

UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Experience of Reading

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Philosophy

by

Alan Tonnies Moore

June 2016

Dissertation Committee:

Dr. Eric Schwitzgebel Chairperson

Dr. John Perry

Dr. Peter Graham

Copyright by
Alan Tonnies Moore
2016

The Dissertation of Alan Tonnies Moore is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

This dissertation often felt like an argument for panpsychism, a sentient creature with experience and intentions of its own, maturing in fits over six years. It owes its life to my dissertation committee. More than anyone, the support and guidance from Eric Schwitzgebel was indispensable. He helped me see through the stretches of nihilism, and in the process, showed me how weird philosophy can be a source of joy. John Perry's unswerving encouragement has made me, and many others, into sharper thinkers, but more importantly, he shows the world what true mentorship looks like. This is what makes a great philosopher. Thank you Peter Graham for the insightful comments, and you are right, the point of philosophy should not be a secret. As you can see, I've taken this to heart. There's no mistaking my reasons here.

I also want to thank the faculty at UC Riverside, the people who made an impersonal institution into an inviting place to learn: Howard Wettstein, Carl Cranor, John Fischer, and Sally Ness. Thank you to Lawrence Rosenblum, Charles Siewert, and Robin Jeshion for participating in my masters, proposition, and prospectus committees. I've also been incredibly lucky to have some amazing fellow graduate students: Megan Stotts for the friendship and place to stay, Megs Gendreau and Daniel Ehlich because they're my favorites, and many others. Without Gerardo Sanchez and Jayne Gales in the office, I doubt I would have filled

out all of the paperwork and forms that needed signing, often on short notice. In addition, I spent a wonderful year as a visiting researcher in Basque Country at the Institute of Logic, Cognition, Language, and Information at the University of Basque Country, or as it is properly called, Euskal Herriko Unibertsitatea. Thank you to the friends and colleagues I made out there: Kepa Korta, Josu Acosta, Eros Corazza, and Jerome Dokic.

The list goes on and on. My heartfelt love goes out to Mom and Dad, Vanessa Carlisle, Alex Rasch, Sasha and Bud Donahue, as well as Noah Peller and Thomas Dwyer for reading the final draft. Most of all, I am indebted to my partner, Charlotte Tonnie, for her continued support, year after year, for a project that often seemed more of like an abstract thought than a concrete piece of writing. Now we both get a break from the nights and weekends on my computer.

To everyone living a precarious life,
May you always find joy in the hustle.

ABSTRACT OF THE DISSERTATION

The Experience of Reading

by

Alan Tonnies Moore

Doctor of Philosophy, Graduate Program in Philosophy
University of California, Riverside, June 2016
Eric Schwitzgebel, Chairperson

What do you experience when you read? You are reading right now, so ask yourself, do you have an inner voice, visual imagery, or a perceptual experience of the words on the page. Perhaps, while you read this abstract, you simply find your mind wandering. Philosophers and psychologist provide radically conflicting descriptions of the commonplace experience of reading, and my own introspection feels incomplete and unclear.

I begin, in Chapter 1, by making the case for a healthy skepticism towards introspective reports, even those of experts describing their current experiences. In Chapter 2, I contrast the experience of reading with perceptual experiences and discuss the sources of introspective error: inference error, overgeneralization, and memory limitations. Due to these sources of error, we can not take introspective reports at face value, and I argue for the need to compare subjective and objective measures of experience using the Subjective and Objective Measures of Experience

(SOME) method. Chapter 3 is an examination of the cognitive processes that underlie the act of reading, and I discuss the phonological loop, the visuospatial sketchpad, situational models of narrative comprehension, flow, and mind wandering. In Chapter 4, I describe a series of experiments on the phenomenology of reading that use SOME method.

Chapter 5 is a general discussion of the results. These experiments suggest that there is extensive variability in the experience of reading. Experts have been blind to this variability because introspective error creates systematic biases in descriptions of experience. I argue that our confidence in introspective reports is always high, regardless of accuracy, so we must evaluate methods of measuring introspective reports, not the reports themselves. Further, these results suggest that coarse-grained aspects of experience, such as modal experiences while reading, are functionally isolated. They are what I call *paraphenomena*. When our experiences are not functionally isolated, the relationship between experience and behavior is often one of interference instead of facilitation. I finish by laying out the implications for the relationship between phenomenal and functional notions of consciousness.

TABLE OF CONTENTS

Preface	xvi
Chapter 1	
What's It Like to Read?	1
I. Guided Introspection	1
II. The (New) Hard Question of Consciousness	10
III. Descriptive Disagreements.....	18
IV. What's to Come	26
Chapter 2	
Concepts + Methods.....	30
I. Action Experience and Phenomenal Worms.....	33
II. Sources of Introspective Error.....	50
III. Methods of Studying Consciousness.....	62
Chapter 3	
Psychology + Reading.....	71
I. The Phonological Loop.....	75
II. The Visuospatial Sketchpad.....	83
III. The Visual System and Flow.....	95
IV. Mind Wandering.....	105
V. Applications and Optimism.....	109

Chapter 4

Experimental Phenomenology	111
I. Experiment One.....	115
II. Experiment Two.....	135
III. Experiment Three.....	142
IV. Experiment Four	153
V. Experiment Five	165
VI. Conclusion	175

Chapter 5

Phenomenal Variability + Functional Isolation	176
I. Introspective Reports	178
II. The Other Moore's Law	188
III. Subjective and Objective Measures of Experience.....	200
IV. Paraphenomena and Phenomenal Interference	205
V. Further Research	216
VI. Autobiographical Introspection.....	219
References	222

LIST OF FIGURES

Figure 3.1	89
The stimuli from Brooks (1967), including the spatial set of sentences, the nonsense set of sentences, and the four-by-four matrix used as a mnemonic device.	
Figure 3.2	99
Examples of the moving window and foveal mask techniques of measuring visual span from Rayner (1998). The asterisk represents the point of fixation.	
Figure 4.1	123
Histogram of general reports for an inner voice, visual imagery, and the words on the page (N = 57).	
Figure 4.2	124
The sum of the three concrete introspective reports for an inner voice, visual imagery, and the words on the page (N = 171).	
Figure 4.3	127
Histogram of the initial set of general reports for the experience of an inner voice, visual imagery, and the words on the page (N = 57).	
Figure 4.4	131
Histogram of positive concrete reports for the descriptive passage, the dialogue, and the abstract passage.	
Figure 4.5	150
Histogram of general reports of mind wandering (N = 228).	
Figure 4.6	159
Histogram of general reports of an inner voice, visual imagery, words on the page, and mind wandering (N = 1,236).	
Figure 4.7	159
Histogram of concrete reports of an inner voice, visual imagery, words on the page, and mind wandering.	
Figure 4.8	169

An example of a mental folding question. Although this one is quite easy, the most difficult questions required up to three non-symmetrical folds. The correct answer is “A”.

Figure 5.1179

Percentage of participants who provided general reports of inner voice, visual imagery, words on the page, and mind wandering from all five experiments.

Figure 5.2180

The combined concrete reports for Experiments 1, 2, 3, and 5, presented as a percentage of each type of report.

Figure 5.3180

The combined concrete reports from Experiments 1 and 2, separated into three categories: the percentage of participants who consistently reported an experience, who consistently denied an experience, and who reported a change in experience.

