sentence. Target sentences containing no error are excluded. (A+B) Local regression models (LOESS) and associated standard errors for the ungrouped data are displayed, as well as means and SEM for bins, each containing data from at least 400 fixations (panel A) or 85 target screens (panel B). (C) Hits (i.e., correctly detected errors) over false alarms (i.e., falsely detected errors) for different pupil sizes, where large points indicate bins with larger pupil sizes.

Next, I performed analyses for different error types (coded via Helmert contrasts) for each experiment separately, and used average pupil size on the 10 words prior to the error sentence to predict overlooking of upcoming errors (see Figure 5-3). As predicted by the levels-of-inattention hypothesis, in both experiments high-level errors were more frequently overlooked compared to low-level errors (for all comparisons: *ps* < .05; except for difference between syntactic and semantic errors in Exp. 2: *p* = .29), suggesting that overlooking high-level errors reflects weak decoupling whereas overlooking low-level errors reflects deep decoupling (cf., Schad, et al., 2012).

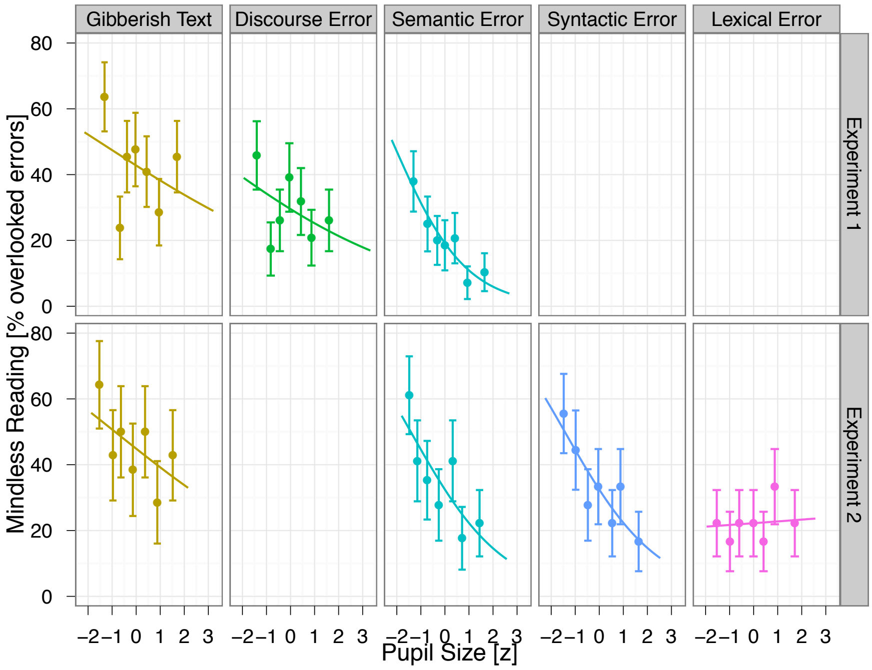


**Figure 5-3**

Percentage of overlooked errors for different error types in two experiments as a function of average pupil size measured during fixations on the 10 words prior to the error sentence. Overlooking different types of errors assesses different levels of attentional decoupling, ranging from weak decoupling (left panels, gibberish text and discourse errors), via medium decoupling (middle panels, semantic and syntactic errors), to deep decoupling (right lower panel, lexical errors). Displayed are predictions from linear logistic regression models as well as means and SEM for pupil-size bins, each containing data from at least 13 text pages.

Moreover, for Exp. 1 (see Figure 5-3, upper panels) small pupil size predicted overlooking of errors ($b$ = -0.40, $SE$ = 0.17, $z$ = -2.4, $p$ < .02) and this prediction was stronger for medium-level (semantic) errors than for high-level errors (gibberish text and discourse errors) ($b$ = -0.18, $SE$ = 0.09, $z$ = -1.94, $p$ = .06), indicating that deeper levels of decoupling were more closely associated with pupil size. In Exp. 2 (see Figure 5-3, lower panels), small pupil size predicted overlooking of high-level (gibberish text) and medium-level (semantic and syntactic) errors ($b$ = -0.55, $SE$ = 0.17, $z$ = -3.2, $p$ < .01), and numerical differences in predictions between these error types were not significant ($ps$ > .25). However, pupil size did not significantly predict overlooking of low-level lexical errors ($b$ = 0.04, $SE$ = 0.23, $z$ = 0.2, $p$ = .85) and this prediction was significantly weaker than the one for medium- and high-level errors ($b$ = 0.50, $SE$ = 0.25, $z$ = 1.97, $p$

$< .05$), suggesting that pupil size may not be reduced when low-level errors are overlooked. To test whether pupil size is more closely related to error detection for medium-level (semantic and syntactic) errors than for high-level errors (gibberish text and discourse errors), I combined data from both experiments and found statistical support for this hypothesis ($b = -0.20$, $SE = 0.09$, $z = -2.3$, $p < .05$).

In an additional analysis, I tested whether during states of deep decoupling pupil size is decoupled to external information. I found that before lexical errors were detected pupil size was larger for low-frequency than for high-frequency upcoming words n+1, which reflects the increased processing load ($b = -0.031$, $SE = 0.016$, $t = -1.98$, $p < .05$). This effect, however, was absent before lexical errors were overlooked ($t = 0.87$, $p > .10$; for the slope-difference: $b = 0.056$, $SE = 0.031$, $t = 1.77$, $p < .10$), supporting the view that relatively large pupils during deep decoupling do not reflect external processing, but may reflect internal cognitive activity.

## 5.4 Discussion

Based on the levels-of-inattention hypothesis, attentional decoupling of task-related thought (TRT) is a graded phenomenon that occurs at different lower and higher levels of cognitive processing (Schad, et al., 2012). Here, I analyzed pupil size as a well-studied neurophysiological indicator of cognitive load or mental effort (Beatty, 1982b; Hampson, et al., 2010; Janisse, 1977) to investigate how different levels of decoupled TRT are related to internal cognitive activity, i.e., task-unrelated thought (TUT). The results support graded decoupling of TRT (i.e., the levels-of-inattention hypothesis) (Schad, et al., 2012) and dichotomous occurrence of TUT (i.e., the decoupling hypothesis) (Smallwood, 2010b; Smallwood, et al., 2011).

I found that weak and (particularly) medium levels of decoupling were closely related to small pupil size, indicating reduced load or intensity of text processing (Just & Carpenter, 1993) during weak and medium levels of mindless reading. This finding is consistent with the hypothesis that weak and medium levels of decoupling represent lapses of external attention that lack resource-demanding TUT (Aston-Jones & Cohen, 2005; McVay & Kane, 2010; Oken, et al., 2006; Smallwood, et al., 2011). Moreover, I found that during deep decoupling pupil size was not reduced. Deep (lexical) decoupling was found to be characterized by strongly reduced influences of lexical and linguistic variables on gaze durations, and by overlooking of lexical errors (see Schad, et al., 2012).

That strongly reduced external processing during deep decoupling does not result in a reduction of overall cognitive load provides support for resource-demanding internal cognitive activity that is unrelated to the external reading task. In support of this interpretation, I found that lexical influences on pupil size, which were present during task-focus, were absent during deep decoupling. These results are consistent with the decoupling hypothesis that deep decoupling results from resource competition between distinct neural systems for online and offline thought (Smallwood, 2010b; Smallwood, et al., 2011) and that internal cognitive activity like TUT results in deep lapses of external attention and strongly reduced TRT (cf. Kam, et al., 2011; Smallwood, et al., 2011; Smallwood, Fishman, et al., 2007). They add to the increasing literature supporting domain-general processes in mind wandering (for review see Smallwood, in press) and suggest their selective engagement during deep, but not during weak levels of decoupling. Moreover, the present findings suggest that weak/medium versus deep levels of decoupling may involve qualitatively different internal cognitive processes, which is inconsistent with the graded-allocation hypothesis that cognitive resources are continuously divided between internal and external information, and with the hypothesis that decoupling always lacks resource-demanding TUT (McVay & Kane, 2009, 2010; Oken, et al., 2006).

The adaptive gain theory of NE functioning (Aston-Jones & Cohen, 2005) has been suggested as a neurophysiological basis of mind wandering (cf. Smallwood, et al., 2012; Smallwood, et al., 2011), and the present findings may extend this view to explain different levels of decoupling: they suggest that weak/medium decoupling may reflect a state of inattention and drowsiness, where small pupil size indicates low tonic NE activity. During deep decoupling, to the contrary, high tonic NE activity may reflect a state of labile attention that is unrelated to the external task. Clearly, future research is needed to evaluate the link between mind wandering and NE.

An interesting question that arises from the present findings concerns the temporal relation between different levels of decoupling. As one possibility, weak and deep levels of decoupling may represent different kinds of decoupling, and an individual episode of decoupling may occur at either a weak or a deep level. As an additional possibility, weak/medium and deep levels of decoupling may both occur during a single episode of mind wandering: an individual episode may occur or develop via a gradual fading of external attention, where states of weak and medium decoupling, which lack internal cognitive activity, may precede and facilitate states of deep decoupling and TUT. In support of the latter possibility, the present results suggest that small pupils during weak/medium levels of decoupling may not only indicate, but may also affect mental processes. Specifically, the balance between attention to external versus internal information may in part be modulated by embodied components of mind wandering, by physically blocking external information at sensory endings, e.g., via blinks (Smilek, Carriere, & Cheyne, 2010c). Likewise, small pupils during weak and medium decoupling reduce the quantity of light entering the eyes, and the resulting slow down in cortical processing of external information (Hawkes & Stow, 1981; Lovasik, Spafford, & Szymkiw, 1985; Martins, Balachandran, Klistorner, Graham, & Billson, 2003; W. Müller, Kollert, & Zachert, 1988) may affect trade-offs between internal and external thought by weakening representations of external information, and may facilitate the occurrence of TUT.

Understanding the link between external attention (TRT) and internal attention (TUT) is also crucial for the development of objective behavioral online-measures to assess internal thought processes via continuous behavioral recordings (Franklin, et al., 2011; Schad, et al., 2012). It is an exciting perspective that such measures may provide empirical access to the onset, the time course, and the offset of an internal train of thought, and may help to elucidate the processes that initiate, support and terminate mind wandering (cf. Smallwood, in press). The present work may contribute to the theoretical and empirical foundation for such an endeavor.

I conclude that graded levels of external inattention may reflect qualitatively different internal states. Attentional decoupling sometimes results from lapses of external attention that lack internal cognitive activity. At other times, however, decoupling seems to be associated with resource demanding internal thought, which reflects spontaneous activity in a default mode (Buckner, et al., 2008; Christoff, et al., 2009; M. F. Mason, et al., 2007). Disentangling these mental processes and how they interact in real time during an individual episode of mind wandering may provide insights into an elusive everyday phenomenon that has long escaped scientific investigation at an exciting and previously unimagined level of detail (Smallwood, in press).

# 6 General discussion

The phenomenon of mindless reading is common. Most people know the experience of reading when suddenly their mind wanders away from the text to other thoughts and feelings, like the tragic end of a countries' participation in the UEFA Champions League. When the mind returns to the reading task readers often have no memory for the text just read (Smallwood, McSpadden, & Schooler, 2008). Although mindless reading is common (Killingsworth & Gilbert, 2010; McVay, et al., 2009; Schooler, et al., 2004) and seems to be widely recognized in lay psychology, its scientific investigation and understanding was long underdeveloped. What happens in a readers' mind and how readers move their eyes when the mind wanders is currently only beginning to be better understood (Reichle, et al., 2010; Schooler, et al., 2011; Smallwood, 2011b, in press).

In the present work, my colleagues and I presented four studies on mindless reading. These studies aimed at providing methodological, empirical, theoretical, and computational contributions to the understanding of mindless reading, mind wandering, and eye movement control. As a main contribution, we introduced new behavioral paradigms to study mindless reading – the shuffled text reading task (Chapters 2 and 3) and the sustained attention to stimulus task (SAST, Chapters 4 and 5) – and demonstrated that it is possible to predict states of mindless reading from eye movement recordings (Chapters 4 and 5). Second, we introduced a theoretical hypothesis on the nature of attentional decoupling – the levels of inattention hypothesis – and provided empirical evidence in support of graded decoupling (Chapters 4 and 5) and of the decoupling hypothesis assuming dichotomous task-unrelated thought (Chapter 5). Third, we studied eye movement control during mindless reading and found reduced influences from higher-level cognitive language processing, but relatively robust lower levels of eye movement control (Chapters 2-4). Fourth, in the shuffled text paradigm, we observed influences of cognitive processing on the zoom lens of attention (Chapters 2) and simulated these results by introducing an advanced version of the SWIFT model of saccade generation (SWIFT 3), which incorporates the zoom lens of attention in a mathematical model of eye movement control (Chapter 3). In the following sections, I will discuss the present findings with respect to empirical approaches to mindless reading (Section 6.1), the nature of attentional decoupling and mind wandering (Section 6.2), and the (non-)cognitive control of eye movements (Section 6.3).

## 6.1 New empirical approaches to mindless reading

In the present work, my colleagues and I introduced new behavioral measures to investigate mindless reading, and I will here discuss the present findings and an outlook for future research.

### 6.1.1 Shuffled-text reading: Weakly decoupled eye movements

To approximate states of weak mindless reading in the laboratory, my colleagues and I studied the reading of shuffled text. The present findings on reduced frequency effects during shuffled text reading suggest that eye movement behavior is decoupled from cognitive processing (Chapter 2). Simulations of the SWIFT 3 model (Chapter 3) traced these reductions to a decoupling of specific cognitive processes: estimation of model parameters suggested that during the reading of shuffled text, lexical influences on word activations and foveal inhibition of autonomous saccade generation was reduced. These findings support the view that shuffled text reading provides a paradigm to study weakly decoupled (i.e., ‘mindless’) control of eye movements. However, the results also suggest that the zoom lens of attention may be more strongly modulated by processing difficulties for shuffled than for normal text. Thus, decoupling of eye movements in shuffled text seems to go along with an increased coupling of the zoom lens of attention, pointing to an interesting dissociation between attention and eye movements in shuffled text reading.

The results on shuffled text reading are currently open to alternative interpretations, including specific influences from linguistic or memory processes. However, our analyses based on the SWIFT model generated several predictions on how various reading processes may be affected during shuffled text reading, and it would be very interesting to test these predictions in future experiments. At present, our findings show that during shuffled text reading, some eye movement effects are similar to those during normal text reading, while others strongly differ. This pattern of findings provides an interesting test case for models of eye movement control, where different models may propose differing explanations for the experimental results (Roberts & Pashler, 2000).

### 6.1.2 Sustained attention to stimulus task (SAST)

To catch spontaneous episodes of mindless reading during natural reading in a behavioral measure we developed the sustained attention to stimulus task (SAST; Chapters 4 and 5), which assesses attentional decoupling via overlooking of errors in the text. The present findings from analyses of error detection (Chapters 4 and 5), lexical and linguistic influences on fixation durations (Chapter 4), and pupil size as a measure of cognitive load or effort (Beatty, 1982b; Janisse, 1977) (Chapter 5) provide converging support that the SAST provides a behavioral measure of mindless reading.

In the present work, we aimed at testing predictions from the levels of inattention hypothesis (Chapter 4) and the cascade model of inattention (Smallwood, 2011b; Smallwood, Fishman, et al., 2007). To this end we designed the SAST to assess different levels of decoupling via errors at different levels of the text. Again, this usage of the SAST was supported by the experimental findings. Consistent with theoretical predictions, we found that readers were more sensitive to low-level than to high-level errors (Chapter 4). Moreover, lower- and higher-level linguistic influences on eye movements were reduced in a graded fashion when different types of errors were overlooked (Chapter 4). Additional support that the SAST measures levels of inattention comes from analyses of pupil size, which suggest that weak and medium levels of decoupling are associated with reduced mental effort, but that deep decoupling may indicate internal cognitive activity (TUT, Chapter 5). These patterns of findings are consistent with the levels of inattention hypothesis that decoupling is a graded phenomenon (Chapters 4 and 5), and the decoupling hypothesis (Smallwood, 2011b; Smallwood, et al., 2011) that TUT are dichotomous and occur during deep, but not during weak levels of decoupling (Chapter 5). Together, these results provide important steps toward establishing the SAST as a measure for graded levels of attentional decoupling during mindless reading.

For future studies, the SAST provides a candidate behavioral measure to assess mindless reading and attentional decoupling in diverse settings. From an applied perspective, the SAST may provide a measure to assess the propensity for mindless reading in individual persons or for specific texts. For example, such a measure may contribute to understanding what is common among individual differences in mind wandering, intelligence (Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011; Mrazek, et al., 2012), and reading comprehension (McVay & Kane, 2012b), or may be helpful in educational (Smallwood, Fishman, et al., 2007) or clinical contexts. Moreover, the SAST may be useful to assess the degree to which a specific text (e.g., a textbook or a novel) attracts readers' interest and attention, and to evaluate the quality of the writing. Given these potential applications, future research is desirable to investigate the reliability, validity and generalizability of the SAST.

First, it would be interesting to test whether measuring the propensity to mindless reading using the SAST is reliable. Important steps toward such applications could be investigations of the internal and retest reliability of the SAST. Second, a large range of research designs and methods could be used to evaluate the convergent, discriminant, and construct validity of the SAST. An important step would be to investigate whether various factors known to influence mind wandering and decoupling (Robertson, et al., 1997; Sayette, et al., 2009; Sayette, et al., 2010; Shaw & Giambra, 1993; Smallwood, Fitzgerald, Miles, & Phillips, 2009; Smallwood, O'Connor, & Heim, 2004-2005) also affect sensitivity for errors in the SAST. Combining or correlating the SAST with other measures of mind wandering would provide additional information on how different levels of attentional decoupling are related to self-reported TUT, SIT, meta-awareness of mind wandering, and to other behavioral and questionnaire measures of attentional decoupling and sustained attention (Smallwood, et al., 2004).

Third, the SAST is based on a very general design principle, namely on assessing levels of decoupling via low sensitivity for different error types. Based on this principle, it would be interesting to generalize the SAST to explore different levels of decoupling in tasks other than reading. For example, it may be possible to design versions of the SAST assessing decoupling in a variety of active vision tasks, like visual search, natural scene viewing, or dynamic scene viewing, by including different kinds of error stimuli in the visual display. As an interesting feature, the SAST technology may even be applicable to assess decoupling in tasks, which involve no active external responses, like passive listening. Last, it would be desirable to design a simple and well-controlled laboratory (vigilance) version of the SAST (using simple stimuli, as they are also used in the SART) to more closely investigate the basic cognitive processes that fail when errors are overlooked. Comparing such basic cognitive processes between the SART and the SAST may provide valuable information on their convergent and discriminant validity.

### 6.1.3 Objective online-markers for mind wandering

One aim of the present work was to find objective online-markers for mindless reading by recording readers' eye movements. We found that eye movements predicted episodes of mindless reading in the SAST. Individual gaze durations that readers made on specific target words predicted whether an upcoming lexical error was detected average five seconds before the error occurred in the text (Chapter 4). Moreover, pupil size measured on the words prior to an error sentence predicted whether upcoming medium- or high-level errors were overlooked (Chapter 5). To my knowledge, these findings provide the first demonstrations that eye movements can be used as objective online-markers to predict states of mindless reading.

Using eye movements to detect mind wandering online provides a promising starting point for future investigations as it has many practical advantages compared to commonly used measures based on performance errors (like the SART or the SAST) or experience sampling. When using behavioral online measures to assess decoupling there is no need to intervene with ongoing task performance. Therefore, based on online measures it is possible to assess how standard measures of mind wandering may alter the occurrence of the phenomenon of interest. Moreover, they may provide access to the cognitive processes involved in subjective self-reports of mind wandering. The capability of online measures to collect more data on mind wandering in a given experiment should generally increase experimental power and be highly useful in neuroimaging studies. Online measures can also be used for advanced experimental designs in eye-tracking experiments by performing gaze-contingent display changes when the mind is absent to find more definite answers on the relation between mind wandering and text processing.

Importantly, such practical advances in the measurement of mind wandering may also contribute to resolve problems in the understanding of the phenomenon, and I will discuss this possibility in the next section (6.2).

## 6.2 The nature of attentional decoupling

During mind wandering, attention turns away from the external environment and cognitive processing is decoupled from perceptual information. It was a major aim of the present work to study this process of attentional (or perceptual) decoupling, and to investigate its' theoretical understanding.

A long-standing question in attention research has been how attentional selection attenuates cognitive processing at early versus late processing stages (Broadbent, 1958; Deutsch & Deutsch, 1963; Driver, 2001). Today evidence exists that attentional selection can operate at either early or late stages (Chun, et al., 2011; Lavie, 2005). In the present work, we introduced the levels of inattention hypothesis (Chapters 1 and 4) suggesting that different levels of early versus late attentional selection spontaneously occur during normal task performance, which reflects different graded levels of weak versus deep attentional decoupling. Tests of this hypothesis – against a dichotomous conception of decoupling – using the SAST were based on analyses of readers' sensitivity for different error types, of lexical and linguistic influences on fixation durations (Chapter 4), and of pupil size as a measure of cognitive load or mental effort (Chapter 5). The results provide support to the levels of inattention hypothesis.

An interesting question that emerges from these findings is whether decoupling in the SAST is associated with internal cognitive activity like TUT, and whether TUT are also of a graded, or rather of an all-or-none nature. In chapter 5, I analyzed pupil size, a widely used measure for cognitive load or mental effort (Beatty, 1982b; Janisse, 1977; Just & Carpenter, 1993). I found that weak and medium levels of decoupling were closely associated with reduced cognitive load, suggesting that the intensity of text processing is reduced (Just & Carpenter, 1993). However, although external processing is strongly reduced during deep levels of decoupling (see Chapter 4) I found no reduction in pupil size (i.e., cognitive load) in this state, suggesting that deep decoupling may be associated with resource-demanding internal cognitive activity (cf. Christoff, et al., 2009; Levinson, et al., 2012; Smallwood, 2010b; Smallwood & Schooler, 2006). These findings are consistent with the decoupling hypothesis that internal and external thought are distinct cognitive modes, which are based on distinct and anti-correlated neural networks, and which compete for limited domain general cognitive resources (Smallwood, et al., 2011; Smallwood & Schooler, 2006). Moreover, the findings are inconsistent with the views (a) that decoupling always lacks internal processing, (b) that TUT is resource-free (McVay & Kane, 2009, 2010), or (c) that reductions in external attention always go along with continuous increases in internal processing.

These findings pose new questions on the nature of attentional decoupling. In the following, I will discuss the levels-of-inattention hypothesis with respect to (a) further empirical tests, (b) specific mechanisms of decoupling, (c) the occurrence versus process of decoupling, and (d) a general cognitive framework.

*Further empirical tests*. Future studies are desirable to further test predictions from the levels-of-inattention hypothesis, and I discuss two possible tests here. First, the present eye movement analyses tested predictions about graded decoupling of low-level lexical and high-level wrap-up processes. However, the levels-of-inattention hypothesis does not only assume two levels of decoupling (high-level vs. low-level), but many different levels. To test this assumption it would be interesting to analyze how influences on eye movements from variables at many different processing levels are decoupled in the SAST. Such levels could include visual, sublexical (e.g., orthographic or phonological), lexical, syntactic, semantic, and discourse processing, and episodic memory encoding. Investigating variables at these different levels could further test theoretical predictions, and may allow to reliably predict from eye movements whether a reader is currently decoupled from the text at a weak or at a deep level. Such a non-intrusive online-measure of the level of decoupling could provide exciting insights in the temporal dynamics of how attention spontaneously fluctuates during task performance. Second, markers for different levels of decoupling can be derived from neuroimaging methods like EEG and fMRI, and studying the SAST in combination with these methods would provide interesting tests of the levels-of-inattention hypothesis and the cascade model of inattention (by investigating reduced influences from external information on neural processing) and the decoupling hypothesis (by measuring the involvement of the default mode network and executive regions in different levels of decoupling).

*Mechanisms of decoupling*. An interesting question for future research concerns the cognitive mechanisms that constitute perceptual decoupling during mindless reading. In the present thesis (Chapter 3), we studied decoupling in a mathematical model of eye movement control (SWIFT). Model simulations for normal and mindless (i.e., shuffled text) reading revealed highly detailed and theoretically explicit information on how cognitive processes linking eye movements to the text might be decoupled during mindless reading (for details, see below, Section 6.3). Clearly, model simulations are desirable to test whether these results generalize to paradigms like the SAST that assess spontaneous decoupling during the reading of normal text. The application of mathematical models provides an interesting and detailed level of theoretical analysis in the study of attentional decoupling.

*Process versus occurrence of decoupling*. In standard attention experiments in the stimulus-response paradigm, like experiments on selective listening (Deutsch & Deutsch, 1963) or attentional cueing (Posner, 1980), attention is manipulated via controlled experimental conditions. Hence, experimental manipulations cause changes in attention. To the contrary, attentional decoupling and mind wandering occur spontaneously. Therefore, the question arises as to what events cause the occurrence of an episode of mind wandering – a question that is often confused with understanding the process of decoupling itself (cf. Smallwood, in press).

Understanding the causes of spontaneous thought processes seems like a difficult endeavor and the precise onset or offset of individual episodes of mind wandering is generally unknown. This difficulty with the measurement of mind wandering makes it hard to investigate the cognitive and brain processes that give rise to, or terminate an internal train of thought. Therefore, methodological advances are desirable to measure mind wandering via online-markers in continuous behavioral (e.g., eye movements, see Chapters 4 and 5) or neurophysiological recordings (see previous Section 6.1.3).

It is a large challenge and an exciting possibility that online-markers may be able to detect the onset, the time course, and the offset of individual episodes of mind wandering. Such a measure would provide the highly interesting possibility to investigate the events and the cognitive and brain processes that initiate and terminate mind wandering. For example, decoupled thought (like TUT or SIT) may be initiated via spontaneous activity in the hippocampus, or in brain areas responsible for motivation or decision-making [e.g., see current concerns hypothesis, (Klinger, Gregoire, & Barta, 1973)]. These activations, in turn, may arise when executive control fails to keep processing resources on task (McVay & Kane, 2009, 2010) or when meta-awareness of mental states vanishes (Schooler, 2002; Schooler, et al., 2011). Likewise, mind wandering may cease when meta-awareness is reestablished or control processes are again directed on the task.

Based on online markers it would moreover be interesting to study the ‘how’ of mind wandering, that is, the processes involved in entertaining an internal train of thought and in insulating internal thought processes from interruptions from external events (Smallwood, et al., 2012; Smallwood, et al., 2011).

The levels-of-inattention hypothesis may provide a fruitful theoretical framework to investigate these questions. First, the events causing the occurrence of attentional decoupling may differ between levels of decoupling. For example, it is a possibility that weak decoupling results from failures of executive control processes, while deep decoupling results from the insulation of daydreams against external interference. Second, attention may dynamically transition between different levels of decoupling during a given decoupling episode, and this possibility raises the question of what processes may cause and mediate transitions between levels of inattention. Here, I speculate that, sometimes, failures of executive control may lead to weak levels of decoupling, and that this weak decoupling may precede episodes of deeper decoupling and TUT. Thus, the emergence of an individual mind wandering episode may be temporally extended, and involve dynamic attentional processes. To the contrary, switches back to the external task may often occur more abruptly when individuals become meta-aware of their mind wandering thoughts and individuals are motivated to direct attention back on task. Combining online-markers of decoupling at different levels (e.g., based on eye tracking) with neuroimaging methods like EEG and fMRI or with cognitive modeling may provide exciting insights in such subtle attentional dynamics.

*A general framework*. Theoretically, the levels-of-inattention hypothesis can be derived from more general principles of the functioning of the cognitive system. As a general theoretical perspective, attentional resources may be freely and continuously distributed between information from external and internal sources (Navon & Gopher, 1979) (note, however, that the results in chapter 5 did not support a continuous view on internal TUT). Moreover, different principles, like hierarchical, serial, and parallel processing, characterize the organization of different cognitive processes in the mind. Combining the assumption of free attention allocation with organizing principles of cognitive processes allows deriving hypotheses on the nature of attentional decoupling. For example, as an important principle, cognitive processes are organized in different hierarchical levels (Craik & Lockhart, 1972; Gazzaniga, 2009). Assuming free attention allocation across different hierarchical levels provides a theoretical derivation of the levels-of-inattention hypothesis. This general theoretical perspective also allows deriving other hypotheses regarding the nature of decoupling. As an important aspect, there is agreement that many cognitive processes operate in parallel and independent from each other (for a more controversial example, see Meyer & Kieras, 1997a).

Combining parallel processing with the assumption of free attention allocation predicts that different parallel cognitive processes may be independently decoupled from perceptual information. For example, it has been argued that lexical-semantic versus syntactic information processing proceed partly independently from each other (Friederici, 1999; Friederici, Opitz, & von Cramon, 2000). It is an interesting possibility that during an episode of mind wandering either one of these processes may be decoupled from perceptual information (e.g., lexical-semantic decoupling), while the other process remains coupled (e.g., syntactic coupling). Free allocation of attentional resources to parallel cognitive processes may thus predict the existence of different types of attentional decoupling.

Importantly, the present findings (Chapter 4) suggest that attentional decoupling may involve free and continuous allocation of cognitive resources. However, some cognitive processes, like access to a global workspace of consciousness, have been postulated to occur in an all-or-none fashion (Baars, 1988; Dehaene & Changeux, 2005; Dehaene, Kerszberg, & Changeux, 1998; Smallwood, 2010b). Such capacity limitations in the human brain and mind may provide a foundation for a dichotomy between conscious stimulus-dependent and conscious stimulus-independent thought (Smallwood, et al., 2011), and the present results from chapter 5 are consistent with this view.

In summary, the present work aimed at contributing to an empirical and theoretical understanding of the processes involved in mindless reading. We introduced and tested the levels-of-inattention hypothesis (Chapters 4 and 5) and found support for predictions from the cascade model of inattention (Chapter 4) and from the decoupling hypothesis (Chapter 5). Simulations of a mathematical model of eye movement control suggested specific cognitive mechanisms that may mediate attentional decoupling of eye movements during mindless reading (Chapter 3). I hope that these studies can contribute to the understanding of mindless reading and of the general phenomena of mind wandering and decoupling.

### 6.3 Cognition in eye movement control

How eye movements are coupled to higher-level cognitive processing via the eye-mind link is and has been debated in theories and models of eye movement control (for short review see the Introduction, Section 1.3.2.1). During mindless reading higher-level cognitive processing is decoupled from the text, and therefore studying eye movements during mindless reading provides an interesting opportunity to inform theories and models on the eye-mind link. In the present work, my colleagues and I investigated two cognitive influences in eye movement control during reading. First, we studied decoupling of eye movements from cognitive text processing in mindless reading paradigms (Chapters 2-4), and I will here discuss central results in the light of previous theories and findings (Section 6.3.1). Second, we investigated the foveal load hypothesis and the zoom lens of attention in a mathematical eye movement model (Chapters 2+3). Here, I will shortly describe an alternative explanation of the foveal load hypothesis based on serial attention allocation (Reichle, et al., 1998), and derive predictions to test both explanations in future research (Section 6.3.2).

### 6.3.1 Mindless reading and the eye-mind link

An important difference between current theories (and models) of eye movement control concerns the question whether higher-level cognition is the driving force generating regular sequences of eye movements during tasks like reading (Starr & Rayner, 2001; also see Section 1.3.2.1). To approach this question empirically, previous studies have approximated mindless reading via non-reading tasks. These studies have elicited an important debate about what conclusions can be drawn from the experimental results. However, whether cognition is needed to generate sequences of eye movements during tasks like reading has not been clear from these previous findings (Inhoff et al., 1993; Nuthmann & Engbert, 2009; Nuthmann et al., 2007; Rayner & Fischer, 1996; Reichle et al., 2012; Vitu et al., 1995). In the present work (Chapter 4), I discuss a new experimental approach to this question that is based on catching episodes of spontaneous mind wandering and decoupling during normal reading.

#### *6.3.1.1 Mind wandering during reading: Absent frequency effects*

During normal reading, word frequency is known to influence fixation durations (Inhoff & Rayner, 1986; Rayner, 2009). Effects of word frequency were found to be absent or reduced during non-reading tasks like target word search (Rayner & Fischer, 1996) and we here found reduced frequency effects during shuffled text reading (Chapter 2). Whether effects of lexical processing (i.e., cognitive influences) persist during episodes of spontaneous mind wandering and decoupling has long been unknown. As the first study on this question, Reichle et al. (2010) investigated eye movements during mindless reading using self-reports of mind wandering. In their study, they replicated word frequency effects on fixation durations for episodes of normal text reading. However, they argued that lexical and linguistic influences on eye movements (and the word frequency effect in particular) were absent during mindless reading.

In the present work (Chapter 4) we provided reliable empirical support for this conclusion. Consistent with previous research we found reliable lexical and linguistic influences on eye movements during normal reading. However, as a central empirical finding from the studies using the SAST, in states when attention was deeply decoupled from the text during deep mindless reading the word frequency effect (on gaze durations) was strongly reduced, and was numerically absent in follow-up analyses of very long target words (Chapter 4, Figure 4-6). (Of course, Bayesian analyses are desirable to test this null-hypothesis in future research.)

Together with the reports by Reichle et al. (2010) these findings suggest that lexical processing influences eye movements during normal reading, but that these influences are absent when external attention spontaneously and deeply lapses in a normal reading task.

#### *6.3.1.2 Is cognition needed for saccade generation? Support for autonomous control*

Theories and models of eye movement control generally agree that some saccades during reading are triggered by non-cognitive processes, for example when the eyes miss their intended target word (Nuthmann, et al., 2005, 2007). However, whether higher-level cognition is necessary to generate regular sequences of eye movements in tasks like reading has been controversial in theories of eye movement control. Cognitive trigger theory (implemented in direct control models; see Introduction, Section 1.3.2.1) suggests that higher-level cognition drives the eyes during tasks like reading. Cognitive-modulation theory (implemented in indirect control models) and POC theory, to the contrary, suggest that non-cognitive, autonomous processes move the eyes across the text.

The present finding that frequency effects are absent during deep attentional lapses strongly suggests that during deep mindless reading saccades are initiated independent of lexical word processing. If lexical processing drives the eyes through the text during normal reading, then absent frequency effects suggest that lexical processing is not the basis for a cognitive triggering process during deep mindless reading. Given that the eyes keep moving across the text even though lexical processing ceases, the question arises as to what other process may initiate saccades. Contrary to non-reading tasks like z-string scanning it seems unlikely that some internal high-level cognitive signal intentionally triggers saccades when the mind spontaneously wanders away from the text. Therefore, even if higher-level cognition should be crucial during normal text reading, the results on mind wandering suggest that some automatic low-level process initiates saccades in regular time intervals when the thoughts are wandering and lexical processing is decoupled from the external environment. The results on deep mindless reading therefore suggest that higher-level cognition may not be principally needed to move the eyes. Instead, an autonomous program seems to be running even when higher-level cognitive influences cease.

Based on this new support for autonomous saccade generation, I will here discuss autonomous control and theories of the eye-mind link (Section 6.3.1.3), transitions between cognitive and autonomous control (Section 6.3.1.4), the average speed of the autonomous timing process (Section 6.3.1.5), mechanisms (Section 6.3.1.6) and origins (Section 6.3.1.7) of autonomous control.

#### *6.3.1.3 Theories of the eye-mind link*

The findings on absent lexical effects during attentional lapses are informative for theories of the eye-mind link (for classification of theories see the Introduction, Section 1.3.2.1; for a summary or results see Table 6-1).

##### Primary oculomotor control (POC) theory

POC theories propose that low-level visuomotor factors are the main driving force moving the eyes across the text, but that high-level cognition has little local influence on eye movements (Reilly & O'Regan, 1998; Yang, 2006; Yang & McConkie, 2001). POC theories are well compatible with our present finding that lexical influences on eye movements are absent during deep levels of mindless reading (Chapter 4), and may successfully explain eye movement control during deep mindless reading. However, the present findings also show strong differences in eye movements between mindless and normal reading, and this supports prominent influences from cognitive processing. I conclude that POC theory may provide a strong baseline against which theories of cognitive control can be evaluated.

### Cognitive modulation theory

The findings on lexical decoupling during deep mindless reading are well in line with cognitive modulation theory (implemented in indirect cognitive control models). Cognitive modulation theory postulates that an "autonomous saccade generator can induce the start of a new saccade program in the absence of lexical processing demands" (Engbert, et al., 2005, p. 792), and that – for normal reading – the progression of this autonomous timing process is modulated by cognitive processing. For episodes of task-focus, cognitive modulation theory therefore predicts that lexical processing affects fixation durations, and the findings on normal reading support this prediction. For episodes of mindless reading, to the contrary, the autonomous timer predicts that the eyes keep moving across the text even when lexical processing ceases, and that lexical influences are reduced or absent. Again, as discussed above, this important prediction was supported by the empirical findings.

### Cognitive trigger theory

Cognitive trigger theory (which provides a mechanism for direct cognitive control models) assumes that a cognitive trigger initiates saccades during reading. This cognitive triggering process is usually assumed to be an early stage of lexical word processing (Reichle, et al., 1998). The present findings suggest that cognitive trigger theory can successfully explain eye movements during episodes of normal reading, where word frequency strongly affects fixation durations. However, absent frequency effects during attentional lapses suggest that lexical processing is not the basis for a cognitive triggering process during deep mindless reading. Moreover, contrary to non-reading tasks like z-string scanning (cf. Section 1.3.2.1) it seems unlikely that some internal high-level cognitive signal intentionally triggers saccades when the mind spontaneously wanders away from the text. Therefore, pure cognitive trigger theory has difficulties to plausibly explain why eye movements are initiated when lexical processing ceases to influence eye movements during deep levels of decoupling (cf. Chapter 4; and Reichle, et al., 2010).

### Dual-trigger theory: Cognitive trigger and autonomous timer

To explain eye movements during mindless reading, cognitive trigger theory may therefore supplement the cognitive trigger with some mechanism initiating saccades at a lower and more automatic level (for possible mechanisms see below, Section 6.3.1.7). The resulting dual-trigger theory therefore assumes that the mechanisms generating saccades during mindless reading differ from those that are in place during normal reading.

**Table 6-1**

Compatibility of theories of the eye-mind link with findings on frequency effects during normal reading and mindless reading.

| | **Pure Oculomotor Control Theory**<br>No cognitive control | **Cognitive Modulation Theory**<br>Indirect cognitive control | **Pure Cognitive Trigger Theory**<br>Direct cognitive control | **Dual Trigger Theory**<br>Direct + no cognitive control |
|---|---|---|---|---|
| **Normal Reading**<br>Frequency effects: ✔ | ✗ | ✔ | ✔ | ✔ |
| **Mindless Reading**<br>Frequency effects: ✗ | ✔ | ✔ | ✗ | ✔ |

### 6.3.1.4 *Transitions between cognitive and autonomous control*

What processes drive the eyes through the text during reading has been a much debated question in the last decades (see Section 1.3.2.1). The previous debate has often focused on the question *whether* higher-level *or* lower-level factors are more important in controlling eye movements during reading (e.g., Starr & Rayner, 2001). However, the mind often wanders during reading (Schooler, et al., 2004), and the present findings (Reichle, et al., 2010, and Chapter 4) suggest that eye movements are sometimes under cognitive control (during normal reading), but sometimes lack cognitive influences (during deep mindless reading). Thus, processes of eye movement control seem to fluctuate during the reading of a single text and to differ between episodes of on-task and off-task thinking. This finding highlights the question of how control switches between levels of control. Here I argue that theories and models on the eye-mind link need to provide mechanisms to explain how such switches occur. Based on my previous analysis (see Table 6-1), I suggest that two different classes of theories can explain frequently occurring transitions between normal and mindless reading.

It seems that models implementing *dual trigger theory* need to assume discrete switches between a direct cognitive trigger and autonomous saccade initiation to explain transitions between mindful and mindless reading. That is, these models seem to predict that an individual saccade is either under cognitive or under autonomous control (Reichle, et al., 2012). How such switches occur and what processes prevent autonomous control when lexical processing is present but slow needs to be specified in dual trigger models.

*Cognitive modulation theory*, to the contrary, assumes that the coupling between the autonomous timer and cognitive processing is gradually tuned between pure autonomous control during deep mindless reading and strong cognitive modulation when attention is fully directed to the reading task. Thus, cognitive modulation theory postulates that the same (autonomous) process initiates saccades during normal and mindless reading.