LIST OF TABLES

Table 4.1	119
The question text for general and concrete introspective reports.	
Table 4.2	126
The number of participants who reported experiences at the extremes. For general introspective reports, these are the participants who responded that they always, or never, have a certain experience. For concrete reports, these are the participants who reported the same experience in all three probes (N = 57).	
Table 4.3	129
Table of two-tailed Pearson correlations between the three sets of introspective reports (N = 57).	
Table 4.4	133
Mean accuracy for the objective measure of experience.	
Table 4.5	134
Table of two-tailed Pearson correlations between subjective and objective measures of experience. None of the results were significant ($p < .05$).	
Table 4.6	139
Mean accuracy for the objective measure of experience.	
Table 4.7	140
Table of Pearson correlations between introspective reports and performance on objective measures of experience.	
Table 4.8	141
Stepwise regression models for predicting objective measures of experience using the corresponding introspective reports and reading comprehension.	
Table 4.9	147
The question text for the two sets of general introspective reports.	
Table 4.10	151
Table of independent samples t-tests on the difference in mean reports for the directed attention conditions and the control group.	
Table 4.11	151

Mean accuracy for the objective measure of experience.	
Table 4.12	153
Table of Pearson correlations between introspective reports and performance on objective measures of experience. The negative correlation between reports of visual imagery and the Stroop Task was expected.	
Table 4.13	156
The question text for general and concrete introspective reports.	
Table 4.14	160
Mean accuracy for the objective measure of experience.	
Table 4.15	161
Table of Pearson correlations between introspective reports and performance on objective measures of experience. None of the results were significant ($p < .05$).	
Table 4.16	162
Table of Pearson correlations between introspective reports and performance on comprehension questions. The negative correlations for reports of mind wandering are expected.	
Table 4.17	162
Stepwise regression models for predicting objective measures of experience using the corresponding reports and reading comprehension.	
Table 4.18	164
Table of Pearson correlations between reports of mind wandering and performance on objective measures of experience.	
Table 4.19	164
Stepwise regression models for predicting objective measures of experience using reports of mind wandering and accuracy on reading comprehension questions.	
Table 4.20	168
The question text of general and concrete introspective reports. Unlike the previous experiments, the initial and final set of general reports use the same wording.	
Table 4.21	170
Mean accuracy for the objective measure of experiences.	
Table 4.22	171

Table of Pearson correlations between introspective reports and performance on objective measures of experience. Unexpectedly, there is a negative correlation between general reports of visual imagery and performance on the Folding Task.	
Table 4.23	173
Table of Pearson correlations between introspective reports and performance on comprehension questions.	
Table 4.24	173
Stepwise regression models for predicting objective measures of experience using the corresponding introspective report and reading comprehension.	
Table 4.25	174
Table of Pearson correlations between performance on objective measures of experience and introspective reports of mind wandering.	
Table 4.26	175
Stepwise regression models for predicting performance on objective measures of experience using reports of mind wandering and reading comprehension.	
Table 5.1	203
The two meta-analyses of the effect size and p values of the correlations between subjective and objective measures of experience.	
Table 5.2	204
The two meta-analyses of the effect size and p values of the correlations between reports of mind wandering and objective measures of experience.	

PREFACE

I began with a question: What do people experience when they read. I turned to the experts, philosophers and psychologists who have written books and peer-reviewed articles on consciousness, to see what they had to say. I was surprised to find a range of descriptions, often standing in direct conflict with each other, nevertheless given with a haughty confidence and categorical certainty as if the answer were obvious to everyone. When I reflected on my own experiences, the answer appeared clouded, as if gazing into a muddied stream. To this day, it reminds me of Jorge Borges' attempt to summon a tiger in his dreams:

Oh, incompetence! My dreams never seem to engender the creature I so hunger for. The tiger does appear, but it is all dried up, or it's flimsy-looking, or it has impure vagaries of shape or an unacceptable size, or it's altogether too ephemeral, or it looks more like a dog or bird than like a tiger (Borges, 1975, p. 294).

I ask you, when you read this quote just now, were you having a visual experience of the words on the page? To me, the answer feels incompetent. My introspection never seems to engender the experience I so hunger for.

I could not find the answer by looking within, and there were widely divergent and conflicting descriptions from others. I needed another source of information, something beyond my own reports and those of the experts. In this way, what began as an intuitive question turned into a piece of experimental

philosophy. Before long, I was engaging with traditional issues in philosophy, such as the extent of our self-knowledge, the relationship between phenomenal and functional notions of consciousness, and the role of descriptive phenomenology throughout the discipline. I also found myself wading deep into areas of cognitive science, from the experimental methods for obtaining introspective reports to the cognitive mechanisms behind the processing of phonological and visuospatial information in a text. I did not think I would end up here, but this is where the path lead. It's refreshing to have found a natural entrance to these academic issues.

My hope is that, if you gain nothing else from this dissertation, you come to have a better understanding of your own experiences while reading. Although I spend most of my time on the theoretical and methodological implications of this research, I care more about the way it can be integrated into daily life. So please, treasure your inner voice, for not everyone has one. Imagine the possibilities open to your visual imagery; fantastical experiences await. The text you are reading is a point of departure, and once you have left the constraints of the perceptual world, you can begin a phenomenal odyssey to wherever your dream tigers reside, imperfect or vaguely defined as they may be.

<3

Alan Tonnie Moore

Oakland, California

March 25, 2016

Chapter 1

WHAT'S IT LIKE TO READ?

As is often the case, today I find myself reading. I am not reading anything important, and I am not overly invested in the content. I am perusing. Perhaps because of my general level of disinterest, it occurs to me that reading is something I do every day, yet I have thought little about my experiences of reading, the “what it’s like” to read. Maybe this is because the answer is obvious. I do not, after all, often think about which way is up. A simple question should settle the matter.

I ask you, as I ask myself, what sort of experiences do you have while reading? You are, in fact, reading at this very moment, so think, what are you experiencing right now? I do not mean this merely as a rhetorical move or an idle query. Take me at my word — for a brief moment, turn your attention inward, to your current conscious states, and lay some fresh eyes upon the flowing content of your stream of experience.

I. GUIDED INTROSPECTION

Are you experiencing some form of visual imagery, perhaps something analogous to a mental slideshow or movie? Daniel Dennett often does, and says “the act of reading, and interpreting, a text such as a novel creates some new things in my imagination: images of the characters doing the deeds” (1991, p.

366). This sentiment is echoed by Ludwig Wittgenstein, who says “[while reading] I have impressions, see pictures in my mind’s eye, etc. I make the story pass before me like pictures, like a cartoon story.”¹ For my part, I feel the pull of these descriptions while reading “stories.” However, this paper is not a narrative in that sense — it is not a story — so maybe this is an inadequate description of your present experiences. After all, visual imagery could take many shapes and forms, extending far beyond the series of images we often associate with narratives in a visual medium such as film. In fact, Charles Siewert says “I enjoy highly variegated visual experience throughout the course of an average day, even while I am reading or talking” (2011, p. 4). E. B. Titchener gives us a vivid illustration of how variegated experience can be. While claiming to have ubiquitous visual imagery, he says “I instinctively arrange the facts or arguments in some visual pattern, and I am as likely to think in terms of this pattern as I am to think in words” (1909, p. 10). He even describes one of these mental patterns as “a suggestion of dull red... of angles rather than curves... [and] pretty clearly, the picture of movement along lines, and of neatness or confusion where the moving lines come together (1909, p. 12).

On hearing Titchener’s description, I begin to think about the numerous texts I have come across in my life: the novels, the philosophical articles, the

¹ This quote comes from the Zettel (1946-1948/1975, p. 44), a collection of Wittgenstein’s notes published posthumously. It is not clear if Wittgenstein endorsed this view or merely found it interesting enough to jot down one day. Regardless of whether or not his official position, the quote is illustrative of an experiential possibility.

newspaper pages, the enigmatic operating instructions, the magazines, the cereal boxes, tattoos, and all of the words that flash across my computer screen daily. The attractiveness diminishes; surely I do not have visual imagery while reading all of these disparate texts. Edmond Burke expresses a similar skepticism and claims that while reading “a very diligent examination of my own mind, and getting others to consider theirs, I do not find that one in twenty times any such picture is formed” (1757/1990, p. 152). Similarly, George Berkeley says that we often read without giving rise to a particular “idea,” and for Berkeley, “ideas” and mental images are synonymous. He gives two examples of imageless terms, “good thing” and “Aristotle,” going on to say that “innumerable examples of this kind may be given.” For Berkeley, this is simply a basic, obvious fact about our experiences, and asks “why should I insist on those things which everyone's experience will, I doubt not, plentifully suggest unto him?” (1710/2009, Intro, Section 20). Whether or not these descriptions capture your experience, they suggest interesting ways it could be. So ask yourself again, are you experiencing visual imagery while reading this page?