#### 6.3.1.5 Findings on autonomous control: average fixation durations

Finding support for autonomous control of eye movements during deep mindless reading poses the important question of how autonomous eye movements can be characterized. An important question concerns the average speed at which an autonomous program generates saccades during mindless reading, and the few existing studies have found diverging results. Reichle et al. (2010) reported that average fixation durations were prolonged during episodes of mind wandering as compared to on-task periods. Contrary to this result, we (Chapter 4) found that global patterns of eye movements, like average fixation durations or average fixation probabilities, were largely unaffected during mindless reading. Moreover, two other studies have investigated eye movements during mindless reading using a probe-caught self-report measure, and both of these studies found no effect of mindless reading on average fixation durations (Smilek, et al., 2010b; Uzzaman & Joordens, 2011). At present, it is unclear what causes the inconsistency between these results, and I here discuss a few potential causes.

In principle, it could be that differences in experimental power or failures to measure mindless reading explain experimental null-effects. However, this explanation is unlikely to account for the current results: in our present study (Chapter 4) we analyzed a much larger dataset compared to Reichle et al. (2010) (Reichle et al. tested four subjects over repeated sessions) and we obtained strong support for decoupled attention. Thus, other factors should exist that explain the differences between studies.

First, it could be that deep but not weak levels of mindless reading are associated with increases in average fixation durations. Consistent with this view, subjects in the study by Reichle et al. (2010) spent ten sessions in the lab reading, and this may have lead to deeper levels of mindless reading compared to the other studies, which tested subjects only during one session. In the present work (Chapter 4), we did not carefully investigate this hypothesis. However, preliminary analyses of the data from the SAST (i.e., from the experiments reported in Chapter 4) seem to support this view.[1] Thus, it could be that prolonged average fixation durations during mindless reading result from deep levels of decoupling.

That average fixation durations may be prolonged during deep decoupling is also interesting for an additional reason: it is a possibility that overlooking errors in the SAST may not only measure attentional decoupling, but may be confounded with speed-accuracy tradeoffs. That is, readers may overlook errors not because their mind is absent, but because they are quickly skimming across the page without carefully processing the text. Note that we took measures to control for such effects in our analysis, and our results suggested that these attempts were successful as average fixation durations were not reduced prior to overlooking errors (see Chapter 4). Moreover, the preliminary analyses reported in Footnote 1 suggest that deep levels of mindless reading – measured via overlooking of lexical errors in Experiment 2 – tended to be associated with prolonged average fixation durations, supporting the view that attentional lapses rather than skimming caused the overlooking of these errors.

Second, prolonged fixation durations during mindless reading may not result from mind wandering per se, but may be related to the cognitive processes mediating awareness of TUT. Evidence for this view comes from Reichle et al.'s (2010) study,

---

[1] In the analyses reported in Chapter 4 in Figure 4-5, average fixation durations (measured on non-final words) were numerically prolonged before low-level errors were overlooked (difference between mindless and mindful reading: 37.9 >= *b*s >= 13.7), and this effect was significant for the 10-word interval (*b* = 37.9, *SE* = 18.2, *t* = 2.08). When medium-level or high-level errors were overlooked, to the contrary, this effect was much smaller (difference between mindless and mindful reading: 11.5 >= *b*s >= -1.2) and non-significant (|*t*s| < 1.5) despite the larger amount of available data. Moreover, in Experiment 2 (reflecting deeper levels of mindless reading) average fixation durations (on non-final words) were marginally significantly prolonged during mindless reading compared to the control condition (see Chapter 4, Supplementary Information, Table F-5).

where prolonged fixation durations during mindless reading were most pronounced for the self-caught measure of mind wandering, which is generally viewed as indicating (meta-)awareness of mind wandering (Schooler, 2002; Schooler, et al., 2011). However, the effect was less reliable for their probe-caught measure of mind wandering, and was not replicated in other research using probe-caught self-report (Smilek, et al., 2010b; Uzzaman & Joordens, 2011) or behavioral measures (SAST, Chapter 4).

A speculative explanation of these differences in results may be that self-caught self-report measures of mind wandering partially reflects the result of a temporally extended decision process, where participants integrate information about their internal thoughts over an extended period of time.[2] When readers catch themselves mind wandering, it seems plausible that before they press a button to indicate their mind wandering, they may monitor their own mind for the occurrence of TUT over an extended period of time. Only when subjects find convincing (and maybe temporally extended or repeated) evidence for the occurrence of TUT in their mind, then they may press the button to indicate a self-caught episode of mind wandering. It seems like an interesting hypothesis that judgments about one's mind wandering state may be based on temporal integration of information from a self-monitoring process, and that this integration process slows saccade generation.

Such an interpretation may also be consistent with the observation that Reichle et al. (2010) reported prolonged fixation durations for longer time intervals (ranging from 10-120 seconds) preceding reports of mind wandering, but that the effect was (statistically and numerically) absent for short time intervals [e.g., 2.5 and 5 sec prior to report; see (Reichle, et al., 2010, Fig. 2)]. This null-effect is consistent with our findings (we analyzed intervals ranging from 10 to 20 words preceding the error sentence, and these were usually read in less than 10 seconds) and was also replicated in the two other existing studies on eye movements during mindless reading, which analyzed an interval of 5 seconds before self-reports of mindless reading (Smilek, et al., 2010b; Uzzaman & Joordens, 2011). As a speculative possibility, the temporal decision process, which may cause prolonged fixation durations during self-caught mind wandering, may be determined a few seconds before the actual response occurs.

---

[2] Note, however, that participants in mind wandering experiments are often instructed to report whether they were mind wandering in the moment immediately preceding the probe, but not during any time since the last probe.

In summary, more research is needed to better understand the conditions under which average fixation durations are prolonged during mindless reading and to clarify inconsistencies between studies by better understanding the relation between self-caught and probe-caught self-report measures of TUT and the SAST as a behavioral measure of different levels of attentional decoupling.

#### *6.3.1.6 Mechanisms of autonomous control*

Empirical support for autonomous saccade initiation also raises the question for the specific mechanisms that may initiate saccades during mindless reading. Here, I discuss some possible mechanisms with respect to theories of eye movement control.

##### Dual trigger theory: cognitive trigger and autonomous timer

As an example of a model assuming a cognitive saccade trigger, Reichle et al. (2012) suggested that when lexical processing fails during mindless reading, a saccade ’deadline’ initiates saccades at a time „corresponding to the maximal time that is normally allotted for lexical processing“ (Reichle, et al., 2012, p. 39) (also see Henderson & Ferreira, 1990). This explanation has been implemented in the E-Z Reader model to explain eye movements in the z-string scanning paradigm (Reichle, et al., 2012) and predicts the finding of prolonged average fixation durations during mindless reading (Reichle, et al., 2010). At present, however, it is unclear whether the predictions from a saccade ’deadline’ match the observed effect sizes at a quantitative level. Specifically, the saccade ’deadline’ predicts that average fixation durations during mindless reading are as long as normal fixations on very low-frequency words. Our findings in the present work, however, did not support this prediction. We found (cf. Chapter 4, Figure 4-3 and Figure 4-6) that on very low-frequency target words, average gaze durations during deep mindless reading were clearly shorter compared to those observed during normal reading. Thus, alternative mechanisms for automatic low-level control may be considered to supplement cognitive trigger theory.

##### Cognitive modulation theory

Cognitive modulation theory assumes that the same autonomous timer initiates saccades during both normal and mindless reading, and that cognitive processing only modulates this timing process during normal reading. This theory predicts that global aspects of eye movements should be very similar between normal and mindless reading, but that cognitive influences on eye movements should be reduced during mindless reading. The current findings from the SAST (see Chapter 4) are very consistent with this prediction: we found that average fixation durations and probabilities were highly similar between normal and mindless reading, but that during mindless reading influences from lexical and linguistic variables were strongly reduced.

### Simulating the SWIFT model

Ralf Engbert and I performed simulations of the SWIFT model to explain eye movements for normal and shuffled text reading (Chapter 3). The results from these simulations suggest specific mechanisms of how eye movements may be decoupled from the text during mind wandering. First, the results suggest that during normal text reading lexical processing strongly and dynamically influences a word-based saliency map (i.e., activations of words in an activation field), while during mindless reading visual saliency may be independent of cognitive processing. Second, the results suggest that during normal reading high foveal word activations inhibit the autonomous saccade timer via foveal inhibition, but that this foveal inhibition may be inactive during mindless reading. These simulations suggest specific processes by which eye movements may be decoupled from cognitive processing when the mind wanders away from the text. Clearly, future model simulations are desirable to test these predictions in paradigms like that SAST that catch episodes of decoupling during normal reading.

### Foveal inhibition versus foveal facilitation

For cognitive modulation theory, in principle cognitive processing may modulate the autonomous saccade timer in two ways: cognitive processing could either inhibit the random timer in the presence of local processing difficulties (e.g., for difficult words) to prolong current fixation durations, or it could speed the random timer when processing is easy in order to shorten the current fixation. While both of these mechanisms seem plausible and efficient from an adaptive control perspective, the SWIFT model (Engbert, et al., 2002; Engbert, et al., 2005; also see Chapter 3), as an important cognitive modulation model, only implements the inhibitory mechanism via a process called 'foveal inhibition'. Thus, in SWIFT local processing difficulties cannot directly speed the progression of the random timer. The present data on reduced frequency effects during mindless reading (Chapter 4, Figure 4-6), however, suggest that cognition may have both consequences: it can (a) slow gaze durations for very difficult words and also (b) fasten gaze durations for easy words. The hypothesis that cognition may fasten fixation durations is also supported by the findings on prolonged average fixation durations during self-caught mindless reading (Reichle, et al., 2010), and potentially during deep mindless reading (see above Section 6.3.1.5). Such results may suggest that cognitive modulation models like SWIFT should not only consider inhibitory influences from cognitive processing, but that cognition may also speed the progression of the autonomous timer. However, other mechanisms in the SWIFT model may potentially explain experimental results. Therefore, additional analyses of experimental data as well as mathematical simulations are needed before further conclusions can be drawn.

### Levels of decoupling

In the present work we provided support for the levels of inattention hypothesis that different levels of cognitive processing are gradually decoupled during different levels of mindless reading (see Chapter 4). Based on these findings, it would be interesting to investigate different levels of decoupling using mathematical models that implement eye movement control at different processing levels. For example, the E-Z Reader model (E-Z Reader 9; Reichle, Warren, et al., 2009) includes explicit assumptions about influences from higher-level language processing in addition to lexical and low-level visuomotor factors. Likewise, the SWIFT model (Engbert, et al., 2005) makes assumptions about higher-level predictability, medium-level lexical, and low-level visuomotor processes. Moreover, work is currently under way to develop or extend existing models to incorporate language processing at higher levels (e.g., Engelmann & Vasishth, 2011). The present findings on different levels of decoupling may provide interesting constraints for these models.

### The role of perceptual processing during autonomous control

Based on the present support for autonomous saccade initiation, an important question concerns whether this autonomous triggering process depends on perceptual information. As one possibility, as suggested by cognitive modulation theory, an autonomous motor program may initiate eye movements independent of visual input. Alternatively, automatic saccade initiation may depend on some kind of low-level visual stimulus processing that is intact even when lexical processing is decoupled.

One way to address this question can be to study influences from low-level visual variables on eye movements during mindless reading. The presently available data on eye movements during deep mindless reading shows that the effect of the visual variable word length on fixation durations was reduced during mindless reading (Reichle, et al., 2010, and Chapter 4). This finding seems compatible with an autonomous timing process that initiates saccades independent of perceptual information. However, further research is needed to clearly test a potential role of visual processing. Such research may further investigate low-level visual influences on eye movements during deep mindless reading, and may test whether processing can be decoupled at even deeper levels (e.g., orthographic decoupling, or decoupling from word boundaries).

#### 6.3.1.7 Origins of autonomous control

Current evidence in support of autonomous low-level saccade initiation during deep mindless reading raises the question for the general nature and origin of this process. Here, I suggest that three possible hypotheses exist about why saccades may be initiated by an autonomous timer in the absence of cognitive processing.

*Global cognitive control*: It may be that autonomous saccade initiation results from a specific task-set that the system employs during reading. According to this view, although saccade initiation is locally independent of cognitive processing during mindless reading, it may be subject to higher-level cognitive settings at a global level. For example, the eye movement control system may strategically define a saccade deadline (Reichle, et al., 2012), initiate a random saccade timer (Engbert & Kliegl, 2001), or pre-define reading-specific 'strategies and tactics' (Reilly & O'Regan, 1998) prior to reading a certain passage of text. A repetitive transcranial magnetic stimulation (rTMS) study (Leff, Scott, Rothwell, & Wise, 2001) suggested that such parameter settings for a specific sensorimotor scanpath may be invoked by the frontal eye fields prior to the reading of an individual line of text. Importantly, the global cognitive control hypothesis suggests that cognition globally sets parameters for a specific eye movement task, and that the settings that are usually invoked for reading lead to an autonomous initiation of saccades when the mind wanders.

*Habit*: It may not be very surprising that the eyes keep moving across the text when lexical processing ceases during mindless reading as a lifelong history of reading experience may generate a habit of constantly moving the eyes across the text. Thus, autonomous saccade initiation during deep mindless reading may not be voluntarily accessible to (global) cognitive settings, but may automatically result from extensive practice in reading. Interestingly, the habit-hypothesis provides further constraints for the nature of the autonomous saccade trigger: To acquire automaticity that outlasts deep failures of external attention, the process initiating saccades in the absence of cognitive processing (i.e., an autonomous timer) should be (passively or actively) involved in the initiation of saccades during normal reading. From the perspective of dual-trigger theory, the habit-hypothesis therefore predicts that the autonomous timer should be located downstream from a potential cognitive triggering mechanism at a low oculomotor level. Moreover, for beginning readers, who have not yet spent enough time to develop a habit, the habit-hypothesis predicts that the eyes may stop moving when attention is deeply disengaged from the external reading task.

*Biologically fixed trigger*: It may be that autonomous saccade generation is a biologically fixed process, which regularly initiates new saccades at a constant rate whenever cognitive control does not inhibit or replace the biological trigger. Such a biological time signal should be active in any task, and would also be engaged during mindless reading. The biologically fixed trigger hypothesis seems to be compatible with the observation that microsaccades occur during visual fixation tasks (Rolfs, 2009). These micro-movements may occur when a biological trigger initiates a new saccade program, which is subsequently inhibited by a cognitive fixation-signal. It seems like an interesting prediction from this hypothesis, that when the mind is deeply wandering during the execution of a fixation task, then this should lead to the generation of saccades because inhibition of the biological trigger fails.

##### 6.3.1.8 Conclusion

It is sometimes assumed that lexical processing is “the ’engine’ driving eye movements during reading” (Reichle, et al., 2003, p. 459). While this analysis provides and interesting account for episodes of normal reading, evidence suggests that cognition may not be needed to initiate saccades during mindless reading. The finding that lexical influences are absent during deep levels of mindless reading supports the existence of an autonomous mechanism for saccade initiation that is independent of cognitive processing. Moreover, observed differences between mindless and normal reading clearly support an important role of cognitive processing for eye movement control during normal reading, and this provides support for models of cognitive control. The findings on coupled processing during normal reading and decoupled processing during mindless reading highlight the flexibility of the eye-mind link when external attention spontaneously fluctuates during performance of a reading task. The specific mechanisms by which cognitive processing controls eye movements during normal reading, however, are not clear from the present findings. Whether two engines exist (cognitive and autonomous) to generate saccades during normal and during mindless reading, or whether a single autonomous engine is running and differentially coupled to cognitive processing is an open question for future research.

### 6.3.2 SWIFT 3: The zoom lens of attention and the foveal load hypothesis

In the present work my colleagues and I also studied how cognitive processing affects eye movements by modulating the zoom lens of visual attention (Eriksen & St. James, 1986). We derived qualitative predictions from the zoom lens for the reading of normal and randomly shuffled text (Chapter 2). Moreover, we introduced an advanced version of the SWIFT model (SWIFT 3) incorporating the zoom lens of attention to test these predictions in a fully quantitative eye movement model (Chapter 3).

Importantly, in our simulations we investigated how the zoom lens of attention impacts on eye movements during reading. We found that the zoom lens contributes to explain effects of current word frequency and of current word length, and provides an account for the observed dissociation between immediate and distributed processing effects in shuffled text reading. Moreover, we found that the zoom lens of attention contributed to the first mathematical explanation of skipping benefits. These results demonstrate that the zoom lens of attention, implemented in the SWIFT model of saccade generation, is an important concept to explain eye movement control (Chapter 3).

#### Serial versus parallel word processing: Tests based on the zoom lens of attention

The zoom lens of attention also provides an important concept to derive experimental tests for the question whether word processing during reading proceeds in a serial or a parallel fashion. In the following, I will derive several qualitative predictions from the zoom lens in processing gradient (PG) models like SWIFT (see Chapter 3) and contrast these with predictions from a competing account based on sequential attention shift (SAS) models like E-Z Reader (Reichle, et al., 1998).

An influential discovery in eye movement research has been that foveal processing difficulties reduce parafoveal processing during reading, and this effect was first described by Henderson and Ferreira (1990) in their foveal load hypothesis (see also Kennison & Clifton, 1995; Schroyens, et al., 1999; White, et al., 2005). To investigate foveal load effects on the perceptual span, Henderson and Ferreira (1990) varied foveal processing difficulty in two eye tracking experiments (by manipulating the fixated word's frequency or syntactic ambiguity) and found a preview benefit effect (in the boundary paradigm, Rayner, 1975) for the low foveal load conditions, but not for the high load conditions. This finding had important theoretical consequences because it was inconsistent with the then dominant Morrison (1984) model of eye movement control, which predicts that preview is independent of foveal load. Today, the effect of foveal load on the perceptual span (Henderson & Ferreira, 1990) is a benchmark result to evaluate mathematical models of eye movement control during reading.

Accounting for foveal load effects on the perceptual span (Henderson & Ferreira, 1990) was an important goal and accomplishment in the development of the E-Z Reader model (Reichle, et al., 1998; Reichle, et al., 2012; Reichle, et al., 2003), as the most prominent sequential attention shift (SAS) model. To explain effects of foveal load on parafoveal processing, the E-Z Reader model assumes that shifts of covert attention are decoupled from the programming of eye movements (Reichle, et al., 1998). The model proposes an early stage of lexical processing (L1, called the 'familiarity check'). When L1 is completed for the attended word, a saccade program is triggered, which initiates a saccade to the next word *n*+1 after a fixed amount of programming time. Moreover, completion of L1 initiates a second stage of lexical processing (L2, 'lexical access'), and after completion of L2 attention shifts to the next word *n*+1. In this model, preview occurs whenever attention shifts to the next word before the saccade program is executed. Importantly, the amount of preview that is obtained for word *n*+1 depends on the duration of L2 relative to the time needed to program the saccade (Reichle, et al., 2003, Figure 4). Because (a) the saccade programming time is independent of lexical processing and (b) L2 is a function of a words' frequency, for low-frequency words attention usually shifts to word *n*+1 only shortly before the eyes move to word *n*+1, and little preview benefit is obtained. For high-frequency words, to the contrary, attention often shifts much earlier, and substantial preview for word *n*+1 can be obtained.

The SWIFT 3 model, which we introduced in Chapter 3, provides an alternative account of the foveal load hypothesis based on the zoom lens of attention (Eriksen & St. James, 1986). In SWIFT 3, attention is allocated to a spatially extended region of the text to support parallel processing of several words at a time. The focus of this attention gradient, i.e., the focus of the zoom lens, depends on foveal word activations. Based on this model, if the foveal word has not yet been processed or is highly activated, the zoom lens is focused and word *n*+1 likely falls outside of the processing gradient. When word activation decreases the zoom lens is defocused and attention is distributed to support processing of upcoming words. Because word activation is influenced by a words' processing difficulty in SWIFT 3, the model qualitatively predicts a large preview benefit for easy words *n* and a small or no preview benefit for difficult words *n*. It would be interesting to test this prediction quantitatively.

It is an interesting task for future research to test (qualitative and quantitative) predictions from these two alternative accounts of the foveal load hypothesis experimentally. First, a critical difference between sequential attention shift (SAS) models like E-Z Reader (Reichle, et al., 1998) and processing gradient (PG) models like SWIFT (Chapter 3) concerns their predictions for parafoveal-on-foveal (PoF, or successor) effects, that is, for influences from upcoming words (e.g., word *n*+1) on fixations on the fixated word *n*. Previous reports about PoF effects (e.g., Hohenstein, Laubrock, & Kliegl, 2010; Kliegl, 2007; Kliegl, et al., 2006; Kliegl, et al., 2007; Risse & Kliegl, 2011, 2012) have been debated (Rayner, 2009). In the present thesis, I provided experimental evidence for reliable and valid PoF effects in shuffled text reading (Chapter 2).

While the general spirit of the serial processing assumption in SAS models does not seem to predict such spatially distributed effects, SAS models can explain PoF effects via 'mislocated fixations', i.e., the widely held assumption that saccades sometimes miss their intended target word due to oculomotor error in saccade targeting (e.g., Drieghe, et al., 2008). An assumption in SAS models is that when a fixation is mislocated, attention is directed to the intended target word, but the eyes fixate on a neighboring word. Via this mechanism processing of neighboring words can influence fixation durations on the fixated word under the assumption of serial word processing. PG models, to the contrary, explain PoF effects by assuming parallel processing of words in the processing span (Kennedy & Pynte, 2005; Kliegl, et al., 2006).

Critically, these explanations fundamentally differ in their predictions for effects of foveal load. PG models predict that based on the zoom lens parafoveal processing should be modulated by foveal processing difficulties. Therefore PoF effects should be more pronounced or likely for easy than for difficult fixated words. To the contrary, mislocated fixations in the SAS account are not influenced by the difficulty of the foveal word (Reichle, et al., 2012), and therefore PoF effects should not vary with foveal load. Some previous studies have failed to find an interaction between foveal load and PoF effects (Henderson & Ferreira, 1990; Kliegl, et al., 2006). However, these null-effects were either based on corpus analyses with reduced experimental control, or lacked statistical power due to small sample sizes. Moreover, some limited evidence suggests that PoF effects may depend on foveal processing difficulties. This was suggested by a recent study showing that PoF effects were reduced under working memory load (Gendt, 2012). In this study, however, foveal load was confounded with parafoveal load, and future research is needed to disentangle these influences.

Based on the load theory of attentional selection (Lavie, 2005), loading low-level perceptual resources, but not loading more central (high-level) resources reduces low-level processing of distractor items. Applying this general idea to the interaction of foveal load and PoF effects predicts that foveal load at lower levels may be particularly efficient in modulating (i.e., reducing) parafoveal processing and PoF effects. To my knowledge, this hypothesis has not been directly tested in previous research. However, somewhat consistent with this view, preliminary corpus analyses showed that when foveal load is measured via lemma frequency (i.e., the frequency of the dictionary form of a word, e.g., ‘laugh’ in ‘laughing’, which may capture early lexical processing) instead of type frequency (i.e., the frequency of the specific word, e.g., the frequency of ‘laughing’ itself, which may reflect later lexical processing stages) then foveal load interacts with lexical successor effects (Kliegl, R., personal communication, 2010). Clearly, more research is desirable to investigate this possibility.

An additional interesting question about models concerns the level of parafoveal processing at which foveal load effects may be expected. At an early visual level, the E-Z Reader model assumes a separate pre-attentive visual stage, which processes low-spatial frequency information (e.g., word boundaries or word shape) for several words in parallel (Reichle, et al., 2003). Due to its pre-attentive nature, the E-Z Reader model does not assume that parafoveal processing of pre-lexical low-spatial frequency information is modulated by foveal processing difficulties. The SWIFT 3 model also implements an early preprocessing stage by assuming minimal preprocessing at a fixed number of 15 letters to the right of fixation. However, it is a plausible hypothesis that foveal load and the zoom lens of attention also modulate parafoveal processing at this early stage. Future research may explore and test such a possibility.

Moreover, SAS and PG accounts of foveal load effects differ with respect to the predictions they make for eye movements during non-reading tasks. Several models have been developed to explain eye movement control during reading. A few studies have now started to generalize these models to explain eye movements in non-reading tasks, like visual search, z-string scanning, or driving (Nuthmann & Engbert, 2009; Reichle, et al., 2012; Salvucci, 2001). Recently, Reichle et al. (2012) performed simulations of the E-Z Reader model for several non-reading tasks and argued that E-Z Reader provides a general framework to explain eye movements in both reading and non-reading tasks. Based on the model simulations Reichle et al. (2012) suggested that while eye movements and shifts of attention are partly decoupled in reading, this is not the case in non-reading tasks. Specifically, they postulate that saccade initiation can be decoupled from shifts of attention only for highly familiar stimuli and much training in a task. Thus, based on the E-Z Reader model, Reichle et al. (2012) make the strong prediction that foveal load does not reduce parafoveal processing in (unpracticed) non-reading tasks. Moreover, the assumed coupling between shits of attention and saccade initiation predicts that parafoveal preview should be substantially increased for novel tasks and stimuli.

In PG models like SWIFT, to the contrary, there seems to be no a priori reason why the zoom lens should not operate during unpracticed non-reading tasks. Instead, the zoom lens predicts that the foveal load hypothesis (Henderson & Ferreira, 1990) is valid also for unpracticed non-reading tasks and for unfamiliar stimuli. Moreover, the zoom lens in PG models predicts that parafoveal preview should decrease for novel tasks and stimuli, because increased foveal load should lead to a stronger focus of the zoom lens.

Thus, generalizing models of eye movement control during reading to non-reading tasks provides new and very interesting possibilities for model tests. It is not easy to clearly distinguish SAS and PG models in eye tracking experiments. Here, I suggest that model mechanisms to account for the foveal load hypothesis, namely the zoom lens and the decoupling of saccade initiation from shifts of attention, may provide some clear experimental tests of the underlying theoretical assumptions.

## 6.4 Conclusions

During reading, many peoples' minds occasionally drift off the text to completely unrelated thoughts and feelings so that they suddenly find themselves someplace in the text without any idea of how they got there. During such episodes of mindless reading the eyes move across the page while attention is decoupled from the text. Although the experience of mindless reading is ubiquitous and evident in subjective experience, it has long escaped vigorous scientific investigation, and has only recently received increased attention in cognitive research.

The present work studied eye movements during mindless reading as a window to what happens in the mind when it wanders. We introduced two new paradigms to investigate the elusive phenomenon of mindless reading. The results demonstrate that the phenomenon can be vigorously investigated in the scientific laboratory. Interestingly, we found that eye tracking can be used to predict states of mindless reading online. This prediction was possible because cognition normally plays an important role in guiding the eyes during reading. For mindless reading, however, the present results suggest that these cognitive influences are absent. To more closely understand the mechanisms that may lie behind the decoupling of eye movements from the text during mindless reading, we performed simulations of a mathematical model of saccade generation (SWIFT 3). We found support for the hypothesis that cognition is not always necessary to generate saccades. Instead, the present results suggest that an autonomous motor timer initiates saccades in the absence of cognitive processing.

Although mind wandering is an elusive experience in the human mind, the present studies suggest that it can be successfully tracked down and understood using tools (like eye tracking, signal detection analysis, and mathematical modeling) and central theoretical concepts (like hierarchical levels of processing, attentional selection, or resource competition) from mainstream cognitive science. We applied these tools and concepts to study mindless reading. The results suggest that reduced external attention during mindless reading is a graded process that involves attentional selection at different early or late stages. Consistent with previous theorizing, however, internal thought may reflect a distinct functional mode that operates in an all-or-none fashion. I conclude that the phenomenon of mindless reading can be incorporated into the theoretical framework of cognitive science, that it informs theories of the mind, and that the steady increase in knowledge and methodology in the study of mindless reading may set the stage for a previously unimagined and detailed cognitive understanding of the elusive phenomenon.

## References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: testing and interpreting interactions*. RR Reno: Sage Publications, Inc.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*(4), 477-492.

Anderson, J. R. (2005). *Cognitive Psychology and Its Implications: Sixth Edition* (6th ed.). New York: Worth Publishing.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (1999). Role of locus coeruleus in attention and behavioral flexibility. *Biological Psychiatry, 46*(9), 1309-1320.

Baars, B. (1988). *A cognitive theory of consciousness*. Cambridge, England: Cambridge University Press.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word Length and Structure of Short-Term-Memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition, 20*(4), 1604-1611.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology-General, 133*(2), 283-316.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364-390.

Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in thought: the effect of mind wandering on the processing of relevant and irrelevant events. *Psychological Science, 22*(5), 596-601.

Bates, D., & Sakar, D. (2008). lme4: Linear mixed-effect models using S4 classes. (Version 0.999375-35).

Beatty, J. (1982a). Phasic not tonic pupillary responses vary with auditory vigilance performance. *Psychophysiology, 19*(2), 167-172.

Beatty, J. (1982b). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276-292.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*(9), 967-983.

Bellgrove, M. A., Hawi, Z., Gill, M., & Robertson, I. H. (2006). The cognitive genetics of attention deficit hyperactivity disorder (ADHD): sustained attention as a candidate phenotype. *Cortex, 42*(6), 838-845.

Bohning, D., Dietz, E., Schlattmann, P., Mendonca, L., & Kirchner, U. (1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society Series a-Statistics in Society, 162*, 195-209.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Oxford University Press.

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 125-147). Oxford, England: Elsevier.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*, 1-38.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: relative contribution of visual, lexical, and contextual factors. *Spanish Journal of Psychology, 5*(1), 66-77.

Cheyne, J. A., Carriere, J. S., Solman, G. J., & Smilek, D. (2011). Challenge and error: critical events and attention-related errors. *Cognition, 121*(3), 437-446.

Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition, 111*(1), 98-113.

Christoff, K. (2012). Undirected thought: neural determinants and correlates. *Brain Research, 1428*, 51-59.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences of the United States of America, 106*(21), 8719-8724.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73-101.

Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. van_Gompel (Ed.), *Eye movements: A window on mind and brain* (pp. 341-372). Amsterdam: Elsevier.

Cohen, G. (2000). Hierarchical models in cognition: Do they have psychological reality? *European Journal of Cognitive Psychology, 12*(1), 1-36.

Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: distorted subjective time during the dual-task bottleneck. *Psychological Science, 19*(11), 1110-1117.

Cornelissen, F. W., Peters, E., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments & Computers, 34*, 613-617.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42 - 45.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

D'Orazio, T., Leo, M., Guaragnella, C., & Distante, A. (2007). A visual approach for driver inattention detection. *Pattern Recognition, 40*(8), 2341-2355.

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research, 1084*, 89-103.

Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R., & Jacobs, A. M. (2009). Event-related potentials reveal rapid verification of predicted visual input. *Plos One, 4*(3).

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., Finnigan, S., & Humphreys, M. S. (2005). fMRI evidence of word frequency and strength effects in recognition memory. *Cognitive Brain Research, 24*(3), 587-598.

Dehaene, S., & Changeux, J. P. (2005). Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. *PLoS Biology, 3*(5), e141.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America, 95*(24), 14529-14534.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*(1-2), 1-37.

Deutsch, J. A., & Deutsch, D. (1963). Attention - some theoretical considerations. *Psychological Review, 70*(1), 80-90.

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology-Learning Memory and Cognition, 32*(4), 805-815.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating. *Behavioral and Brain Sciences, 11*(3), 357-374.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. Matin (Eds.), *Mechanisms of Attention: Attention and Performance XI* (pp. 171-187). Hillsdale, NJ: Erlbaum.

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: does an extra space make a difference? *Vision Research, 45*(13), 1693-1706.

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16*(1-2), 79-103.

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *The Quarterly Journal of Experimental Psychology, 61*(8), 1239-1249.

Driver, J. (2001). A selective review of selective attention research from the past century.

*British Journal of Psychology, 92 Part 1*, 53-78.

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences of the United States of America, 108*(19), 7716-7720.

EDAS II (Version 1.6.22). (2009). URL http://www.entroware.com/edas2.html.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*(6), 641-655.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America, 105*(5), 1704-1709.

Engbert, R. (2007). *Reading with a dynamic processing span*. Paper presented at the 14th European Conference on Eye Movements, ECEM.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: a possible role for autonomous saccades. *Biological Cybernetics, 85*(2), 77-87.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035-1045.

Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. D. Gilchrist & S. Everling (Eds.), *Oxford Handbook of Eye Movements* (pp. 787–800). Oxford, UK: Oxford University Press.

Engbert, R., & Krügel, A. (2010). Readers use bayesian estimation for eye movement control. *Psychological Science, 21*(3), 366-371.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*(5), 621-636.

Engbert, R., Longtin, A., & Kliegl, R. (2004). Complexity of eye movements in reading. *International Journal of Bifurcation and Chaos, 14*(2), 493-503.

Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America, 103*(18), 7192-7197.

Engbert, R., & Nuthmann, A. (2008). Self-consistent estimation of mislocated fixations during reading. *Plos One, 3*(2), e1534.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777-813.

Engelmann, F., & Vasishth, S. (2011). *Language processing and eye-movement models*. Paper presented at the ECEM.

Eriksen, C. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics, 12*(2B), 201-204.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention - a zoom lens model. *Perception & Psychophysics, 40*(4), 225-240.

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review, 109*(3), 545-572.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11-15.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22*(4), 661-721.

Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass: MIT Press.

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: the role of perceptual load. *Cognition, 111*(3), 345-355.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., et al. (2011). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 1-17.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America, 102*(27), 9673-9678.

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review, 18*(5), 992-997.

Frawley , W. J. (Ed.). (2003). *International Encyclopedia of Linguistics*. New York: Oxford University Press.

Friederici, A. D. (1999). The neurobiology of language comprehension. In A. D. Friederici (Ed.), *Language comprehension: a biological perspective* (pp. 263-301). Berlin: Springer.

Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cerebral Cortex, 10*(7), 698-705.

Friedman, M. (Ed.). (2004). *Kant: Metaphysical Foundations of Natural Science*. Cambridge, UK: Cambridge University Press.

Gazzaniga, M. S. (Ed.). (2009). *The Cognitive Neurosciences* (4th ed.): MIT Press.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2*(4), 1360-1383.

Gendt, A. (2012). *Eye movements under the control of working memory: The challenge of a reading span task.* Unpublished Ph.D. Thesis, University of Potsdam, Potsdam, Germany.

Geyken, A. (2006). The DWDS-Corpus: A reference corpus for the German language of the 20th century. In C. Fellbaum (Ed.), *Collocations and Idioms: Linguistic, Lexicographic, and Computational Aspects.* London: Continuum Press.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4*(1), 1-21.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry, 81*(25), 2340-2361.

Gillespie, D. T. (1978). Monte Carlo simulation of random walks with residence time dependent transition probability rates. *Journal of Computational Physics, 28*(3), 395-407.

Glass, L., & Mackey, M. C. (1988). *From clocks to chaos. The rhythms of life*. Princeton, NJ: Princeton University Press.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.

Graesser, A. C., Olde, B., & Klettke, B. (2002). How does the mind construct and represent stories? In M. C. Green, J. J. Strange & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 229-262). Mahwah, NJ: Erlbaum.

Grodsky, A., & Giambra, L. (1990). The consistency across vigilance and reading tasks of individual differences in the occurrence of task unrelated and task related images and words. *Imagination, Cognition and Personality, 10*, 39-52.

Hampson, R. E., Opris, I., & Deadwyler, S. A. (2010). Neural correlates of fast pupil dilation in nonhuman primates: Relation to behavioral performance and cognitive workload. *Behavioural Brain Research, 212*(1), 1-11.

Hawkes, C. H., & Stow, B. (1981). Pupil size and the pattern evoked visual response. *Journal of Neurology, Neurosurgery and Psychiatry, 44*(1), 90-91.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau, 62*(1), 10-20.

Henderson, J. M. (2012). Introduction to "Computational Approaches to Reading and Scene Perception". *Visual Cognition, 20*(4-5), 357-359.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading - Implications for attention and eye-movement control. *Journal of Experimental Psychology-Learning Memory and Cognition, 16*(3), 417-429.

Henderson, J. M., & Ferreira, F. (1993). Eye-movement control during reading - Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 47*(2), 201-221.

Hogaboam, T. W. (1983). Reading patterns in eye movements. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 309-332). New York: Academic Press.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology-Learning, Memory, and Cognition, 36*(5), 1150-1170.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers - effects of word-length and word-frequency. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*(6), 1430-1440.

Hyönä, J., Tommola, J., & Alaja, A. M. (1995). Pupil-dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 48*(3), 598-

612.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology-Human Perception and Performance, 31*(5), 979-995.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye-movements during reading. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 41*(1), 63-89.

Inhoff, A. W., Radach, R., & Eiter, B. (2006). Temporal overlap in the linguistic processing of successive words in reading: Reply to Pollatsek, Reichle, and Rayner (2006a). *Journal of Experimental Psychology-Human Perception and Performance, 32*(6), 1490-1495.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word-processing during eye fixations in reading - Effects of word-frequency. *Perception & Psychophysics, 40*(6), 431-439.

Inhoff, A. W., Topolski, R., Vitu, F., & O'Regan, J. K. (1993). Attention demands during reading and the occurrence of brief (express) fixations. *Perception & Psychophysics, 54*(6), 814-823.

Janisse, M. P. (1977). *Pupillometry: The psychology of the pupillary response.* Washington: Hemisphere Publishing.

Johansson, P., Hall, L., Sikstrom, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310*(5745), 116-119.

Johnson, K. A., Kelly, S. P., Bellgrove, M. A., Barry, E., Cox, M., Gill, M., et al. (2007). Response variability in attention deficit hyperactivity disorder: evidence for neuropsychological heterogeneity. *Neuropsychologia, 45*(4), 630-638.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading - From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought - Pupillometric indexes of sentence processing. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 47*(2), 310-339.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology-General, 111*(2), 228-238.

Kam, J. W., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., et al. (2011). Slow fluctuations in attentional control of sensory cortex. *Journal of Cognitive Neuroscience, 23*(2), 460-470.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when - An experience-sampling study of working memory and executive control in daily life. *Psychological Science, 18*(7), 614-621.

Kant, I. (1786). *Metaphysische Anfangsgründe der Naturwissenschaft*. Riga: Johann Friedrich Hartknoch.

Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology, 44*(1), 128-144.

Kennedy, A. (2008). Parafoveal-on-foveal effects are not an artifact of mislocated saccades. *Journal of Eye Movement Research, 2*(1), 1-10.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*(2), 153-168.

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working-memory capacity readers - Implications for eye-movement control. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*(1), 68-81.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*(6006), 932-932.

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

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology-General, 136*(3), 530-537.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review, 12*(1), 132-138.

Kliegl, R., & Engbert, R. (in press). Evaluating a Computational Model of Eye-Movement Control in Reading.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 262-284.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18*(5), 655-681.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology-General, 135*(1), 12-35.

Kliegl, R., Ping, W., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in linear mixed models: Estimating the relation of spatial, object, and attraction effects in visual attention. *Frontiers in Psychology, 2*, 12.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+2. *Journal of Experimental Psychology-Human Perception and Performance, 33*(5), 1250-1255.

Klinger, E., Gregoire, K. C., & Barta, S. G. (1973). Physiological correlates of mental activity: eye movements, alpha, and heart rate during imagining, suppression, concentration, search, and choice. *Psychophysiology, 10*(5), 471-477.

Krügel, A., & Engbert, R. (2010). On the launch-site effect for skipped words during reading. *Vision Research, 50*(16), 1532-1539.

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *Quarterly Journal of Experimental Psychology, 63*(9), 1838-1857.

LaBerge, D. (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology-Human Perception and Performance, 9*(3), 371-379.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*(1), 101-124.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7*(1), 12-18.

Laubrock, J., Kliegl, R., & Engbert, R. (2006). SWIFT explorations of age differences in eye movements during reading. *Neuroscience and Biobehavioral Reviews, 30*(6), 872-884.

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences, 9*(2), 75-82.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology-General, 133*(3), 339-354.

Leff, A. P., Scott, S. K., Rothwell, J. C., & Wise, R. J. S. (2001). The planning and guiding of reading saccades: a repetitive transcranial magnetic stimulation study. *Cerebral Cortex, 11*(10), 918-923.

Levinson, D. B., Smallwood, J., & Davidson, R. J. (2012). The persistence of thought: evidence for a role of working memory in the maintenance of task-unrelated thinking. *Psychological Science, 23*(4), 375-380.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6-14.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (Eds.). (2011). *Oxford Handbook of Eye Movements*. Oxford, UK: Oxford University Press.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research, 44*(10), 1013-1024.