Broadening the phenomenal possibilities to include other modalities, perhaps you are experiencing inner speech at this moment. Bernard Baars is confident that you are, claiming “human beings talk to themselves every moment of the waking day. Most readers of this sentence are doing it just now” (2003a, p. 7). Although Baars is a preeminent American psychologist, he provides no empirical support. Like Berkeley before him, Baars takes his experience of reading

to be a universal and self-evident truth. In emphasizing the auditory dimension of consciousness, Baars is in company with other influential psychologists such as William James, who describes reading as “an uninterrupted and protracted recall of sounds by sights which have always been coupled with them in the past” (1890, p. 361). Alain Morin, another psychologist, makes an even stronger claim, saying “all aspects of normal language functions (e.g., reading, writing, speaking, and calculating) require intact inner speech” (2009, p. 395). Psychologists are not the only proponents of pervasive inner speech while reading, and the philosopher Max Velmans gives an exacting description, going as far as to tell us which syllables our inner speech will stress (and not stress) while reading the two instance of “refuse” in the sentence “If we do not increase the dustmen’s wages, they will refuse to take the refuse” (2009, p. 249).

Is inner speech a pervasive aspect of your experience while reading? R. S. Woodworth describes times in which he reads with “clear consciousness of a particular thought, and no [auditory] images” (1906, p. 704). This sentiment was echoed in the early 20th century psychology labs aligned with Wilhelm Wundt in the famous imageless thought controversy. Out of decades of experimentation, there are numerous examples of subjects reading without reporting an experience of inner speech (Thomas, 2012). Further, to complicate his previous statement above, Morin goes on to give an example of someone with normal cognitive functioning who lacks inner speech, explaining this by claiming that he uses visual imagery as a compensation strategy. If you are experiencing inner speech while

reading this paper, does it feel so obvious? And if not, does it feel like a cognitive or experiential deficiency?

Now consider your perceptual experiences. While reading, do you have an experience of the black words on the white page, or are the words, in a sense, transparent? Julian Jaynes tells us that “right at this moment... as you read, you are not conscious of the letters or even of the words or even of the syntax, or the sentences, and punctuation, but only of their meaning” (1976, p. 26), suggesting that we consistently get *lost* in the content of the prose. Siewert, on the other hand, responds that “this extreme denial of visual consciousness, once made plain, [is] very strange, and just about as obviously false a remark as one could make about visual experience” (1998, p. 249). Siewert argues that we do not dwell on each individual letter, yet skilled readers are nevertheless visual conscious of “vaguely identifiable groups of characters.” And yet Jayne’s basic idea is endorsed by the psychologist Russ Hurlburt, who argues that the retinal representations of the words on the page takes place outside of consciousness, and that “the actual experience while reading... has little or nothing to do with such things” (Hurlburt & Schwitzgebel, 2007, p. 50).

Of course, these possibilities are not mutually exclusive, and you could be having a complicated, cross-modal experience. Returning to the ornate descriptions of experience given by Titchener, he claims that, in addition to his colorful visual imagery, he also has “vivid and persistent auditory imagery,” saying “my natural tendency is to employ internal speech; and there are occasions when my voice rings

out clearly to the mental ear and my throat feels stiff as if with much talking” (1909, p. 9). In fact, Titchener’s reported auditory experiences extend beyond that of an inner voice; he reports pervasive inner music as well. Hurlburt also conveys this type of dual experience through the words of Melanie, one of his experimental participants, saying that some people “apparently simply read, comprehending the meaning without images or speech. Melanie’s general view... is that she starts a passage in inner speech and then ‘takes off’ into images” (2007, p. 101).

Although these three options — visual imagery, inner voice, and a visual experience of the words on the page — are the most commonly expressed (or denied) experiences in the academic literature, they do not determine the phenomenal bounds of inner life while reading. Consider the following passage from the philosopher William Robinson, describing his experience of reading the sentence “the boy the man the girl saw chased fled.”

For example, upon reading [this sentence] for the first time, my feeling was indeed the predicted “Huh? Does this make sense? What is it saying happened?” After the explanation, I had a mental image, namely, of a man chasing a boy with a girl looking on. This was followed by a ‘correction’ (to reflect the past tense form, ‘saw’), namely an image of a girl looking through a doorway at a man (to visual right, as it happened) and another image of a man chasing a (running) boy. There was also a vague sense of relief or relaxation, as of a problem solved. Finally, after the explanation, my subvocal utterance of (1) had a ‘phonological shape’ (pauses, emphases and tonal modulations) that it lacked when first read. (Robinson, 2005, p. 549)

Of the experiences described in this vignette, visual imagery and inner voice are only a part. What about the others, the feeling of confusion, relief, or relaxation?

They are all there, in Robinson's experience, and they occur simultaneous with more commonly cited modal experiences. Is there a principled way to separate some of these experiences as uniquely *reading experiences*? What about becoming aware that the clock has been striking while reading (Blackmore, 2002, p. 24)? What about mind wandering? Do these possibilities constitute the discrete building blocks of the experience of reading, some of which we can rule out or disqualify, or do the individual strands form a braid that is too intertwined to parse out, woven into a sturdy phenomenal fiber?

This is where the current literature leaves us, deep in a muddled question bog. We have, with few exceptions, a collection of armchair claims about the experiences of one individual that is then generalized to apply to all. Casual introspection will not settle the matter, otherwise the truth would be self-evident. The testimony of experts does not settle the matter, for experts have already failed to find agreement.

I have spent years asking "what's it like to read," and my difficulty providing an adequate response feels embarrassing. Perhaps, after the previous section, you feel a portion of this embarrassment yourself. So let's back up, away from the experience of reading, and consider consciousness in a more general form. "What's it like to read" would seem to fall among the easiest category of questions: things you are experiencing right now. For a moment, think about the vibrant range of experiences that fit this description. Some are straightforwardly

perceptual, either concerning your past, current, or counter-factual experiences of perception:

What does your face look like in the mirror?

How does itchy salt water feel on on your skin?

What would a radio sound like playing in the apartment next door?

Others are perceptual-like, that is, they are not perceptual experiences but are constructed out of the same phenomenal materials:

What happened in your dream last night?

What color were the metal elves you saw last time you took LSD?

Still other experiences go beyond the limits of perception and concern propositional thoughts, emotional experiences, and experiential blanks.

What is it like to think “David Lewis couldn’t possibly have believed his own account of possible worlds.”

What is a nagging feeling of boredom?

What is the bodily feeling of anger?

What is the phenomenal blank associated with spacing out?

As we move further away from perceptual experiences, I start to lose my grasp on — not the questions themselves — but how to answer them. Even something as difficult to describe as the olfactory experience of fresh-baked cookies seems more concrete than the phenomenology associated with the feeling of confidence or the experience (or lack thereof) associated with spacing out. Putting these tough cases aside for the moment, even the questions concerning perception or perception-like

experiences, the ones that are supposed to be answerable, even obvious, have led to a fascinating, confusing, and contradictory picture of our conscious lives.