Lovasik, J. V., Spafford, M., & Szymkiw, M. (1985). Modification of pattern reversal vers by ocular accommodation. *Vision Research, 25*(4), 599-608.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin, 109*(2), 163-203.

Malmkjaer, K. (Ed.). (2002). *The Linguistics Encyclopedia* (2nd ed.). London: Routledge.

Mangun, G. R., & Hillyard, S. A. (1995). Mechanisms and models of selective attention. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 40-85). New York, NY: Oxford University Press.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia, 37*(6), 661-670.

Martins, A., Balachandran, C., Klistorner, A. I., Graham, S. L., & Billson, F. A. (2003). Effect of pupil size on multifocal pattern visual evoked potentials. *Clinincal and Experimental Ophthalmology, 31*(4), 354-356.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science, 315*(5810), 393-395.

Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text. *Psychological Science, 15*(1), 1-7.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological Bulletin, 81*(12), 899-917.

May, R. M. (2004). Uses and abuses of mathematics in biology. *Science, 303*(5659), 790-793.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1*(1), 11-38.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychological Review, 88*(5), 375-407.

McConkie, G. W., Kerr, P. W., & Dyre, B. P. (1994). What are "normal" eye movements during reading: Toward a mathematical description. In J. Ygge & G. Lennestrand (Eds.), *Eye movements in reading* (pp. 315-327). Oxford: Pergamon.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye-Movement Control during Reading: I. The Location of Initial Eye Fixations on Words. *Vision Research, 28*(10), 1107-1118.

McConkie, G. W., & Rayner, K. (1975). Span of effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*(6), 578-586.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(1), 196-204.

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin, 136*(2), 188-197.

McVay, J. C., & Kane, M. J. (2012a). Drifting from slow to "d'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology-Learning Memory and Cognition, 38*(3), 525-549.

McVay, J. C., & Kane, M. J. (2012b). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology-General, 141*(2), 302-320.

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review, 16*(5), 857-863.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance. 1. Basic mechanisms. *Psychological Review, 104*(1), 3-65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance. 2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*(4), 749-791.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology-General, 132*(2), 228-252.

Molenberghs, P., Gillebert, C. R., Schoofs, H., Dupont, P., Peeters, R., & Vandenberghe, R. (2009). Lesion neuroanatomy of the Sustained Attention to Response task. *Neuropsychologia, 47*(13), 2866-2875.

Monsell, S. (1991). The nature and locus of the word frequency effect in reading. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148 –197). Hillsdale, NJ: Erlbaum.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading - evidence for parallel programming of saccades. *Journal of Experimental Psychology-Human Perception and Performance, 10*(5), 667-682.

Morton, J. (1964). The effects of context upon speed of reading, eye-movements and eye-voice span. *Quarterly Journal of Experimental Psychology, 16*(4), 340-354.

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology-General*.

Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "Zoom Lens" of visual attention. *The Journal of Neuroscience, 23*(9), 3561-3565.

Müller, W., Kollert, A., & Zachert, C. (1988). Pupil size and the steady-state pattern reversal visual evoked cortical potential. *Documenta Ophthalmologica, 68*(3-4), 357-361.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111*(3), 721-756.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*(4), 453-465.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology-Learning Memory and Cognition,* *24*(5), 1091-1104.

Naveh-Benjamin, M., & Guez, J. (2000). Effects of divided attention on encoding and retrieval processes: Assessment of attentional costs and a componential analysis. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(6), 1461-1482.

Navon, D., & Gopher, D. (1979). Economy of the human-processing system. *Psychological Review, 86*(3), 214-255.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know - verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113*(2), 327-357.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*(3), 322-336.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45*(17), 2201-2217.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research, 47*(7), 990-1002.

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9*(5), 28.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience, 29*(26), 8604-8611.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and Performance XII: The Psychology of Reading* (pp. 363–383). Hillsdale, NJ: Erlbaum.

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology, 117*(9), 1885-1901.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsoy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition, 15*(4), 700-708.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction, 1*, 117-175.

Pashler, H. E. (1998). *The Psychology of Attention*. Cambridge, MA: MIT Press.

Pinheiro, J., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye-movement control from the perceptual span in reading. *Perception & Psychophysics, 40*(2), 123-130.

Pollatsek, A., Reichle, E. D., Juhasz, B. J., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology-Human Perception and Performance, 34*(3), 726-750.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006a). Attention to one word at a time in reading is still a viable hypothesis: Rejoinder to Inhoff, Radach, and Eiter (2006). *Journal of Experimental Psychology-Human Perception and Performance, 32*(6), 1496-1500.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006b). Serial processing is consistent with the time course of linguistic information extraction from consecutive words during eye fixations in reading: A response to Inhoff, Eiter, and Radach (2005). *Journal of Experimental Psychology-Human Perception and Performance, 32*(6), 1485-1489.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006c). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1-56.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 178-202.

R Development Core Team. (2012). R: A language and environment for statistical computing. (Version 2.15.1). R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org; under the GNU General Public License, Version 2, June 1991.

Rabovsky, M., Alvarez, C. J., Hohlfeld, A., & Sommer, W. (2008). Is lexical access autonomous? Evidence from combining overlapping tasks with recording event-related brain potentials. *Brain Research, 1222*, 156-165.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller & J. Pynte (Eds.), *Reading* *as a perceptual process* (pp. 165-191). Amsterdam: Elsevier Science.

Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences, 14*(4), 180-190.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98*(2), 676-682.

Raney, G. E., & Rayner, K. (1995). Word-frequency effects and eye-movements during 2 readings of a text. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 49*(2), 151-173.

Rao, K. V., & Proctor, R. W. (1984). Study-phase processing and the word-frequency effect in recognition memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 10*(3), 386-394.

Rayner, K. (1975). Perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65-81.

Rayner, K. (1977). Visual-attention in reading - Eye-movements reflect cognitive-processes. *Memory & Cognition, 5*(4), 443-448.

Rayner, K. (1979). Eye guidance in reading - fixation locations within words. *Perception, 8*(1), 21-30.

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

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology-Human Perception and Performance, 30*(4), 720-732.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*(4417), 468-469.

Rayner, K., Binder, K. S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: word predictability has little influence on initial landing positions in words. *Vision Research, 41*(7), 943-954.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading - Effects of word-frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*(5), 734-747.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology-Human Perception and Performance, 7*(1), 167-179.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word N + 2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology-Human Perception and Performance, 33*(1), 230-245.

Rayner, K., & Pollatsek, A. (1981). Eye-movement control during reading - evidence for direct control. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 33*(Nov), 351-373.

Rayner, K., & Pollatsek, A. (1987). Eye-movements in reading - A tutorial review. *Attention and Performance*(12), 327-362.

Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. New York: Prentice Hall.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology-General, 136*(3), 520-529.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3*(2), 245-248.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology-Human Perception and Performance, 22*(5), 1188-1200.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31*(6), 537-550.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 213-234). Amsterdam: Elsevier.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(2), 294-320.

Reichle, E. D. (2006). Computational models of eye-movement control during reading: Theories of the "eye-mind" link. *Cognitive Systems Research, 7*(1), 2-3.

Reichle, E. D. (2011). Serial attention models of reading. In S. P. Liversedge, I. D. Gilchrist & S. Everling (Eds.), *Oxford Handbook on Eye Movements* (pp. 767-786). Oxford, UK: Oxford University Press.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences, 13*(3), 115-119.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125-157.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*(1), 4-22.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: a unified framework for understanding the eye-mind link. *Psychological Review, 119*(1), 155-185.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445-526.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science, 21*(9), 1300-1310.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*(1), 1-21.

Reilly, R. G., & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research, 38*(2), 303-317.

Reilly, R. G., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognition and applied aspects of eye movement research* (pp. 429–455). Oxford: Elsevier.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7*(1), 34-55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17*(9), 742-746.

Riby, L. M., Smallwood, J., & Gunn, V. P. (2008). Mind wandering and retrieval from episodic memory: A pilot event-related potential study. *Psychological Reports, 102*(3), 805-818.

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7*(1), 23-33.

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging, 26*(2), 451-460.

Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word n+2 in reading. *Journal of Experimental Psychology-Human Perception and Performance, 38*(4), 1026-1042.

Ritchie, D. M. (1993). The Development of the C Language. Retrieved July 23rd, 2012, from Bell Labs/Lucent Technologies: http://cm.bell-labs.com/who/dmr/chist.html

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107*(2), 358-367.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747-758.

Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Research, 49*(20), 2415-2441.

Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research, 1*(4), 201-220.

Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the sauce: the effects of alcohol on mind wandering. *Psychological Science, 20*(6), 747-752.

Sayette, M. A., Schooler, J. W., & Reichle, E. D. (2010). Out for a smoke: the impact of cigarette craving on zoning out during reading. *Psychological Science, 21*(1), 26-30.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*(4-5, Special Issue on Computational Approaches to Reading and Scene Perception), 391.

Schad, D. J., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research, 50*(23), 2600-2616.

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: Objective measures reveal mindless reading at different levels. *Cognition, 125*(2), 179-194.

Schad, D. J., Risse, S., Slattery, T. J., & Rayner, K. (2012). *Early word frequency effects in the fast priming paradigm support immediate cognitive control of eye movements during reading*. Unpublished manuscript, Potsdam.

Schnitzer, B. S., & Kowler, E. (2006). Eye movements during multiple readings of the same text. *Vision Research, 46*(10), 1611-1632.

Schöner, G. (2008). Dynamical systems approaches to cognition. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 101-126). New York: Cambridge University Press.

Schooler, J. W. (2002). Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences, 6*(8), 339-344.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 203-226). Cambridge, MA: MIT Press.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences, 15*(7), 319-326.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 52*(4), 1021-1046.

Seli, P., Cheyne, J. A., Barton, K. R., & Smilek, D. (2012). Consistency of sustained attention across modalities: comparing visual and auditory versions of the SART. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale, 66*(1), 44-50.

Seli, P., Cheyne, J. A., & Smilek, D. (2012). Attention failures versus misplaced diligence: separating attention lapses from speed-accuracy trade-offs. *Consciousness and Cognition, 21*(1), 277-291.

Shaw, G. A., & Giambra, L. (1993). Task-unrelated thoughts of college students diagnosed as hyperactive in childhood. *Developmental Neuropsychology, 9*(1), 17-30.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information-processing. 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127-190.

Shulman, G. L., Corbetta, M., Buckner, R. L., Fiez, J. A., Miezin, F. M., Raichle, M. E., et al. (1997). Common blood flow changes across visual tasks: I. Increases in subcortical structures and cerebellum but not in nonvisual cortex. *Journal of Cognitive Neuroscience, 9*(5), 624-647.

Shulman, G. L., Wilson, J., & Sheehy, J. B. (1985). Spatial determinants of the distribution of attention. *Perception & Psychophysics, 37*(1), 59-65.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage, 20*(1), 114-124.

Skinner, B. F. (1986). Is it behaviorism? *Behavioral and Brain Sciences, 9*(4), 716-716.

Smallwood, J. (2010a). The footprints of a wandering mind: Further examination of the time course of an attentional lapse. *Cognitive Neuroscience, 2*(2), 91-97.

Smallwood, J. (2010b). Why the global availability of mind wandering necessitates resource competition: Reply to McVay and Kane (2010). *Psychological Bulletin, 136*(2), 202-207.

Smallwood, J. (2011a). *Distinguishing how from why the mind wanders: A process - motivation hypothesis for the mind wandering state*. Unpublished manuscript.

Smallwood, J. (2011b). Mind-wandering while reading: attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistics Compass, 5*(2), 63-77.

Smallwood, J. (in press). Distinguishing how from why the mind wanders. A process-occurrence framework for self-generated mental activity. *Psychological Bulletin*.

Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*(3), 452-484.

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience, 20*(3), 458-469.

Smallwood, J., Brown, K. S., Baird, B., Mrazek, M. D., Franklin, M. S., & Schooler, J. W. (2012). Insulation for daydreams: a role for tonic norepinephrine in the facilitation of internally guided thought. *Plos One, 7*(4), e33706.

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., et al. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *Plos One, 6*(3).

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., et al. (2004). Subjective experience and the attentional lapse: task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*(4), 657-690.

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review, 14*(2), 230-236.

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: negative moods lead the mind to wander. *Emotion, 9*(2), 271-276.

Smallwood, J., McSpadden, M., Luus, B., & Schooler, J. (2008). Segmenting the stream of consciousness: The psychological correlates of temporal structures in the time series data of a continuous performance task. *Brain and Cognition, 66*(1), 50-56.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review, 14*(3), 527-533.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*(6), 1144-1150.

Smallwood, J., O'Connor, R. C., & Heim, D. (2004-2005). Rumination, dysphoria, and subjective experience. *Imagination, Cognition and Personality, 24*(4), 355 - 367.

Smallwood, J., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. *Consciousness and Cognition, 15*(1), 218-231.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958.

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010a). Failures of sustained attention in life, lab, and brain: ecological validity of the SART. *Neuropsychologia, 48*(9), 2564-2570.

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010b). Out of mind, out of sight: eye blinking as indicator and embodiment of mind wandering. *Psychological Science, 21*(6), 786-789.

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010c). Out of mind, out of sight: eye blinking as indicator and embodiment of mind wandering. *Psychological Science*.

Spearman, C. (1914). The theory of two factors. *Psychological Review, 21*(2), 101-115.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: some current controversies. *Trends in Cognitive Sciences, 5*(4), 156-163.

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General, 140*(3), 407-433.

Stawarczyk, D., Majerus, S., Maquet, P., & D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. *Plos One, 6*(2), e16997.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*(4), 242-248.

Trukenbrod, H. A., & Engbert, R. (2009). The control of fixation duration: Time-course of the response to stepwise changes in processing difficulty. *Journal of Vision, 9*(8), 437-437.

Trukenbrod, H. A., & Engbert, R. (submitted). *ICAT: A computational model for the adaptive control of fixation durations*. Unpublished manuscript.

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*(2), 153-162.

Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Consciousness and Cognition, 20*(4), 1882-1886.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.). New York: Springer.

Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology, 16*(1-2), 154-177.

Vitu, F., Lancelin, D., & d'Unienville, V. M. (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology-Human Perception and Performance, 33*, 1220-1249.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research, 41*(25-26), 3513-3533.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading - Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57*(3), 352-364.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology, 100*(3), 426-432.

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E-Z Reader. *Cognition, 111*(1), 132-137.

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience, 9*(7), 971-978.

Whaley, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17*(2), 143-154.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review, 12*(5), 891-896.

Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. New York: Springer.

Wotschack, C. (2009). *How reading strategies modulate effects of distributed processing and oculomotor control: an eye movement study.*, University of Potsdam, Potsdam.

Wotschack, C., & Kliegl, R. (2011). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology*.

Wright, D. B., Horry, R., & Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods, 41*(2), 257-267.

Wyman, D., & Steinman, R. M. (1973). Latency characteristics of small saccades. *Vision Research, 13*(11), 2173-2175.

Yang, S. N. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research, 7*(1), 56-69.

Yang, S. N., & McConkie, G. W. (2001). Eye movements during reading: a theory of saccade initiation times. *Vision Research, 41*(25-26), 3567-3585.

Zola, D. (1984). Redundancy and word perception during reading. *Perception & Psychophysics, 36*(3), 277-284.

# Appendix

## A Global analyses (Chapter 2)

**Table A-1**

Eye-movement statistics for reading shuffled and normal text.

| Variable | shuffled text |  | normal text |  | *t*-test |  |  |
|---|---|---|---|---|---|---|---|
|  | *M* | (*SD*) | *M* | (*SD*) | *t* | *df* | *p* |
| N of readers | 30 |  | 30 |  |  |  |  |
| N of fixations/sentence | 10.1 | (2.0) | 7.8 | (1.4) | 5.14 | 51 | < 0.001 |
| N of sentences | **139** | (7) | **137** | (8) |  |  |  |
| *Fixation probabilities* |  |  |  |  |  |  |  |
| skipping (p0) | 0.10 | (0.06) | 0.21 | (0.08) | -6.45 | 55 | < 0.001 |
| single fixation (p1) | 0.70 | (0.06) | 0.67 | (0.06) | 1.69 | 58 | 0.10 |
| double-plus fixation (p2+) | 0.16 | (0.07) | 0.08 | (0.04) | 5.41 | 48 | < 0.001 |
| regression (prg) | **0.06** | (0.04) | **0.06** | (0.04) |  |  |  |
| mean saccade length (letters) | 6.1 | (0.9) | 7.6 | (1.2) | -5.85 | *55* | < 0.001 |
| *Fixation position (letter)* |  |  |  |  |  |  |  |
| single fixation (l0) | 2.5 | (0.3) | 2.7 | (0.2) | -3.22 | 53 | < 0.01 |
| 1st of multiple (l1) | **2.0** | (0.4) | **2.2** | (0.6) |  |  |  |
| 2nd of multiple (l2) | **5.5** | (0.4) | **5.5** | (0.8) |  |  |  |
| *Fixation duration (ms)* |  |  |  |  |  |  |  |
| single fixation (d0) | 254 | (40) | 213 | (32) | 4.37 | 55 | < 0.001 |
| 1st of multiple (d1) | 227 | (32) | 199 | (30) | 3.50 | 58 | < 0.001 |
| 2nd of multiple (d2) | 197 | (36) | 172 | (36) | 2.74 | 58 | < 0.01 |
| gaze duration | 293 | (58) | 231 | (37) | 4.89 | 50 | < 0.001 |
| Reading rate (words/min) | 193 | (47) | 250 | (46) | -4.74 | 58 | < 0.001 |

*Note (continued)*. No invalid fixations were removed. Data are from right eye. Mean n of fixations (N), regression probability (prg) and reading rate are based on all fixations; all other measures are based on first-pass reading. Welch t-tests over participants were used to test differences between normal and shuffled text reading. Values of **non**-significant differences (*ps* > .25) are printed in **bold**.

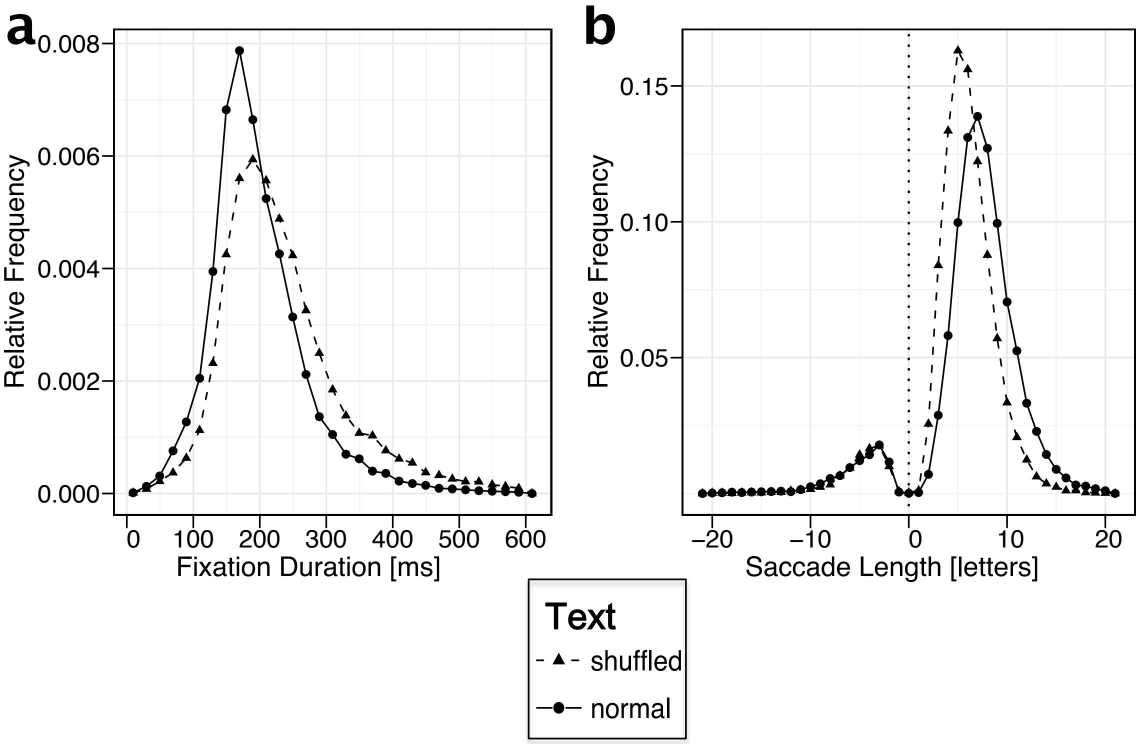


**Figure A-1**

Global analyses. (a) Distribution of all observed valid fixation durations during reading of randomly shuffled text (triangles and dashed line) vs. normal reading (circles and solid line). Displays the corresponding mean frequency distributions. Relative proportions of fixation durations are displayed for 31 levels (from 0 ms up to 620 ms in 20-ms steps). (b) Distributions of all observed saccade lengths. Negative saccade lengths indicate regressive saccades.