Consider the ways people have described the experience of seeing a round object at an obtuse angle, perhaps a plate lying on the table in front of you. First, take a moment to think about it. Put a plate in front of you. Introspect: what does it look like? Would you describe the visual experience as an ellipse, after all, an elliptical object would occlude it from view, or would you describe it simply as a round object on its side? John Locke (1690/1975) and David Hume (1739/1978) argue the former, A. J. Ayer (1940) and Charles Siewert (2006) the latter, H. H. Price (1932) thinks it depends on how far away the object is, Michael Tye (2000) and Alva Noë (2004) argue that both descriptions are correct, and Sean Kelly (2008) holds that we shift between the two. Whatever initial plausibility to the claim that it looks elliptical, this description is bizarre in light of the fact that, contrary to three hundred years of introspective reports, there is no elliptical retinal projection involved (Schwitzgebel, 2011, ch. 2).

These are all answers to question about a single, rather simple, perceptual experience. Not to fall into cynical despair, but the disagreement here is rather troubling. Experience might *often* be ineffable, but the visual experience of a plate seems to be terrain that is as favorable toward introspection as it gets: the description of a uni-modal perceptual experience, with ample time to reflect, given by men who have devoted decades to studying consciousness. Here we are, with a prestigious group of experts — all of whom have written peer-reviewed articles

and academic books on consciousness — giving us five different descriptions of a plate.

Examples like this are numerous. Introspection is tricky business. Examine the details of your current visual experience. Is your peripheral vision detailed or fuzzy (Dennett, 1991, ch 3)? Ask yourself if you had tactile sensation of the shoe on your left foot two seconds ago (Schwitzgebel, 2011, ch. 6). Of course you feel your foot in your shoe now, but what did you feel before attending to it? Now think about your emotional or physical states. Are you sad, happy, excited, hungry? If you are like me, these questions are frustratingly difficult to answer, and this is especially surprising given their simplicity. These are not abnormal experiences, something only found under artificial conditions in a psychology lab or a thought experiment. These are a wide sampling of the experiences we have on a daily basis. How can knowing the answers be so *hard*?

II. THE (NEW) HARD QUESTION OF CONSCIOUSNESS

This is the new “hard” question of consciousness: what do you experience, even right now? In this dissertation, I argue that there is widespread bias in our introspective reports, even among experts on consciousness, from the philosopher’s armchair to the psychologist’s lab. Further, I will make the case that pervasive and salient aspects of experience exist disconnected from the way we process textual information. Not only is your inner voice, if you have one, causally impotent, it may not even covary with phonological processing. In arguing for these

conclusions, I will be frequently invoking terms of art, ones rooted in common sense, but terms of art nonetheless. Here I will clarify some of these concepts and survey the general landscape. In the process, I will give examples of how poor our self-knowledge truly is.

Since introspection is our sole epistemic access to experiences, we first need to have a basic idea of what introspection is. Although I am not committed to a particular model of introspection, I do rely on a family of views that treat introspection as a self-monitoring process (e.g. Armstrong, 1968; Goldman, 2006; Lycan, 1996; Nichols & Stich, 2003). I am something of an introspective pluralist, and believe that, at bottom, introspection is any self-monitoring process with the aim of forming knowledge of our conscious states (Schwitzgebel, 2011). Now, there are many self-monitoring processes in human cognition, many of which have nothing to do with consciousness. I could monitor the movements of my hand using my visual perceptual system by simply looking at it. I could monitor the quality of my new haircut by observing the reactions of others, operating on the assumption that only a truly spectacular or horrendous haircut will get noticed. These are all self-monitoring processes, but because they are not directed at experience, they are not examples of introspection.

The classic model of introspection is an intellectual, serial process. It involves an intention to know what you are experiencing right now, the formation of a judgment, and then forming a corresponding propositional attitude, such as “I am experiencing a visual perception of my white computer screen.” This is more of

a caricature of introspection rather than the dominant or limiting case. Although it is not central to this paper, it is helpful to understand what I take to be the varieties of introspection. Of course, you could introspect your emotional states in the classic way, but you could also do it by observing your own behavior, perhaps learning that you are depressed by noticing that you have not left the bed all day. You could introspect by reading your own stream of consciousness prose. You could introspect your poorly detailed periphery vision by performing Dennett's classic playing card experiment (1991, p. 53).

At least, this is how we come to know our own conscious experiences. Because we can not directly measure the experience of others, or even the raw, unexpressed output of introspection, there is an extra step involved when we study consciousness in the third person. I call the product of introspection *introspective judgments*, and *introspective reports* are the articulation of these judgment. They can be verbal or written, linguistic or isomorphic. Some examples of introspective reports: writing "I'm experiencing mental imagery," saying "I'm experiencing mental imagery," drawing the visual imagery you are experiencing, singing the auditory imagery you are experiencing, and evoking metaphors to describe your olfactory experience. Just as with self-knowledge, learning about experience in others is a multi-layered process. Introspection aims at knowledge of conscious experiences, and introspective reports are the articulation of the introspective judgments that result.

Any time one cognitive process tracks another, there is room for error, and introspection is no different. The conditions under which introspection is more or less reliable is an empirical question, and from decades of experimental work we have extensive evidence that suggests a skeptical stance towards a wide range of self-reports. In their famous study, Nisbett and Wilson (1977) showed that we often do not know why we make the judgments we do, and even more alarmingly, that we are often blind to our own ignorance. Subjects given four identical stockings or nightgowns devised elaborate explanations for why they preferred one over the rest, often referring to specific details, such as texture, color, or quality. Despite the objective sameness of the articles of clothing, subjects reported radically different experiences of them. This set off an explosion of research regarding knowledge of our own cognitive processes, a survey of the limits to whatever core infallibility our self-knowledge possesses. Nisbett and Wilson are commonly read as maintaining that we are mistaken about everything going on in our heads, be it conscious, unconscious, or the underlying computation, but they actually argue for a rather narrow form of skepticism concerning *knowledge of cognitive processes*. For them, cognitive processes are distinct from “mental content,” such as feelings, sensations, perceptions, and judgments. Mental content, as they claim, can be known with “near certainty” (p. 255).

In the subsequent decades, the empirical data on our poor self-knowledge has been expanding. To get a feel for the many ways that we are routinely wrong

about ourselves, consider the exploding research on implicit biases and moral judgments.

Are you racist, or more carefully, do you have a disposition towards racist behavior? The answer is probably yes, although it is unlikely that you are aware of it. In study after study, participants who wholeheartedly and sincerely claimed they were not racist nevertheless exhibited racist behavior in an experimental setting. From the research on implicit biases, in this case, implicit racism, we know that the underlying cognitive process is both automatic and unconscious. In a classic study, subjects were primed by showing them photos of black and white faces. These photos were followed by an adjective with a positive valence (e.g. attractive, likable, and wonderful) or a negative one (e.g. annoying, disgusting, and offensive), and they measured each subject's reaction time in categorizing them as such. White subjects were fastest at categorizing negative words after being primed with a black face, and conversely, positive words after being primed with a white face. Because these results did not correspond with other tests of racism that looked at conscious behavior, such as racial attitudes, this experiment suggests that implicit racism is an automatic process that floats free from our conscious judgments about race (Fazio, Jackson, Dunton, & Williams, 1995).

Other studies have looked at our awareness of this automatic process, that is, how conscious we are of our implicit racial biases. Consider the results from Greenwald, McGhee, and Schwartz (1998), in which subjects were shown a name that they had already identified as white or black and then asked to categorize an

adjective as either pleasant or unpleasant. Subjects were significantly faster at identifying negative adjectives after a black name and positive adjectives after a white name. A close look at the results show how confounding this outcome is. In subsequent questioning, nearly 3/4 of the subjects disavowed having any racial prejudices against blacks, yet the fact remains that, of these nineteen subjects, all but one showed a significant positive association with whiteness. As this study shows, people often have a racial bias against black people that they are entirely unaware of it, even sincerely disavowing its existence. These results are robust across numerous experiments. Despite what we might think about ourselves, we have unconscious and automatic racial prejudices that we are entirely ignorant of (Devine & Monteith, 1999; Dovidio & Gaertner, 2004; Greenwald & Banaji, 1995). Extending the idea further, the general effect at play, implicit bias, applies in other situations as well, such as in implicit gender bias (Banaji, Hardin, & Rothman, 1993). Edouard Machery (manuscript) has circulated the strongest critique of this literature, arguing against explaining the results in terms of bias. Instead, Machery argues for a trait picture of attitudes, but for my purposes, this distinction does not matter. Machery's traits are just as unknown to their owners as implicit attitudes, and it is this ignorance of a piece of self-knowledge that I am interested in.