## B LME models (Chapter 2)

**Table B-1**

Means, standard errors, and *t*-values of fixed effects on fixation durations; Variances and standard deviations of the random effects. Linear mixed model fit by restricted maximum likelihood (REML).

| | Log gaze durations | | | Log single fixation durations | | |
|---|---|---|---|---|---|---|
| **Fixed Effects:** | Estimate | SE | t-value | Estimate | SE | t-value |
| Intercept | 5.540 | 0.029 | 191.3 | 5.489 | 0.030 | 182.2 |
| *Word n* | | | | | | |
| frequency (frq) | **-0.0002** | **0.006** | **-0.04** | *0.009* | *0.005* | *1.77* |
| frq*frq | [a] 0.020 | 0.005 | 4.3 | [a] 0.017 | 0.004 | 4.5 |
| 1/length (lgth) | -0.739 | 0.072 | -10.2 | 0.265 | 0.063 | 4.2 |
| *Word n – 1* | | | | | | |
| frequency | -0.035 | 0.002 | -15.1 | -0.034 | 0.003 | -12.5 |
| 1/length | [a] 0.247 | 0.026 | 9.5 | 0.207 | 0.035 | 5.9 |
| *Word n + 1* | | | | | | |
| frequency | -0.016 | 0.002 | -6.9 | -0.016 | 0.002 | -6.6 |
| 1/length | [a] 0.119 | 0.025 | 4.8 | [a] 0.114 | 0.026 | 4.4 |
| *Viewing position* | | | | | | |
| last sacc. amplit. | 0.017 | 0.001 | 13.6 | [a] 0.027 | 0.001 | 30.1 |
| pos in word | -0.138 | 0.017 | -8.3 | [a] -0.082 | 0.013 | -6.2 |
| pos*pos | -1.088 | 0.050 | -21.6 | [a] -0.348 | 0.038 | -9.3 |
| next sacc. amplit. | -0.007 | 0.001 | -5.6 | [a] 0.011 | 0.001 | 10.4 |
| *Interactions* | | | | | | |
| (frq n)/(lgth n) | [a] 0.401 | 0.055 | 7.3 | **-0.063** | **0.048** | **-1.3** |
| (frq n)*(frq n-1) | 0.015 | 0.001 | 10.4 | 0.019 | 0.002 | 12.3 |
| (frq n)*(frq n+1) | 0.010 | 0.002 | 5.5 | 0.009 | 0.002 | 5.1 |
| (frq n+1)/(lgth n) | 0.060 | 0.025 | 2.4 | [a] 0.099 | 0.021 | 4.7 |
| **Slope-differences between shuffled and normal text reading** | | | | | | |
| Experim. Cond. (Exp) | -0.296 | 0.040 | -7.4 | -0.270 | 0.042 | -6.4 |
| *Word n* | | | | | | |
| Exp*frq | -0.026 | 0.004 | -6.7 | -0.028 | 0.004 | -7.0 |
| Exp*frq*frq | [a] | | | [a] | | |
| Exp*lgth | 0.483 | 0.054 | 9.0 | **-0.009** | **0.057** | **-0.2** |
| *Word n – 1* | | | | | | |
| Exp*frq | **-0.004** | **0.003** | **-1.2** | *-0.008* | *0.004* | *-1.7* |
| Exp*lgth | [a] | | | *0.105* | *0.054* | *1.94* |
| *Word n + 1* | | | | | | |
| Exp*frq | *0.006* | *0.003* | *1.86* | **0.005** | **0.003** | **1.6** |
| Exp*lgth | [a] | | | [a] | | |
| *Viewing position* | | | | | | |
| Exp*last sacc. amp. | 0.007 | 0.002 | 4.0 | [a] | | |
| Exp*pos in word | **0.033** | **0.023** | **1.5** | [a] | | |
| Exp*pos*pos | 0.453 | 0.070 | 6.5 | [a] | | |
| Exp*next sac. amp. | 0.013 | 0.002 | 7.8 | [a] | | |
| *Interactions* | | | | | | |
| Exp*(frq n)/(lgth n) | [a] | | | 0.177 | 0.037 | 4.8 |
| Exp*(frq n)*(frq n-1) | -0.019 | 0.002 | -7.9 | -0.018 | 0.002 | -7.2 |
| Exp*(frq n)*(frq n+1) | -0.010 | 0.003 | -3.3 | -0.010 | 0.002 | -4.0 |
| Exp*(frq n+1)/(lgth n) | 0.102 | 0.043 | 2.4 | [a] | | |

**Random effects:**

| Groups | Name | Variance | Std.Dev. | Variance | Std.Dev. |
|---|---|---|---|---|---|
| Word ID | Intercept | 0.0085 | 0.092 | 0.0045 | 0.067 |
| Reader | Intercept | 0.0237 | 0.153 | 0.0262 | 0.162 |
| Residual | | 0.1232 | 0.351 | 0.0839 | 0.290 |

| | Log gaze durations | Log single fixation durations |
|---|---|---|
| *N* of fixations | 38,738 | 24,433 |
| AIC | 30,282 | 9,921 |
| BIC | 30,548 | 10,148 |
| logLik | -15,110 | -4,933 |

*Note (continued)*. All data are from right eye [60 readers; 550 unique word IDs]. **Non-**significant coefficients are set in bold ($t$ < 1.96). *Marginally* significant coefficients are set in italics ($t$ ≤ 1.645). Shuffled text reading is the reference condition. Experimental condition (Exp) depicts the contrast between that reference condition and the normal text reading condition using a dummy-coded factor. [a] The slope-difference between shuffled and normal text reading was not significant for these effects, thus the interactions of the respective effect with experimental condition was dropped from the model. The main effect reflects the average effect in shuffled and normal reading.

## C Supplementary material (Chapter 2)

The same statistical models were used as in the primary analyses to test relative effects of word frequency (see Figure 2-3) and the results from these analyses are reported in Table C-1.

For the analyses for content words *n*, all fixations on function words were discarded from the analyses.

We controlled for effects of orthographic neighbors of word *n* by adding the number and cumulative frequency of orthographic neighbors (*Coltheart distance* = 1) as well as the number and cumulative frequency of higher-frequency orthographic neighbors (*Coltheart distance* = 1) and the interaction of these four variables with experimental condition (shuffled versus normal text) into the (g)lme models. Afterwards, non-significant predictors involving orthographic neighbors were dropped from the model. This is how we controlled for effects of orthographic neighborhood when testing effects of word frequency.

In both control analyses, frequency effects for short words were significantly reversed during shuffled text reading for current-word gaze duration and regression probability, but not for long words or normal text, in accordance with the respective effects in the main analyses. In addition, the lag-frequency effect was also stronger for short words in the shuffled text, as was the case in the primary analyses. Thus, critical differences between reading conditions did not depend on content/function words and were not driven by orthographic neighbors.

**Table C-1**

Post-hoc tests for relative frequency effects (cf. Figure 2-3). Analyses of content words *n*; and controlling for effects of orthographic neighbors.

| **Dependent variable / Effect** | **Shuffled text** | | **Normal text** | |
|---|---|---|---|---|
| **Analyses for content words *n*** | | | | |
| | Short words *n* | Long words *n* | Short words *n* | Long words *n* |
| (Log) Gaze duration on word *n* | | | | |
| Linear frequency effect | *b* = 0.022<br>*SE* = 0.012<br>*t* = 1.83 | *b* = -0.042<br>*SE* = 0.015<br>*t* = -2.79 | *b* = 0.0004<br>*SE* = 0.012<br>*t* = 0.04 | *b* = -0.63<br>*SE* = 0.014<br>*t* = -4.58 |
| Probability for a regression to word *n* | | | | |
| Linear frequency effect | *b* = 0.255<br>*SE* = 0.093<br>*p* < .01 | *b* = -0.032<br>*SE* = 0.106<br>*p* = .76 | *b* = -0.250<br>*SE* = 0.111<br>*p* < .05 | *b* = -0.129<br>*SE* = 0.124<br>*p* = 0.30 |
| (Log) Duration of the first fixation on word *n* + 1 after having made a single fixation on word *n* | | | | |
| | Effect in shuffled text | | Difference shuffled vs. normal text | |
| | All words *n* | Word *n* = Content word | All words *n* | Word *n* = Content word |
| Interaction of word frequency and word length | *b* = -0.179<br>*SE* = 0.033<br>*t* = -5.40 | *b* = -0.283<br>*SE* = 0.053<br>*t* = -5.30 | *b* = 0.120<br>*SE* = 0.055<br>*t* = 2.17 | *b* = 0.268<br>*SE* = 0.087<br>*t* = 3.07 |
| **Controlling for effects of orthographic neighbors of word *n*** | | | | |
| | Short words *n* | Long words *n* | Short words *n* | Long words *n* |
| (Log) Gaze duration on word n | | | | |
| Linear frequency effect | *b* = 0.033<br>*SE* = 0.009<br>*t* = 3.79 | *b* = -0.014<br>*SE* =0.013<br>*t* = -1.05 | *b* = 0.011<br>*SE* = 0.009<br>*t* = 1.24 | *b* = -0.036<br>*SE* = 0.013<br>*t* = -2.88 |
| Probability for a regression to word n | | | | |
| Linear frequency effect | *b* = 0.279<br>*SE* = 0.092<br>*p* < .01 | *b* = -0.112<br>*SE* = 0.095<br>*p* = .24 | *b* = -0.031<br>*SE* = 0.106<br>*p* = .77 | *b* = -0.045<br>*SE* = 0.112<br>*p* = .69 |
| (Log) Duration of the first fixation on word *n* + 1 after having made a single fixation on word *n* | | | | |
| | Effect in shuffled text | | Difference shuffled vs. normal text | |
| | Standard analysis | Controlling for neighborhood | Standard analysis | Controlling for neighborhood |
| Interaction of word frequency and word length | *b* = -0.179<br>*SE* = 0.033<br>*t* = -5.40 | *b* = -0.170<br>*SE* = 0.035<br>*t* = -4.85 | *b* = 0.120<br>*SE* = 0.055<br>*t* = 2.17 | *b* = 0.121<br>*SE* = 0.058<br>*t* = 2.10 |

## D Using parameters of cognitive models for hypothesis testing (Chapter 3)

In this Appendix, we propose a minimum set of criteria for valid model comparisons. First, fitting models to experimental data always comprises the risk of overfitting error variance instead of capturing valid and reliable effects. This is particularly problematic for high-dimensional models containing many free parameters such as current models of eye movement control during reading. To guard against overfitting, we implement a cross-validation by splitting data into subsets containing half of the data. A training set is used for the estimation of model parameters. Estimated parameters are then used in Monte Carlo simulations to predict eye movements on a distinct and independent evaluation set (or test sample), where model predictions are compared to experimental results.

Second, we suggest that several basic eye movement phenomena should be checked for each estimated parameter set to ensure that model behaviour is reasonable for standard eye-movement effects. We suggest that it is particularly informative to investigate distributions of fixation durations, saccade lengths, and landing positions, as well as basic oculomotor effects like the optimal viewing position (OVP) effect on refixation probabilities and inverted optimal viewing position (IOVP) effects on fixation durations (Vitu, et al., 2001). Moreover, effects of word length and frequency on various measures of fixation durations and fixation probabilities provide benchmark results for model evaluation.

Third, experimental results in different reading conditions are often quite similar in many respects. In our present analyses, several effects in eye movements were present in both shuffled and normal text conditions. For example, readers in both conditions exhibited Gaussian landing site distributions, an OVP effect on refixation probabilities, IOVP effects on measures of fixation durations, and effects of word length and frequency on fixation durations and fixation probabilities. Qualitatively replicating experimental effects in each task with numerical model simulations therefore does not guarantee that estimated model parameters capture variance that is specific to both tasks. We here suggest two ways how more specific model predictions can be tested. As a first step, we consider it critical to investigate effects that (1) specifically differ between tasks and (2) are meaningfully related to the estimated model parameters. In the present work, we are interested in specific differences in how word frequency and word length influence fixation durations during normal and shuffled text reading. We have previously proposed hypotheses about what cognitive processes may cause these effects, namely effects of foveal load on the perceptual span.

Fourth, an even closer model test should be performed before task differences in parameter estimates can be relied upon. Such a test provides evidence that parameter estimates for the SWIFT 3 model capture valid task-specific differences in eye movements. As a minimal criterion, we suggest deriving (1) model predictions for data observed in the test sample of a task. These predictions should be better than (2) predictions from the model for the other task and better than (3) predictions derived from the experimental data from the other task. For example, we predict experimental eye movement data in the test sample for shuffled text based on (1) simulations of the shuffled-SWIFT model, (2) simulations of the normal-SWIFT model, and (3) experimental data on normal text reading. When comparing these three predictions, one could postulate that predictions from the shuffled-SWIFT model must be as good or better than predictions from the normal-SWIFT model, and than predictions from experimental data on normal text reading. We will also test this criterion when predicting experimental data from normal text reading.

# E Supplementary information: Model predictions on standard eye movement effects (Chapter 3)

## E.1 Distributions of Fixation Durations and Saccade Lengths

As a first statistical result, we compared the distributions of fixation durations for model simulations to the respective experimental data. Figure E-1 shows that the random walk assumptions can explain the variance contained in fixation durations. Also the experimentally observed task differences, i.e., an increase in mean and variance of fixation durations for shuffled text reading, was captured by the shuffled-SWIFT model.

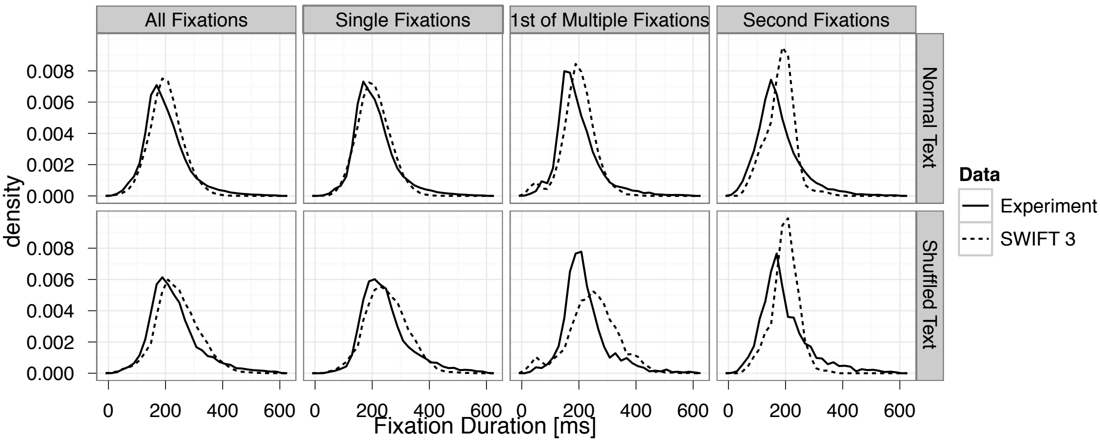


**Figure E-1**

Distributions of fixation durations for experimental data (solid lines) and model simulations (dashed lines) for normal (upper panel) and shuffled (lower panel) text. Fixation duration measures: All valid fixation durations (left panel), single fixation durations (middle left panel), first of multiple fixation durations (middle right panel), second fixation durations (right panel).

Distributions of saccade lengths were well reproduced for reading of normal text, particularly for forward-directed saccades (Figure E-2). Note that the SWIFT model produces distributions for forward-directed and regressive saccades based on one single mechanism. Saccade lengths during shuffled text reading were somewhat shorter, which was well reproduced by the shuffled-SWIFT model. The experimentally observed reduction in variance of forward saccades, however, was not evident in the model simulations. Note that the distributions of fixation durations and saccade lengths were not included into the function for optimizing model parameters.

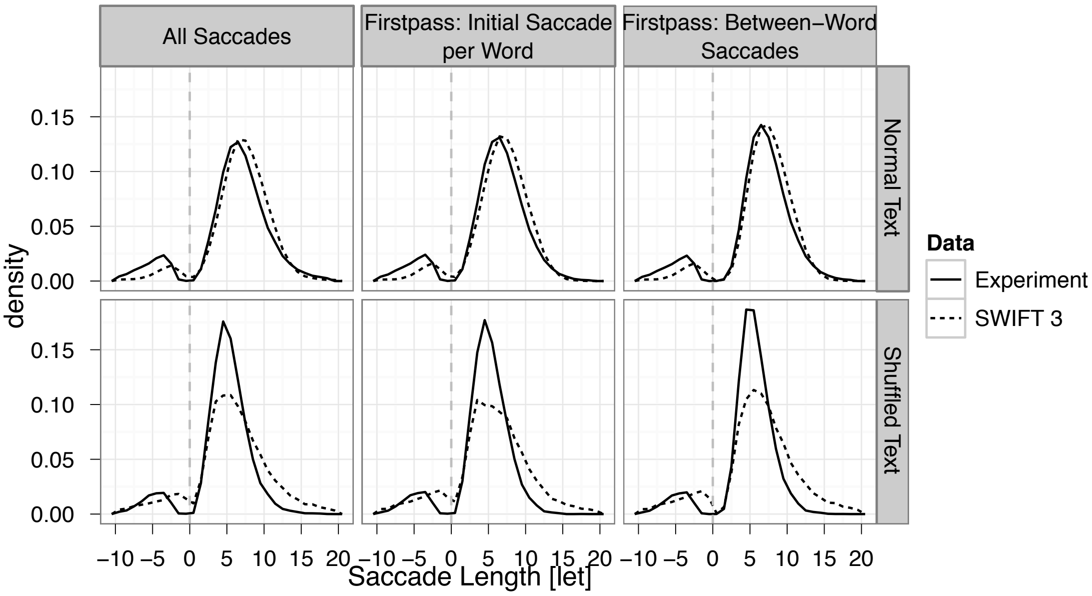


**Figure E-2**

Distributions of saccade lengths for experimental data (solid lines) and model simulations (dashed lines) for normal (upper panel) and shuffled (lower panel) text. Positive values indicate forward-directed saccades, negative values indicate the lengths of regressive saccades. Measures of saccade lengths: All valid saccades (left panel), initial saccade after initially fixating a word in firstpass (middle panel), between-word saccades in firstpass (right panel).

### E.2 Initial Landing Positions

Given a good agreement of distributions of fixation durations, we now investigate basic oculomotor assumptions in the SWIFT model. The landing positions of initial fixations on a word during reading approximately follow Gaussian distributions and exhibit a considerable variance. Important factors influencing the maximum and variance of landing position distributions are the launch site distance and word length (McConkie et al., 1988).

Model predictions are generally in good agreement with the experimental data (see Figure E-3). Model simulations reproduced the effects a) that the maxima of the landing site distributions were shifted toward the word beginning for large launch site distances and were shifted toward word endings for small launch site distances and b) that the variance of landing site distributions increased with increasing launch site distance and word length. Thus, effects of saccade range error were clearly present in the simulated data for both shuffled and normal text.

In addition, the simulated data reproduced differences in landing site distributions between shuffled and normal text reading. Maxima of landing site distributions were shifted toward word beginnings for readers of shuffled text, and the shuffled-SWIFT model reproduced this shift. The shift was mainly present for small (launch sites -1, -3, -5), but not so much for large launch site distances (see launch site -7), indicating that the effect of launch site distance on the maxima of landing site distributions was reduced for shuffled text. Also, variances of landing site distributions were somewhat reduced during shuffled text reading. All of these effects were also visible in the model simulations.