The existence of implicit racism is well established, but it bears saying a little more about its importance. Negative racial association, although unconscious and, in the most cited studies, affecting nothing more than response time, is exactly

the sort of psychological pattern that contributes to wider social issues such as police killings. When the decision to use lethal force is made on a snap judgment, the immediacy and speed of the association between black faces and negative concepts could be the deciding factor between life and death. Indeed, when shown pictures of people holding a variety of objects, subjects were considerably more likely to mistake a harmless object for a gun when the person holding it was black (Payne, 2006). Most people do not know that they have this seed of racism within them. Note that, you do not learn about your own implicit racism through classical introspection of your belief states. Instead, you need to apply the methods of psychology.

Experimental philosophy, as a methodological and sociological movement in philosophy, became popular due to brilliant work on the source of our moral judgments (e.g. Green, 2008; Haidt, 2001). Prior to this body of research, it was relatively uncontroversial that moral judgments were the logical application of moral principles to a given situation. On this picture, something is right just if it maximizes happiness, or results from duty, or is a virtuous action. Sentimentalism, the view that moral judgments are determined by emotional responses, had few followers. However, a series of studies have shown a remarkably tight relationship between morality and emotion, such as a high correlation between disgust and moral judgments (Haidt, 2001; Haidt, Rozin, McCauley, & Imada, 1997), the odd inability for participants to give rational justification for moral judgments (Haidt, Bjorklund, & Murphy, manuscript), and the activation of affective parts of the

brain when making moral judgments (e.g. Greene & Haidt, 2002; Moll, de Oliveira-Souza, & Eslinger, 2003). As these studies suggest, many of our moral judgments are driven by the emotional reaction to a situation, not by a line of reasoning that extends from moral principle to reflected judgment. I only wish to gesture at the bigger empirical and philosophical issues, to put the recent work around moral judgments into a larger context of our own self-ignorance. What is the cause of our moral judgments? Whatever the answer, we are missing a crucial piece if we solely rely on introspection of our own decision making processes.

In the research on implicit racism and the origin of moral judgments, classical introspection and psychological experimentation lead to starkly different conclusions. Both groups of studies call into question our knowledge of the mental processes behind the judgments we make about our beliefs and desires, but what about these judgments themselves? Or sensations, perceptions, emotions, pains — all of the things that Nisbett and Wilson claim we know with near certainty? After all, skepticism regarding knowledge of the cognitive processes that underly our experiences is different from skepticism regarding knowledge of our experiences themselves.

This final, radical skepticism is the deepest and most fundamental form of introspective skepticism. It calls into question a pervasive domain of self-knowledge, from what we are seeing in front of us to what we are feeling right now. Putting aside philosophical scenarios such as a brain in a vat or an evil deceiver, it is radical indeed to claim we are wrong about such daily going-ons. Yet

skepticism is not the same thing as unreliabilism. A skeptic does not have to think that we are always mistaken about the contents of our mental life, only that we are often mistaken, or maybe just that we are wrong more often than we realize. As I see it, radical introspective skepticism is more of an orientation than a definitive view, a way of engaging with introspective reports that does not take them, any of them, at face value. Rather, the reliability of reports is an empirical matter, best approached using a variety of methods and tools, and it is to these that we now turn.

III. DESCRIPTIVE DISAGREEMENTS

In order to outline the limits of introspective knowledge, it is useful to look at areas of introspective disagreement. What follows is a brief history of introspection, a cautionary tale of what can happen when experts turn their gaze inward and provide different descriptions of what they see.

For much of the 20th century, consciousness was rarely studied outside of anglophone philosophy departments. It is something of a historical accident that consciousness fell under the domain of philosophy, a result of the behaviorist revolution that exiled experience from the labs of psychologists. There has always been a virtuous philosophical passion for hard questions that have been orphaned by other institutions, and when combined with a history of writing on consciousness that can be traced back to Aristotle, philosophy was a natural fit. From the inception of analytic philosophy, as championed by G. E. Moore and

Bertrand Russell, armchair reasoning has been the dominant approach. Given that Moore and Russell were interested in tightening up the concepts we employ in everyday language, translating them into a more precise and regimented logical form, the focus on armchair reasoning is entirely reasonable. Parsing through linguistic meaning, in the way that early analytic philosophers were interested in, does not require empirical investigation. Simply knowing the meaning of the words is enough to follow the arguments from premises to conclusion. The armchair reasoning they were engaged in was, if not a priori, tethered to it on a short rope. Consciousness lies in the gray area around clear examples of a priori reasoning. After all, consciousness simply is experience, thus precluding it from being a priori on some views. Yet being a priori and being the result of armchair reasoning are two different things, one metaphysical, the other epistemological. Because armchair reasoning is an informal methodology, it is not entirely clear what it amounts to. I view it as an archetype more than a clearly defined experimental method. Typically, armchair reasoning floats free from the formal scientific standards of hypothesis testing, control groups, and modern statistical analysis. It involves a minimum of empirical work, and when experimental results do find their way in, they are merely premises in a larger argument.

As I said, armchair reasoning dominated anglophone philosophy departments and the study of consciousness that occurred inside. Approaching consciousness using armchair reflection is understandable. Armchair reasoning works well for so many other difficult and orphaned issues: conceptual analysis,

exegesis and interpretation in the history of philosophy, metaethics, and the study of formal logical systems. Although the recent influx of empirical and experimental work has greatly improved our understanding of these issues, this should not be taken as a refutation of armchair philosophy. Indeed, armchair reasoning has a distinguished history with a long list of accomplishments.

Accurate introspective reports, however, do not appear to be among them. We have been discussing the diabolically simple question of what it's like to read. We are not talking about a complicated philosophical theory or finer points of consciousness, such as the perception of difference tones or dream states, but rather something basic, maybe even fundamental, and presumably something that we should just *know*. Yet, as obvious as many people believe their claims about experience to be, the armchair methodology has not been producing results upon which people can agree.

Why has an experience as pedestrian as reading — or whether there is imageless thought, or what a plate looks like on a table — led to such a menagerie of descriptions? If there are disagreements here, there must be dozens of other such impasses. Experience seems as if it should be a straightforward and easy phenomena to study, yet these disagreements are pervasive and deep. In addition to the experimental work that calls into question our knowledge of cognitive processes, implicit biases, and the origin of moral judgment, there is a growing body of philosophical and experimental work that casts doubt on the accuracy and reliability of introspection itself.

It bears pointing out that even in the heyday of introspective psychology, untrained introspection was largely viewed with a skeptical eye, and subjects often had to undergo months of training before they were considered to be proficient introspectors. Wilhelm Wundt required his subjects to perform ten thousand laboratory introspections before he would use their data, and E.B. Titchener's sprawling 1,600 page lab manual spells out detailed introspective training exercises, ranging from listening for difference tones to describing after-images. This extensive training was thought to hone and calibrate introspection, a faculty that was viewed with a skeptical eye by the same psychologists who built the foundation of experimental psychology upon it (see Schwitzgebel 2004, 2005 for a detailed look at introspective training).

The introspective psychologists were devotees to one particular experimental approach to consciousness, but this is not the only, or even the most influential, orientation toward consciousness in the 20th century. That distinction goes to the behaviorist's philosophical and methodological rejection of consciousness as an appropriate topic for scientific study. The origin of behaviorism as a self-conscious movement can be traced back to John Watson's (1913) influential behaviorist manifesto. In this and subsequent writing, Watson directly criticized introspective reports, specifically those relating to mental imagery, with theatrical epithets such as "old wives tales," "savagery," and "sheer bunk" (Watson, 1928). During the half century that behaviorism dominated psychology in the United States and abroad, the experimental approach to

consciousness was marginalized. Research on consciousness continued, but it was relegated to fringe issues, such as the study of dreaming, and the results languished in minor journals.