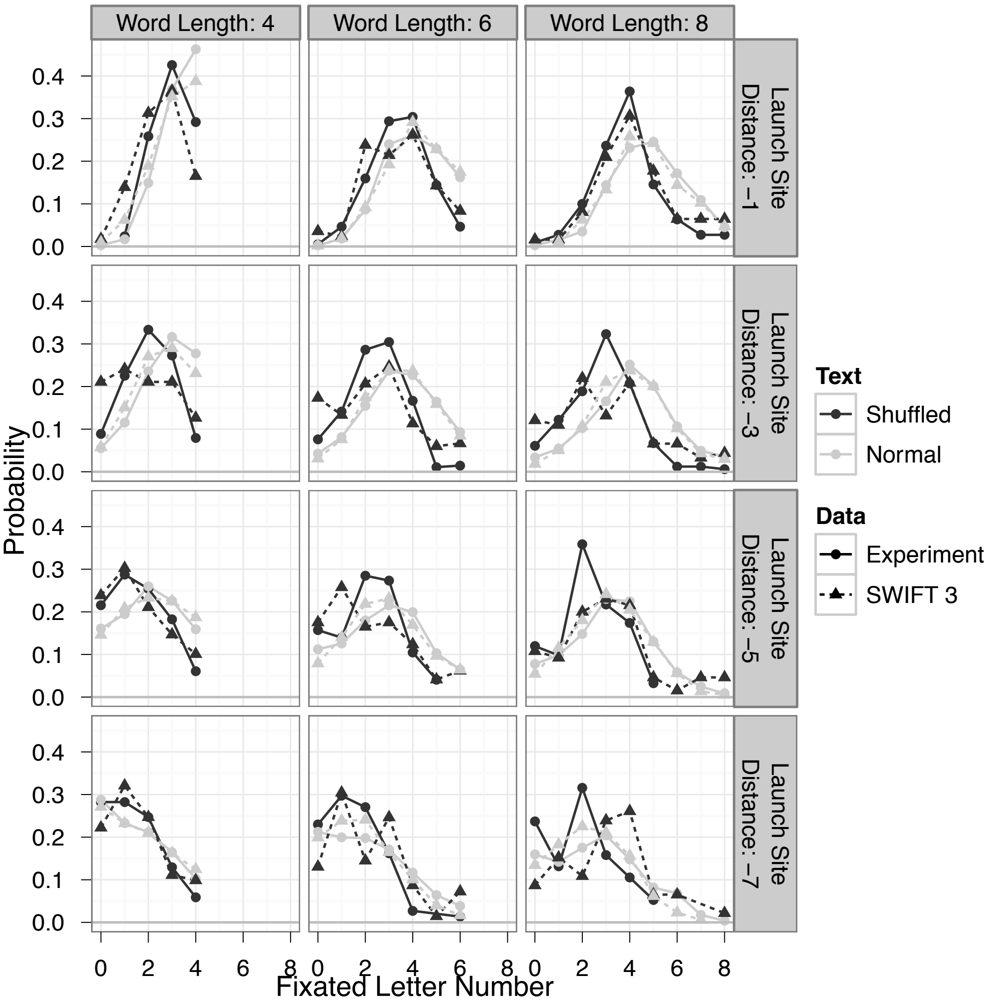


**Figure E-3**

Distributions of initial landing positions by word length and launch site distance. The columns of panels show distributions for word lengths 4, 6, and 8, and the rows of panels indicate distributions for launch sites -1, -3, -5, and -7.

### E.3 Refixation Probability

Distributions of refixation probabilities over different landing positions indicate the optimal viewing position (OVP) during reading (Vitu et al., 2001). The landing position that is associated with the minimal refixation probability indicates the location that is optimal to process the fixated word during one fixation. In the SWIFT model (see also Engbert et al., 2005), the optimal viewing position emerges as fixations are distributed over landing sites according to assumptions about oculomotor control (McConkie et al., 1988). The assumption of a processing gradient is then sufficient to reproduce the U-shaped forms of the refixation distributions.

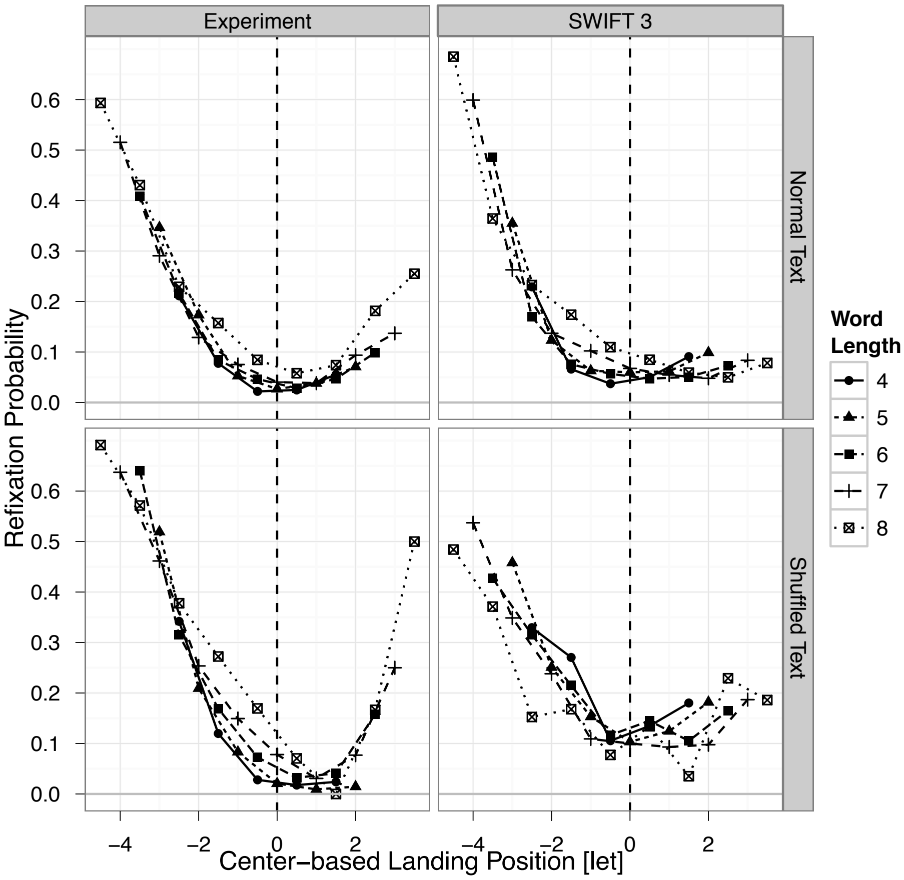


**Figure E-4**

Refixation probabilities after an initial fixation in firstpass as a function of center-based landing position plotted for different word lengths. Experimental data (left panel) show U-shaped curves without an influence of word length. In the model simulations (right panel) these curves are qualitatively reproduced.

For long words, the model produces refixations as long words do not fit into the perceptual span and therefore need to be refixated for complete word processing. Short words, however, fully fit into the perceptual span, and no refixations should be needed to complete visual word processing. The SWIFT model (Engbert et al., 2005) assumes that the random saccade timer also triggers saccades early for short words such that refixations are necessary for complete processing. The SWIFT 3 model qualitatively reproduced refixation probabilities during normal and shuffled text reading (see Figure E-4). For normal text reading, the distributions are nicely met by model simulations. The shuffled-SWIFT model correctly reproduced the observed increase in refixation probabilities as compared to normal text reading. However, refixation probability at word centres was overestimated by shuffled-SWIFT, and refixations were underestimated at word beginnings. Note, however, that the effect of landing position was not included in the parameter fits.

### E.4 Inverted Optimal Viewing Position

Based on refixation results for the OVP, one may expect that fixation durations are shortest for fixations at word centers and longer at the edges of words. This, however, is not the case. Vitu and colleagues (2001) were the first to report that fixation durations are longer for fixations at word centers and shorter and the edges of words, which was called an *inverted* OVP effect (IOVP) of fixation durations (see also Nuthmann et al., 2005; Kliegl et al., 2005). For single fixation durations, the SWIFT model explains the IOVP effect via error-correcting after misguided saccades (Engbert et al., 2005; Nuthmann et al., 2005). If a saccade fails the intended word target and lands on a neighboring, unintended word, then immediately a new saccade program is triggered, which can potentially lead to error correction. The likelihood for mislocated fixations is highest at word boundaries, at the first or last letters for each word. This mechanism reduces mean fixation durations at word edges, which can explain the single fixation duration IOVP effect (Engbert et al., 2005).

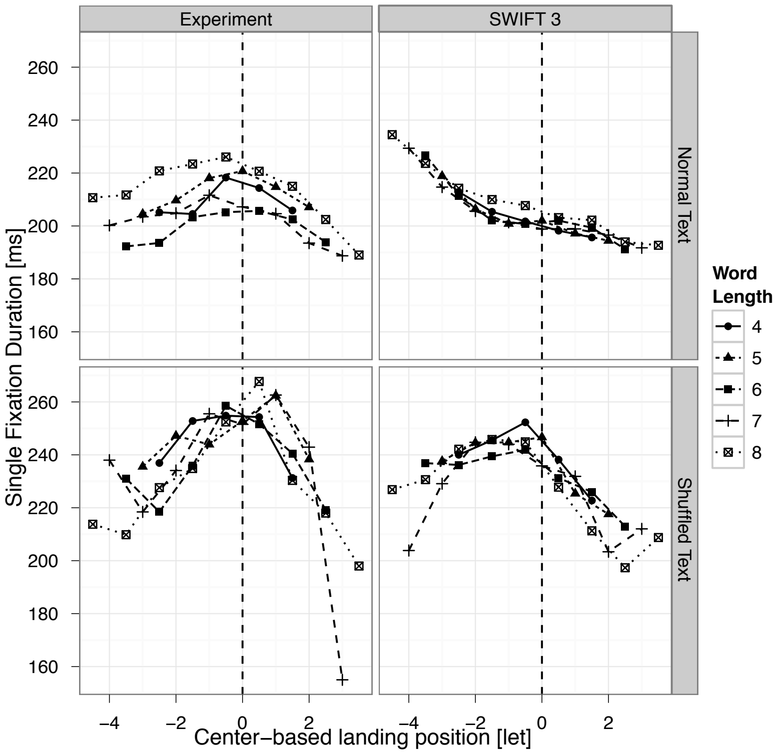


**Figure E-5**

Effects of inverted optimal viewing position for single fixation durations as a function of initial landing site. Effects are shown for model simulations (right panels) and experimental data (left panels), for the reading of normal text (upper panels) and the reading of shuffled text (lower panels).

In two fixation cases, an IOVP effect is observed for the first fixation duration. Plotting the average second fixation duration as a function of first fixation landing site shows a U-shaped effect (Figure E-6). Assuming error-correction after misguided saccades is not sufficient to explain these effects. The SWIFT (Engbert et al., 2005) introduces an explanatory mechanism for this complicated pattern of first and second fixation durations: saccade latencies are modulated by intended saccade length. This assumption is motivated by findings from neurophysiology showing that programming a very short saccade is a difficult task for the oculomotor system (Adams et al., 2000; Kalesnykas & Hallett, 1994; Wyman & Steinman, 1973) because an extremely short neuronal pulse must be produced by the brainstem saccade generator (e.g., Spark, 2002). This additional assumption is sufficient to produce the compensatory interaction between first and second fixation durations (see Figure E-6; for separate simulations involving each mechanism, see Engbert et al., 2005).

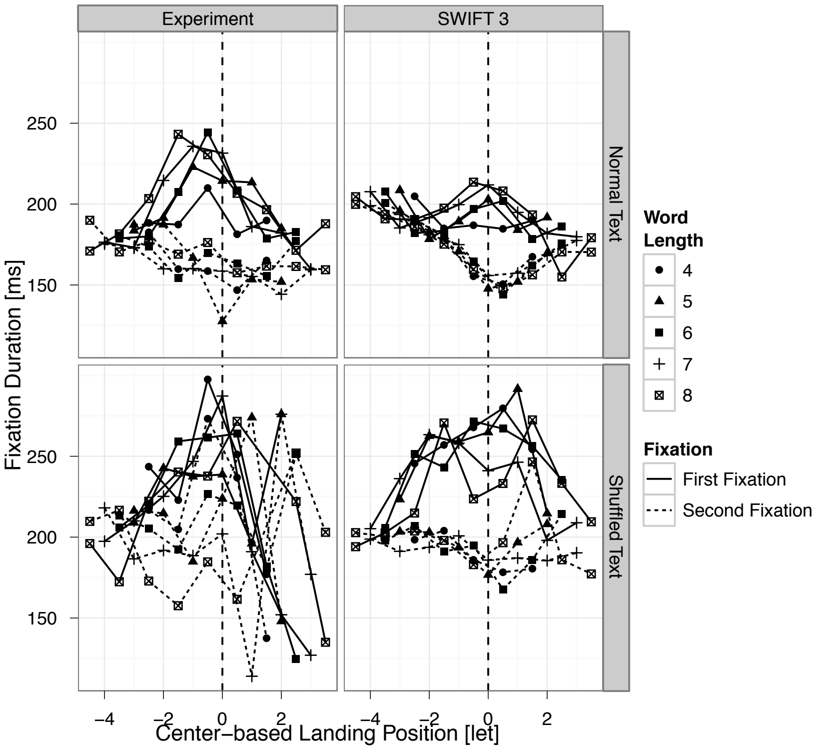


Figure E-6

Effects of inverted optimal viewing position for first and second fixation durations in two fixation cases, as a function of initial landing site. Effects are shown for model simulations (right panels) and experimental data (left panels), for the reading of normal text (upper panels) and the reading of shuffled text (lower panels).

For shuffled text reading, the SWIFT model makes the clear prediction that the single fixation duration IOVP should be stronger for longer average fixation durations. This prediction is based on the fact that the mechanism triggering error-correcting saccades works at a fixed speed, independent of the average fixation durations in a task. Fixations at word edges should be relatively independent from cognitive processing demands and constant over different tasks. Fixations at word centers, to the contrary, should be primarily under control of the random timer, and thus adapt to varying task difficulties. This prediction from the SWIFT model is displayed in Figure E-5 (right panel), and the findings for shuffled text reading qualitatively correspond to the model prediction. For normal text reading, the model did not capture the IOVP effect in single fixation durations well, presumably because the IOVP effect was not included into the procedure for finding optimal model parameters.

The SWIFT model also makes the analogous prediction that the IOVP effect in first fixation durations should be enhanced during reading of shuffled text. During shuffled text reading, readers make shorter saccades on average, and accordingly saccade latencies are more strongly reduced. This prediction from the SWIFT model, as displayed in Figure E-6, was also supported by the experimental results, as the first fixation duration IOVP effect was stronger for shuffled than for normal text reading. Model predictions concerning the IOVP effect were well in line with the observed data.

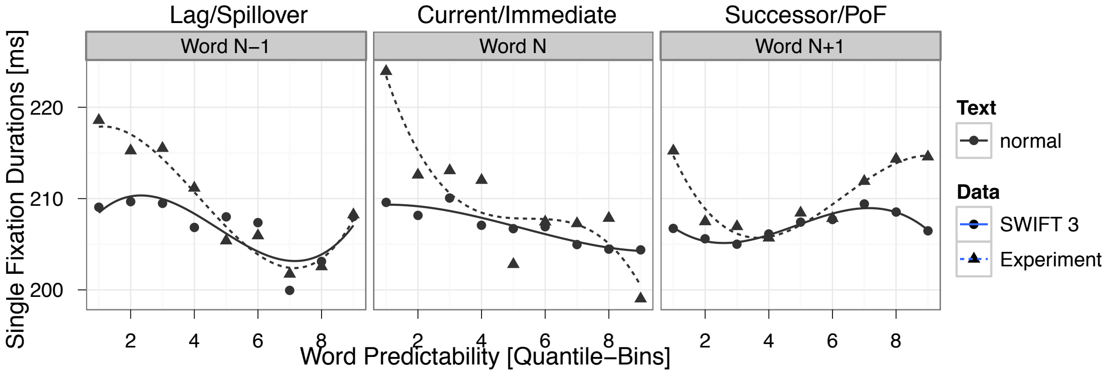


**Figure E-7**

Spatially distributed effect of word predictabilities of words *N*-1, *N*, and *N*+1 on single fixation durations on word *N* for observed (triangles, dashed lines) and simulated (points, solid lines) data. Continuous word predictabilities were categorized into nine quantile-based bins.

### E.5 Word predictability effects in normal text reading

In normal text reading, fixation durations are influenced by whether it's possible to predict upcoming words from their preceding context. Single fixation durations are longer on high predictable words than on low predictable words, and we replicated this standard finding for our experimental data on normal text reading (see Figure E-7). Beyond the current-word predictability effect, word predictability also shows effects of distributed processing: single fixation durations are shorter if the last word *N*-1 was highly predictable and longer if word *N*-1 was of low predictability. The effect of upcoming word *N*+1 predictability on fixation durations on word *N* (i.e., successor effects), however, is reversed (Kliegl et al., 2006): fixation durations are longer before high predictable words, which may indicate processes of memory retrieval for the predicted upcoming word. Although predictability effects were somewhat stronger in observed than in simulated data, the SWIFT 3 model qualitatively reproduced distributed effects of word predictability on single fixation durations (see Figure E-7).

# F Supplementary information (Chapter 4)

## F.1 Statistical analysis method

For statistical analysis, we used (generalized) linear mixed effects models [(G)LMMs] with crossed random effects for subjects, words, and screens. In the *R* system for statistical computing (R Development Core Team, 2012), LMMs were fit using the lmer program of the lme4 package (Bates & Sakar, 2008), and GLMMs were fit using the glmmADMB program (Bohning, Dietz, Schlattmann, Mendonca, & Kirchner, 1999), which provides an interface to the ADMB software (Fournier et al., 2011). (G)LMMs model the dependent variable at the level of single responses (e.g., single eye movements or error detections) and handle imbalance in the design automatically (Baayen, 2008; Baayen, et al., 2008; Kliegl, et al., 2010; Kliegl, et al., 2011).

## F.2 Results

### *F.2.1 Error detection*

To statistically analyze readers' sensitivity for errors (that is, their propensity for mindful reading) and response bias, we fitted logistic generalized linear mixed effects models using the probit link (Wright, et al., 2009). In the GLMM, we regressed participants' responses to error sentences on response bias (*c*) and sensitivity for errors (d'). Differences in these effects between experiments were included as predictors (Δc:Experiment; Δd':Experiment; Experiment was dummy coded with Exp. 2 serving as the reference). Moreover, we tested how response bias and sensitivity for errors depended on the time on task (that is, page number, centered within each experiment separately; Δc:time; Δd':time) and how these effects differed between experiments (Δc:time:Experiment; Δd':time:Experiment). Finally, we tested planned contrasts for differences in the sensitivity for different types of errors within Exp. 1 and within Exp. 2 separately (using sliding difference contrasts nested within experiments). Moreover, random variation of response bias and sensitivity over subjects and target screens were included into the GLMM. The detailed results from this analysis are reported in Table F-1.

To test whether sensitivity for errors differed between error types, we fit alternative GLMMs lacking (a) any effects of error types, (b) the effects of error types within Exp. 1, and (c) within Exp. 2. These alternative (nested) models were compared to the original model via likelihood ratio tests, and all three models differed from the full model [(a) overall: $\chi^2(5) = 116.9$; $p < 7\text{e-}16$; (b) Exp. 1: $\chi^2(2) = 69.9$; $p < 7\text{e-}16$; (c) Exp. 2: $\chi^2(3) = 47.7$; $p < 3\text{e-}10$). Thus, sensitivity highly significantly differed between error types.

**Table F-1**

Detection of Inconsistencies

| **Fixed effects** | | **Estimate** | **S.E.** | ***z*** | ***p*** | |
|---|---|---|---|---|---|---|
| | | Bias (c) and Sensitivity (d’) | | | | |
| Experiment 2 | | | | | | |
| c | | -1.979 | 0.231 | -8.57 | < 2e-16 | *** |
| d’ | | 2.154 | 0.248 | 8.70 | < 2e-16 | ** |
| Experiment 1 – Experiment 2 | | | | | | |
| Δc | × Exp | 0.433 | 0.258 | 1.68 | .093 | + |
| Δd’ | × Exp | -0.251 | 0.280 | -0.90 | .370 | |
| | | Time on Task | | | | |
| Experiment 2 | | | | | | |
| Δc × Time | | 0.009 | 0.007 | 1.27 | .205 | |
| Δd’ × Time | | -0.014 | 0.007 | -1.94 | .052 | + |
| Experiment 1 – Experiment 2 | | | | | | |
| Δc × Time | × Exp | -0.012 | 0.008 | -1.45 | .147 | |
| Δd’ × Time | × Exp | 0.017 | 0.009 | 2.01 | .045 | * |
| | | Error Type | | | | |
| Experiment 1 | | | | | | |
| Δd’ × (semantic – discourse) | | 0.796 | 0.139 | 5.73 | 1e-08 | *** |
| Δd’ × (discourse – gibberish) | | 0.306 | 0.127 | 2.40 | .016 | * |
| Experiment 2 | | | | | | |
| Δd’ × (lexical – syntactic) | | 0.459 | 0.152 | 3.02 | .003 | ** |
| Δd’ × (syntactic – semantic) | | 0.039 | 0.146 | 0.26 | .791 | |
| Δd’ × (semantic – gibberish) | | 0.535 | 0.148 | 3.63 | .0003 | *** |

| **Random effects**<br>**Groups** | **Name** | **Std. Dev.** |
|---|---|---|
| subject | c | 0.264 |
| target screen | c | 0.216 |
| subject | d’ | 0.357 |
| target screen | d’ | 0.370 |

*Note*. Total number of observations: 1793. Groups: target screens, *N* = 62; subjects, *N* = 30. Significance codes: ***$p < .001$, **$p < .01$, *$p < .05$, +$p < .10$; significance tests are based on the wald-statistic and were checked using parametric bootstrapping, which confirmed levels of significance.

### *F.2.2 Eye movements: Determine valid gaze durations*

Fixations were assigned to lines of text, individual words, and letters, using elaborate algorithms implemented in the EDAS II software ("EDAS II," 2009). For text pages containing error sentences, the automatized assignments were checked visually and, if necessary, fixation assignments were corrected manually. To control for skimming, we analyzed for each trial the percentage of words in the error sentence that was not fixated by the reader (cf. Fig. F-1), and removed trials where less than 50% of words were fixated (4% of trials). First-pass fixations comprise all fixations on a word prior to making regressions back to this word or previous words in the text. Gaze duration is the cumulative duration of all first-pass fixations per word. From the eye movement recordings, we determined valid gaze durations using standard selection criteria (Table F-2; Table F-3). Across all 30 readers, a total of 20,498 words were included in the experiment within valid trials (see Methods section for details). Out of these words, 24.2 % were not fixated in firstpass reading, leaving 15,539 words for analyses. 39.3 % of the fixated words did not receive valid gaze durations according to standard criteria (see Table F-2, lines 4 to 9; Table F-3), resulting in a total of 9,435 valid gaze durations. We found a few more invalid gaze durations during mindless reading than during mindful reading or the control, but selectivity overall did not strongly depend on the state of mind.

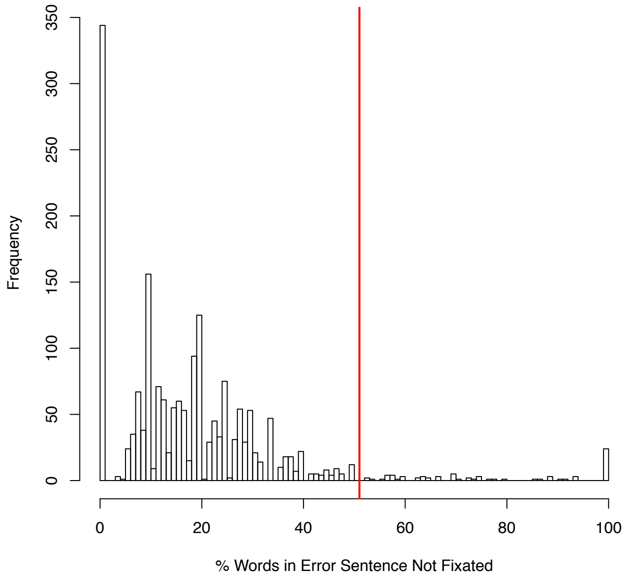


**Figure F-1**

Histogram for the percentage of words in the error sentence per trial that were not fixated.

**Table F-2**

Selection criteria used to determine valid gaze durations during mindless and mindful reading and a control condition, and the number and percentage of words in each condition.

| | Selection Criteria | Words | Total | Mindless Reading | Control | Mindful Reading |
|---|---|---|---|---|---|---|
| 1 | Valid Trials | N | 20498 | 5161 | 4999 | 10338 |
| 2 | Not fixated in Firstpass | N | 4959 | 1319 | 1221 | 2419 |
| | | % | 24.2 | 25.6 | 24.4 | 23.4 |
| 3 | Fixated in Firstpass | N | 15539 | 3842 | 3778 | 7919 |
| | | % | 75.8 | 74.4 | 75.6 | 76.6 |
| 4 | Line Not Fixated | N | 1769 | 475 | 404 | 890 |
| | | % | 11.4 | 12.4 | 10.7 | 11.2 |
| 5 | Invalid Fixations in Firstpass and Blinks | N | 1324 | 358 | 331 | 635 |
| | | % | 9.6 | 10.6 | 9.8 | 9.0 |
| 6 | Long Regressions (in/out) | N | 2465 | 592 | 612 | 1261 |
| | | % | 19.8 | 19.7 | 20.1 | 19.7 |
| 7 | First/Last Fixation in Line | N | 225 | 58 | 48 | 119 |
| | | % | 2.3 | 2.4 | 2.0 | 2.3 |
| 8 | Long/short Gaze Duration | N | 42 | 10 | 7 | 25 |
| | | % | 0.4 | 0.4 | 0.3 | 0.5 |
| 9 | Long/short Saccade (in/out) | N | 279 | 79 | 69 | 131 |
| | | % | 2.9 | 3.4 | 2.9 | 2.6 |
| 10 | Valid Gaze Durations | N | 9435 | 2270 | 2307 | 4858 |
| | | % | 60.7 | 59.1 | 61.1 | 61.3 |

*Note*. Row 1 = 2 + 3; row 3 = 4 + 5 + 6 + 7 + 8 + 9 + 10. Data are from 30 readers. Row numbers 4-9 indicate the order in which selection criteria were applied. Accordingly, the number and percentage of words for these criteria was calculated after excluding words based on previously listed criteria.

**Table F-3**

Definition of criteria used to determine valid gaze durations (cf. Table F-2).

| Criterion | Detailed description |
|---|---|
| Not Fixated in Firstpass | Words that were not fixated in firstpass reading |
| Fixated in Firstpass | Words that were fixated in firstpass reading |
| Line Not Fixated | Words in lines of text in which less than 50% of words were ever fixate |
| Invalid Fixations in Firstpass and Blinks | Words on which blinks started/ended in firstpass reading (detected manually or by the EyeLink 1000 system) and words with firstpass fixations manually marked invalid due to calibration problem |
| Long Regressions (in/out) | Words with long (> 20 characters) left-directed incoming/outgoing saccades (mainly return sweeps) |
| First/Last Fixation in Line | Words with First/Last Fixation in a Line |
| Long/Short Gaze Duration | Words with long (> 1500 ms) or short (< 50 ms) gaze durations |
| Long/Short Saccades | Words with long (> 20 characters) or short (< 1 character) incoming/outgoing saccades |

### *F.2.3 Global analyses*

For global analyses, we computed various eye movement measures. Initial fixation duration is the duration of the first fixation on a word (in first-pass), irrespective of later eye movements. Single fixation durations are the subset of initial fixation durations where a word was only fixated once in first-pass reading. Gaze duration is the cumulative duration of all first-pass fixations per word. Total reading time is the cumulative duration of all fixations on a word. The number of passes indicates how often a word is read in total. Average values of these measures of processing difficulty during mindless and mindful reading and the control are presented in Table F-4. The results indicate a small trend toward faster reading during mindless episodes in some variables like total reading time, word skipping, and regressions. For statistical testing we fitted (G)LMMs: each of the eye movement measures was regressed on an intercept, the difference between mindless reading and the control, and the difference between mindless and mindful reading. In addition, crossed random intercepts were used for subjects and words. Only one measure showed a significant difference between mindless and mindful reading: Readers made less passes during mindless as compared to mindful reading ($t$ = 5.0; see Table F-4). All other global eye movement measures did not significantly differ between mindless reading and mindful reading or the control ($Ps > .10$; $|ts| < 1.15$).

**Table F-4**

Global analyses: Differences in eye movements between mindless and mindful reading, and a control condition

| **Variable** | **Mindless Reading** | **Control** | **Mindful Reading** |
|---|---|---|---|
| Error | not detect. | none | detected |
| **Fixation Duration (ms)** | | | |
| Initial Fixation Duration | 256 | 257 | 257 |
| Single Fixation Duration | 256 | 258 | 257 |
| Gaze Duration | 298 | 299 | 299 |
| Total Reading Time | 353 | 355 | 359 |
| **Saccade Probabilities after Initial Firstpass Fixation (%)** | | | |
| Skipping | 20.8 | 20.2 | 19.9 |
| Refixations | 14.7 | 15.8 | 15.5 |
| Regressions | 12.1 | 12.7 | 13.4 |
| # of passes | 1.25 | 1.27 | 1.39 |

**Table F-5**

LMMs testing the influence of linguistic variables on gaze durations (untransformed and log10-transformed) during mindful, control, and mindless reading in experiments 1 and 2 on a 14-word interval prior to the error sentence.

| | | | **(A) Gaze Durations** | | | | **(B) Log10 Gaze Durations** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Fixed effects** | | | **Estim.** | **S.E.** | ***t*** | | **Estim.** | **S.E.** | ***t*** | |
| Experiment 2 | | | | | | | | | | |
| Overall effect (average of mindless reading, control, and mindful reading) | | | | | | | | | | |
| (Intercept) | | | 267.0 | 8.3 | 32.30 | *** | 5.519 | 0.228 | 199.5 | *** |
| 1/wl | | | -268.6 | 60.6 | -4.43 | *** | -0.923 | 0.164 | -5.64 | ** |
| freq | | | -12.0 | 4.1 | -2.97 | ** | -0.024 | 0.011 | -2.09 | * |
| (freq)/(wl) | | | 237.8 | 31.8 | 7.47 | *** | 0.596 | 0.084 | 7.13 | *** |
| Differences between mindful and mindless reading | | | | | | | | | | |
| (Int) | × (mindless – mindful) | | 9.7 | 6.7 | 1.44 | | 0.022 | 0.019 | 1.14 | |
| 1/wl | × (mindless – mindful) | | 154.7 | 80.3 | 1.93 | + | 0.320 | 0.228 | 1.40 | |
| freq | × (mindless – mindful) | | -6.0 | 5.1 | -1.17 | | -0.019 | 0.014 | -1.32 | |
| (freq)/(wl) | × (mindless – mindful) | | -175.0 | 37.9 | -4.61 | *** | -0.404 | 0.109 | -3.72 | *** |
| Difference between mindless reading and control | | | | | | | | | | |
| (Int) | × (control – mindless) | | -13.8 | 8.0 | -1.72 | + | -0.042 | 0.023 | -1.83 | + |
| 1/wl | × (control – mindless) | | -194.9 | 100.8 | -1.93 | + | -0.604 | 0.287 | -2.10 | * |
| freq | × (control – mindless) | | 8.9 | 6.4 | 1.40 | | 0.024 | 0.018 | 1.32 | |
| (freq)/(wl) | × (control – mindless) | | 91.9 | 47.2 | 1.95 | + | 0.259 | 0.135 | 1.92 | + |
| Experiments 1 and 2 | | | | | | | | | | |
| final(phrase/sentence final words) | | | -8.7 | 6.0 | -1.47 | | -0.028 | 0.017 | -1.67 | + |
| (final) | × (mindless – mindful) | | -18.0 | 9.5 | -1.90 | + | -0.043 | 0.027 | -1.58 | |
| (final) | × (control – mindless) | | 23.7 | 11.0 | 2.15 | * | 0.062 | 0.031 | 1.96 | * |
| Differences between Experiment 1 and 2 | | | | | | | | | | |
| Overall effect (average of mindless reading, control, and mindful reading) | | | | | | | | | | |
| (Intercept) | | × (Exp) | 9.5 | 5.8 | 1.65 | + | 0.024 | 0.017 | 1.45 | |
| 1/wl | | × (Exp) | 0.1 | 68.2 | 0.002 | | 0.193 | 0.192 | 1.01 | |
| freq | | × (Exp) | -6.0 | 4.2 | -1.44 | | -0.022 | 0.012 | -1.88 | |
| (freq)/(wl) | | × (Exp) | 72.2 | 33.6 | 2.15 | * | 0.142 | 0.095 | 1.50 | |
| Differences between mindful and mindless reading | | | | | | | | | | |
| (Int) | × (mindless – mindful) | × (Exp) | 5.5 | 9.0 | 0.61 | | 0.016 | 0.026 | 0.61 | |
| 1/wl | × (mindless – mindful) | × (Exp) | -154.2 | 110.3 | -1.40 | | -0.282 | 0.312 | -0.90 | |
| freq | × (mindless – mindful) | × (Exp) | 3.6 | 7.0 | 0.52 | | 0.019 | 0.020 | 0.98 | |
| (freq)/(wl) | × (mindless – mindful) | × (Exp) | 166.3 | 55.1 | 3.02 | ** | 0.369 | 0.157 | 2.34 | * |
| Difference between mindless reading and control | | | | | | | | | | |
| (Int) | × (control – mindless) | × (Exp) | 2.4 | 10.5 | 0.23 | | 0.020 | 0.030 | 0.67 | |
| 1/wl | × (control – mindless) | × (Exp) | 128.1 | 133.7 | 0.96 | | 0.372 | 0.378 | 0.98 | |
| freq | × (control – mindless) | × (Exp) | -1.3 | 8.4 | -0.15 | | -0.005 | 0.024 | -0.21 | |
| (freq)/(wl) | × (control – mindless) | × (Exp) | -56.6 | 66.1 | -0.86 | | -0.195 | 0.188 | -1.04 | |

**Random effects**

| **Groups** | **Name** | **Std. Dev.** | **Correlations** | | | **Std. Dev.** | **Correlations** | | |
|---|---|---|---|---|---|---|---|---|---|
| Words | (Intercept) | 46.1 | subject-varying | | | 0.125 | subject-varying | | |
| Screens | (Intercept) | 4.9 | | | | 0.023 | | | |
| Subjects | (Intercept) | 36.8 | Int | 1/wl | freq | 0.131 | Int | 1/wl | freq |
| | 1/wl | 112.6 | -.49 | | | 0.200 | -.38 | | |
| | freq | 7.1 | -.36 | .68 | | 0.020 | .06 | .90 | |
| | freq/wl | 71.8 | .81 | -.90 | -.69 | 0.118 | .37 | -.64 | -.52 |
| Residual | | 131.3 | | | | 0.375 | | | |

*Note*. Total number of observations: 9427. Groups: words, *N* = 490; screens, *N* = 62; subjects, *N* = 30. Significance codes: ***|*t*| > 3.291, **|*t*| > 2.576, *|*t*| > 1.96, +|*t*| > 1.645

### F.2.4 Local analyses

We aimed at assessing whether the influence of linguistic and lexical variables on fixation durations differed between mindless and mindful reading and the control condition. We retrieved word frequency norms from the dlexDB-database based on the *Digitales Wörterbuch der Deutschen Sprache des 20. Jahrhunderts* corpus (Geyken, 2006; Heister et al., 2011). We used a LMM to regress gaze durations on the fixed effects word length (1/wl), word frequency (log10 freq.), their interaction, and whether words occur at the end of a clause or a sentence (using dummy coding with non-final words as the reference), as well as the interaction of these variables with mindless reading [coded as sliding differences using the contr.sdiff() function of the MASS package (Venables & Ripley, 2002), where the intercept reflects the average over all mindless conditions]. We also tested whether the effects depended on the experiment number (using dummy coding with Experiment 2 as reference). To control for correlated error variance, random intercepts for words and target screens were used. In addition, we tested how the intercept as well as the effects of word length, word frequency, and their interaction varied over subjects, and found these random slopes to be reliable. The effect of final versus non-final words did not significantly vary between subjects and the random effect was dropped from the model. The results from this analysis are presented in Table F-5*A*.

To test the frequency effect for long and for short words separately, we centered the inverse word length variable at word lengths four and twelve for use in two post-hoc LMMs (Aiken & West, 1991). This was done by subtracting the inverse of word length 12 (or word length 4) from each value of the predictor variable. The modified word length variables were used as predictors in otherwise identical LMMs. Given that the interaction of word length and word frequency was included in the model, the main effect of word frequency now reflects the effect of word frequency at the specified word length, in our case at word length 12 (or 4), and thus provides a statistical test for the frequency effect among long (or short) words. As is visible in Figure 4-3*B*, for long words the frequency effect was strongly reduced during mindless reading (slope-difference to mindful reading for twelve letter words: $\Delta b = 15$, $SE = 5.6$, $t = 2.8$). For short words, to the contrary, the frequency effect was not significant during mindful reading (slope for four letter words: $t = 0.62$), but marginally significant during mindless reading ($t$ = -1.71; slope-difference: $\Delta b$ = -14, $SE$ = 5.9, $t$ = -2.3).

We also fitted the original LMM to log-transformed gaze durations to reduce problems with heteroscedasticity (Kliegl, et al., 2010). The results from this model are presented in Table F-5*B*. The main results from the model on untransformed data were replicated in this additional test of multiplicative effects.

### *F.2.5 Inferring Mindless Reading from Eye Movements*

#### Bayes' Theorem

Bayes' Theorem specifies how to derive the *posterior* probability P(H | E) for a hypothesis (H) after observing some empirical evidence (E).

$$P(H \mid E) = P(H) * P(E \mid H) / P(E) \qquad (1)$$

The posterior probability is determined based on the *prior* expectation P(H) that the hypothesis is true before any evidence is observed. P(E | H), often called the *likelihood*, is the probability of observing the evidence given that the hypothesis is true. Lastly, the probability P(E) to observe the evidence under any hypothesis is a normalizing constant, and is determined via

$$P(E) = SUM_i[P(H_i) * P(E \mid H_i)] \qquad (1a).$$

#### Estimating Probabilities in the Present Experiment

In the present experiment, we focused on the situation where eye movements should provide the strongest indicator for the state of mind. In a region comprising the 14 words prior to the error sentence, the state of mindless reading affected gaze durations particularly for long words (Figure 4-3B). We thus chose target words for analysis that were ten or more characters long. Gaze durations on these target words should reflect effects of lexical processing. Likewise, detection of lexical errors should be strongly driven by lexical word processing (Figure 4-5). Accordingly, we only analyzed eye movements before a lexical error was encountered. This strong selection resulted in a total of $N$ = 123 gaze durations that were made on 35 target words. While this selection went along with a strong reduction in the amount of available data, we focused the analysis on a condition where the effects are strong (and consistent with the general analyses; Figure 4-3B) and the state of mindless reading can be inferred from individual eye movements. For this specific subset of the data, we estimated Bayesian probabilities.

The *prior* probability for mindless reading, P(mindless), was estimated based on the rate with which readers overlooked any errors in Exp. 2, and was chosen as the likelihood of the intercept parameter in a simple logistic regression analysis. To obtain the *posterior* probability for mindless reading, P(mindless | gaze), we performed a Bayesian logistic regression analysis using the bayesglm program from the arm package (Gelman, et al., 2008), which is supplied in the *R* system for statistical computing (R Development Core Team, 2012).

# Supplement

## Erklärung

Hiermit erkläre ich, dass ich diese Arbeit selbständig, nur unter Mitwirkung der genannten Koauthoren, und ohne unzulässige Hilfe Dritter verfasst habe. Bei der Abfassung wurden nur die in der Dissertation angegebenen Hilfsmittel benutzt. Alle wörtlich oder inhaltlich übernommenen Stellen wurden als solche gekennzeichnet.

Ich erkläre zudem, dass diese Dissertation in der gegenwärtigen oder einer anderen Fassung nicht einer anderen Fakultät einer wissenschaftlichen Hochschule zur Begutachtung im Rahmen eines Promotionsverfahrens vorgelegen hat.

Ich erkläre zudem, dass ich an keiner anderen Hochschule ein Promotionsverfahren eröffnet habe.

Daniel J. Schad, Potsdam, den 29.10.2012