During the 1990s, research on consciousness picked up, this time in cognitive science labs as well as philosophical armchairs. In a relatively short period of time, consciousness once again found itself into the top departments and placed in the top journals. Far from a radical rejection of inner life, this renaissance brought about a radical acceptance of introspective reports. In the process, much of the nuance in the views of the introspective psychologists was lost — the introspective skepticism and introspective training — and today, reports on consciousness go largely unquestioned. In the cases where someone criticizes a description of experience, there is rarely a principled foundation from which to levy a strong critique. Consider again the visual experience of a plate viewed at an obtuse angle through an exchange that begins with Alva Noë's description of this experience in *Action in Perception* (2004). Noë holds what he calls a dual-aspect view of experience, in which our visual experience of objects has a “size and shape” as well as a “perspectival size and shape.” When talking about the perspectival aspect of an experiences, he often highlights the subjective perspective of the viewer, saying things like “we experience that the plate is round and that it looks elliptical from here” (p. 172).

Sean Kelly (2008), in his review of *Action in Perception*, simply claims that this description of the experience is false. In Kelly's own words, “now, to be clear I

should emphasize that I agree with Noë that it is possible to experience the circularity of the plate, and I agree also that it is possible to experience its apparent ellipticalness. What I disagree with him about, however, is that we always experience both of these at once” (p. 685). For Noë and Kelly, the same basic building blocks of the experience are present, but they are organized into entirely different, mutually exclusive descriptions. To complicate matters further, Charles Siewert (2006) adds his own unique take on the debate, saying “true, the plate in some way and in some sense *appears differently* as its surface is positioned at various angles relative to the viewer... However, the plate does not appear to *change* shape; it looks *constant* in shape — the *same* shape during the viewing as a whole. And *nothing* (not even a so-called “sense-datum” distinct from the plate) appears to change shape during this time“ (p. 3).

These three philosophers have three different theories of consciousness, theories that are supposed to explain the various things we experience throughout our conscious lives. Descriptive phenomenology is not the only motivation behind these theories, they are also addressing philosophical puzzles, engaging with metaphysical debates, and responding to the latest scientific results. Still, we can think of our experiences as crucial pieces of data that these theories are supposed to make sense of by categorizing central concepts and illuminating connections, in short, by explaining them. The problem is that, as we see with the example of the plate, these philosophers disagree on the experiential data itself. Noë thinks the plate looks simultaneously round and elliptical from here, Kelley thinks it looks

both ways, just not at the same time, and Siewert denies that the plate looks elliptical at all. The raw experience — or at least, the introspective report that motivates the various descriptions — is supposed to be a material fact that can be settled. Ideally, once the necessary distinctions have been made, the correct description should be apparent to us all. Yet this is not happening. The same situation leads to different introspective reports; the raw data for the theories are themselves different. It is as if competing scientific labs were running the same experiment and getting three different data sets as a result. Note that, in this exchange, although the isolated introspective reports are open for criticism, both introspection and the armchair methodology itself lie outside the scope of this critique.

Next, consider conflicting claims about the richness of experience, that is, the extent to which we are conscious of detailed information from our sensory modalities (e.g. Schwitzgebel, 2011, Ch. 6). James (1890/1981) and Searle (1992) believe that we are constantly conscious of the object of attention as well as multiple periphery sensory modalities, while Jaynes (1976/2000), Dennett (1991), and Mack and Rock (1998) argue that we are rarely conscious of more than a few sensory modalities, tasks, or thoughts at a time. These views are drastically different, and despite the conflict in introspective reports, the truth about consciousness is supposed to be discernible through introspection. Of course, everyone admits that people's introspective reports can be mistaken, but whenever

introspection is criticized, it is always someone else who is doing the poor introspecting. Introspection, as a whole, lies outside the scope of criticism.

In addition to the methodological introspection of the introspective psychologists, the consciousness denying of the behaviorists, and the current introspective practices of modern philosophy, there is a small group that argues for an introspective skepticism. Introspective skeptics believe that consciousness is real enough, but its precise nature, or even its general nature, remains relatively unknown. Schwitzgebel (2011) has perhaps the most well developed skeptical position. He does not argue that we are wrong about much of our self-knowledge, rather that we do not know if we are right. I will return to introspective skepticism in Chapter 2, but for the moment I wish to only address a practical concern. You might wonder what use it is to push and prod on introspective reports when they are our sole epistemic access to experience. They are the only show in town, so what do we have to gain from questioning the validity of introspective reports aside from a debilitating pessimism? It was exactly this sort of skepticism that led to the behaviorist overthrow of introspective psychology. It was skepticism like this that led Hume to stop doing philosophy and pick up a game of backgammon instead. Personally, I like studying consciousness, and I want to know more about my own experiences. I would like to avoid the complete elimination of consciousness from my ontology. Much of this dissertation will be a methodological treatise, built off of a cognitive science foundation and accounting for known sources of introspective error. Reasons for being skeptical of naive

introspective may not (and I will argue that they do not) apply to all instances of introspection.

One final note. Although it might seem as if these general approaches break down along institutional lines, with philosophers on one side and psychologists on the other, both have a fairly mixed track record with regard to introspection. For each philosopher who makes an introspective report from the solitude of her armchair, there is a psychologist making a report from the solitude of her lab. The apparent obviousness and directness of our own conscious states makes it seductive to pronounce truths and not spend the time to empirically investigate. Because of this, psychologists are often as closed to other minds as the philosophers. Practitioners in the two fields are siblings, and there is something to the idea that armchair introspection is a tiny piece of experimental work with a subject pool of one. Speaking in generalities, both are guilty of failing to view introspective reports, and introspection in general, through a critical lens. This credulity extends into the present, and even people doing contemporary experimental work on consciousness take many reports at face value (e.g. Hurlburt & Heavey, 2006). The answer to introspective skepticism is not going to be as simple as switching departments.

IV. WHAT'S TO COME

Are we reaching into our stream of consciousness and coming out with a seductive myth or are we all getting our hands around the contents of different

streams? Perhaps a fistful of soft river mud is the relevant metaphor. The project of this dissertation is to devise and advertise a new method of studying consciousness. I am optimistic, and of the various orientations toward consciousness, I have the greatest affinity with the methodological introspection of the introspective psychologists. There are multiple ways of approaching inner life, and the best method depends on the exact question you are interested in, the specific details of the context, even the idiosyncratic peculiarities of the individual having the experience. Given that many of the answers can not be known a priori, we must search for the conditions under which our introspective reports are reliable. Many of the issues will be decided on the battlefield of experimentation, many others will require old fashioned conceptual analysis from the armchair.

What is my goal here? I want to learn more about my own experiences of reading, and I hope that in the process you will learn a little more about your own. I believe that applying an experimental methodology to the study of basic aspects of experience will have a profound effect across a large swath of philosophy and psychology. This project also acts as an egalitarian critique of philosophy, cracking open the small male-dominated circle of elites to the experiences of the masses. (A casual look at the references in this dissertation or a survey of conference attendance will show that experimental philosophy and aesthetics are two of the most male dominated regions within philosophy). With any luck, this dissertation will lay out a schema for the experimental study of consciousness that can be used

by others to advance the poor understanding of our own, individualistic and subjective, conscious experiences.

The outline. Chapter 2 lays the foundation for experimentally studying the phenomenology of reading. I begin with conceptual clarifications, arguing that the experience of reading is an action-experience that extends through time, what I call a phenomenal worm. I explore the ways in which this changes our orientation towards consciousness, contrasting it with theories that take simple cases of visual perception as their starting point. I detail the sources of introspective error and describe a research strategy for reducing them, called the Subjective and Objective Measures of Experience (SOME) method. Chapter 3 explores the cognitive processes involved in reading. I survey research on the phonological loop, experimental evidence for a visuospatial sketchpad, situational theories of narrative comprehension, the literature on our visual system, the experience of flow, and recent work on mind wandering. Chapter 4 describes and reports the results of five experiments using SOME method that were conducted 2009–2015. In these experiments, subjects read short passages and reported on their experiences before, during, and after. I then looked for behavioral corroboration of these reports by comparing them to objective measures of experience. In Chapter 5, I discuss the implications of these results for philosophy, psychology, and the understanding of our own experiences. These experiments suggest that there is a large amount of variability in the experiences while reading, and I argue that the inability of experts to see this range of experience is due to deep and pervasive

introspective bias. Further, the results suggest that coarse-grained aspects of experience, such as modal experiences while reading, are functionally isolated. They are what I call *paraphenomena*. When our experiences are not functionally isolated, the relationship between experience and behavior is often one of interference instead of facilitation. I finish by laying out the implications for combining phenomenal and functional notions of consciousness.

Chapter 2

CONCEPTS + METHODS

In many ways, the experience of reading is unlike many of the classic examples of consciousness in the philosophical and psychological literature. It floats unmoored from the perceptual stimuli that give rise to it, a lively phenomenal galley drifting and bobbing along the stream of experience. This chapter is devoted to an exploration of this vessel and its course. A common theme is variability. There is a wide range of introspective reports people provide while reading, and understanding variability is key to properly studying this experience. Poor introspection has the potential to create the mere appearance of conflict, as when there are two competing descriptions of the same experience. It also has the power to paper over different experiences with the same introspective report. At the same time, there could be individual differences between people, or for one person over time. This is all to say, without a proper analysis of the experience of reading, we could be seeing variability when there is none and sameness when there are differences.

Philosophy is a discipline full of thought experiments, yet there is surprisingly little experimentation on our own experiences. When consciousness is the target of a thought experiment, we are asked to consider what a hypothetical person would experience in a given counter-factual scenario, be they our “twin” or

someone with an impoverished conceptual framework. We are rarely asked to consider other experiences that we could have. I suspect this stems from the feeling that our own consciousness is obvious and universal, as if the precise experiences we have are logical consequences of psychological laws that govern human phenomenology. For a moment, humor me with one. Imagine a phenomenal world that is an exact duplicate of the one you currently inhabit, save for one difference — you can not read. (This phenomenal state is not rare, and to put it in perspective, according to UNESCO there are one billion illiterate adults in the world). Now, from this hypothetical perspective, what would you expect the experience of reading to be like? You have seen black marks on a piece of paper, you have heard numerous voices throughout your life (maybe even an inner one), and you have had visual imagery of many sorts. You are not conceptually impoverished (e.g. Mary in Jackson, 1983; 1986), but rather have knowledge and experience of all the proper phenomenal pieces. So, given that Illiterate-You has not experienced reading for themselves, what are their thoughts on the phenomenology of reading?

My intuition, and maybe you share it, is that Illiterate-You would expect the experience of reading to contain nothing beyond the visual experience of the words on the page. After all, your eyes must scan the page while reading, an essential action that differentiates reading from imagining, writing, mental rehearsing, listening, remembering, and hallucinating. To put yourself in a

situation that is somewhat analogous, imagine reading a sentence in a language you do not understand. Perhaps a bit of classic Chinese philosophy:

大人者，不失其赤子之心者也。

Did you have an auditory experience? How about visual imagery? Or did you only see black marks on the page, meaningless to those who have not learned the conventional connections that bind them to the world? For reading, a causal relationship between your eyes and the words is necessary, yet oddly, the experience of the text itself is not commonly cited nor universally accepted. As we saw in Chapter 1, Julian Jaynes rejects the idea that we have this experience. Similarly, descriptions of getting lost while reading, of “taking off” into a narrative, often lack a perceptual dimension. Contrary to the naive thoughts of Illiterate-You — and perhaps your experience of reading Chinese — the academic literature on consciousness privileges mental imagery over visual perception. By the way, this Chinese sentence was written by Mencius, and translates as “the great man is he who does not lose his child’s-heart” (Mencius, 4th c. BCE/1895, 4B12).

As the thought experiment brings out, there is a strong sense in which reading should not give rise to visual imagery or an inner voice. Apart from the fact that people report these experiences, what reasons do we have for thinking that reading about a person would lead to a visual experience of them, let alone whatever visual experiences we might have while reading abstract philosophy?

Why would silently reading a word lead to an auditory experience of the word being read — not out loud — but in loud? And what about reading crossword puzzles, text messages, or tattoos, non-narrative forms of text that seem even less associated with a visual or auditory experience.

The first section is devoted to conceptual clarification. I argue that the experience of reading is subject to a high degree of variability because it is *an action-experience* and a *phenomenal worm*, then I sketch out some of the ways that our experiences change: over time, across individuals, across different texts, and depending on our goals and purposes. The second section lays out the sources of introspective error: inference error, over-generalization, and short-term memory limitations. In the final section I argue that we must compare subjective and objective measures of experience using the (aptly named) Subjective and Objective Measures of Experience (SOME) method. SOME method minimizes the sources of error by focusing on reports of short, retrospective, specific moments of inner life and corroborating them through behavioral measures.

I. ACTION EXPERIENCE AND PHENOMENAL WORMS

You can look at a river and see a torrent of water. Someone else can look at the river and see a trickle. The two experiences sound different, maybe even contradictory, but both can be experiences of the same river. Perhaps one of you visited the river in the winter, the other in the summer. Or you are upstream near the source, before the river combines with its tributaries, while the other is at the

river's mouth, after this augmentation. Later, one of you might make grandiose claims about what you saw, saying "the experience of the river is one of cascading water every moment of the waking day." The other could retort "a very diligent examination of the waters, and getting others to consider them, I do not find that one in twenty times any such flow is formed." Either of you could ask "why should I insist on those things which everyone's experience will, I doubt not, plentifully suggest unto him?" Such declarations are ludicrous in this context, yet they are the watery equivalents of the claims people have made about the experience of reading from Chapter 1. Some of the strangeness stems from an unexamined, implicit idea of what the experience of reading amount to. Once we have more clarity on the concepts involved, we will see that these types of comments are just as senseless when applied to a stream of water as they are when applied to our stream of experience.

Heraclitus famously said that you cannot step into the same river twice. Change was a major aspect of Heraclitus' philosophical view, and he thought, paradoxically, that constant flux is a necessary condition for constancy. It is only because a river has water flowing through it that it continues to exist at all, and should the flow of water come to a halt, it would become something else entirely, perhaps a lake or a dry creek bed. Similarly, we continue to exist as unique individuals because the cells that make up our bodies are constantly cycling in and out every seven years or so. To put an introspective spin on Heraclitus, the constantly changing stream of experience could lead to a constantly morphing

stream of reports. There are differences in the way people describe the experience of reading. Sometimes there is a substantive issue in dispute, but I suspect that most debates are the result of an inherent variability in consciousness itself. Now, a metaphor is not an argument, even if the metaphor has ancient roots. But this general framework is useful in making sense of the conflicting introspective reports from Chapter 1. I will give some a priori reasons for thinking that the experience of reading is a stream in constant flux, then I will give concrete examples of the protean nature of experience. Accounts that differ are (partially) accounts of different things, or accounts of a changeable, mutable aspect of the same basic experience.

Much of the literature on consciousness focuses on perceptual experiences, so consider an everyday example, such as the visual experience of redness from looking at a stop sign. Here is an intuitive statement: the stop sign is red because it will look red to a normal perceiver in normal lighting conditions. Such a commonsense view is advocated by Peacocke (1983), among others. I want to highlight, not so much a problem with any specific analysis of this statement, but features of this experience that make it a poor starting point for understanding the experience of reading. I will compare the perceptual experience of a stop sign with the experience of reading, and argue that emphasizing perception obscures the wide range of variability in our phenomenal lives. I begin by describing four features of the common experience of seeing a red stop sign.

First, the experience of redness is what I call a *qualia-experience*, a category of experiences that are defined by the particular phenomenal characteristics. The claim that a stop sign is red because it looks red defines what it is to be red in terms of a specific experience, that is, the experience of redness. For a qualia-experience of color, it could not have another phenomenal characteristic of color without being a different experience entirely. There are ways of interpreting a phrase like “what if the experience of redness looked black?” but it is a borderline nonsense question. Of course, something that looked red once can now look black, but this is a claim about an object, not the experience of it. Redness can not look black any more than red can be black. Qualia-experiences are constrained by the logical implications of their defining phenomenal characteristics, and in the case of the experience of color, this implies mutual exclusivity with other colors. This makes qualia-experiences, such as the experience of redness, context-independent. Insofar as they are instantiations of their core experiences, you always know what you are getting when you experience redness. The shade of red might change or a pain might shift from a dull ache to a throb, but the phenomenal characteristics that make them experiences of redness or experiences of pain remain the same. In other words, you can look at a stop sign and see the same color, over and over again. Redness is not going to change on you.

Second, with the visual experience of a stop sign, objective facts about the world provide firm guidance for the experience we will have. If I know that you are looking at the sign, that your visual system is functioning correctly, that the

sign is made out of a material that reflects lightwaves of a certain length, and that there is light to be reflected, I know what your experience of it will be. Actually, I do not even need to know this much. For everyday talk, it is enough for me to know that you are looking at an object that is red. Circular as this folk statement may be, I intend it only as a contrast with other modal experiences. We do not expect, for example, an auditory experience of the red sign. In fact, we would say that an auditory experience is simply not an instance of visual perception. The modality is all wrong, and when it comes to the perception of a stop sign, we know what the modality of the experience is going to be.

Third, this commonsense claim requires ideal conditions to ground the connection between properties of the object (being red) and properties of the experience (being of redness). This expectation comes from the robust causal connections that exist between our experiences and the world we perceive. We frequently think of these connections as having a normative dimension. If I see a red stop sign, I *should* have a visual experience of redness. Of course, a stop sign can look many different colors depending on lighting and a whole host of philosophical thought experiments, but there is an intuitive appeal to thinking of redness as the normative color to be experienced. In fact, we often say basic, uncontroversial things like “stop signs look red,” even though they do not literally always look this color. However, they appear red often enough in *normal* circumstances that we say this without objection from anyone but a diehard skeptic. In most contexts, red is the *expected* or *ideal* color of a stop sign, which is

all I mean in saying that there is a normative dimension to the visual experience. Deviations from experiencing a stop sign as red are added onto the theory, something to be accounted for after the fact, like air resistance in Newtonian physics.

Finally, this description of the perceptual experience of a stop sign does not have a temporal aspect. I call this sort of momentary snapshot of experience, exemplified by the simple case of color perception, a *phenomenal slice*. As with qualia-experiences, there is a fixedness to phenomenal slices, they have some sort of character, and this character is unchanging. Of course, your experience or the stimulus could change, but then you would have a new phenomenal slice. I do not think a phenomenal slice necessarily captures a moment of objective time, if such a thing exists, but rather tracks phenomenal simultaneity, everything experienced as happening at once. Although phenomenal slices are not the only way to think about the experience of colors, for many philosophical questions, it is perfectly fine to focus on them exclusively. What is the experience of pain? Were you angry a moment ago? Do you feel the shirt on your back now? These questions can often be answered with reference to a specific moment of inner life.

Contrast the perceptual experience of the stop sign with the experience of reading. As we have discussed already, it is an experience that can take numerous forms: an inner voice, visual imagery, the visual perception of the words on the page, as well as others less commonly reported. Unlike the experience of redness, the experience of reading is not defined a priori by any given phenomenal

characteristics. Rather, it is determined by an action, the act of reading, and is what I call an *action-experience*. It sounds trivial to say that, for something to be an experience of reading, it has to be experienced while reading, but defining things in this way is highly substantive. One consequence is that this builds in a measure of variability over time that does not exist in something like the experience of redness. Of course, the precise shade of red can change, but all the visual experiences of these shades are still experiences of redness. The experience can not stray too far, for instance, auditory experiences are excluded. Reading is different, and the possibilities can take drastically different cross-modal forms. One way to see this is to return to the question that began this thesis, what are you experiencing right now, while reading? Is your experience of reading these sentences the same as it was while reading the opening pages? If you want, try a little experiment, a measure of test-retest reliability. Return to the first page and read it again. Has the form of your experience transformed over the course of two chapters?

This basic point, that the experience of reading is an action-experience and can change over time, can be extended beyond rereading the same sentence. The experience of reading could be vastly different now than it was in ancient Greece. After all, the behaviors associated with reading have changed dramatically. It is well documented that in ancient Greece, books were not read silently, but were rather read out loud, by yourself or in front of a group (Saenger, 1997). Silently reading was uncommon enough that in his *Confessions*, Augustine of Hippo

commented on the amazing ability of Ambrose of Milan to read silently to himself (4th c. CE/1991, Book VI, Section III). If the actions associated with reading have changed this much, our experience of it could have undergone an equally dramatic transformation.

Second, the act of reading does not provide detailed guidance towards the type of experiences you will have. Action-experiences are variable, and because of this, unpredictable. Why would looking at black marks on a piece of paper lead to an auditory or visual experience? There are black marks on the white wall of my office, yet these merely lead to a run-of-the-mill visual experience. There is something special about words — they have meaning, they refer to things, attribute properties to these things, have connotations, cause emotions, and bring up memories. Even still, none of this necessarily entails auditory or visual experiences. Given the number of people who have described their experience while reading as having cross-modal components, it would be remiss to discount these experiences merely because they are not implied by the external stimulus. Still, from a third-person perspective, a perceptual experience is all we might expect. In this discussion, I have been trying my best to avoid controversy, sticking to claims that should be widely acceptable to most readers. However, once we consider the numerous ways in which our experiences can deviate from the properties of the world, a more radical position presents itself, and it could be the case that people have radically different experiences when they look at a stop sign. The inverted spectrum hypothesis is one such example (e.g. Shoemaker, 1982; Block, 1990).

Third, there is no normative or ideal experience of reading that can make it analogous to the visual experience of the stop sign. People are not in a priori agreement as to what the core experience should be. Further, there is no agreement on what constitutes ideal conditions. We can read fictional narratives, poems, historical non-fiction, and argumentative essays. We can read them by skimming, reading silently, or reading out loud. We can read in a loud room, while watching TV, during a seminar, or while drunk. We can proofread, we can read for content, we can read to get lost in a story. These types of reading, ways of reading, situations in which we read, and reasons for reading could lead to drastically different experiences, but none of them stand out as being on top of a phenomenal hierarchy. Consider again the experience of mind wandering. If we are spacing out a sizable amount of the time, does that make spacing out a part of the experience of reading? There is no principled line to draw. After all, visual imagery and inner speech are no more a priori exclusive to reading as spacing out. These are all experiences we have through out the day. Some of the time we have them while reading, and since the experience of reading is defined by an action, is an action-experience, they all have an equal claim to legitimacy. Besides, it is not as if ideal conditions are easy to spell out in the case of visual perception either. Once you get into the details of spelling out ideal conditions, it might turn out that the experience of the stop sign is on the same epistemic footing as the experience of reading.

Because action-experiences are defined by an action, they extend through time in a way many experiences do not. Thus reading (and the experience of reading) necessarily has a duration. It is an action through and through, and actions take time to complete. Because of this, the experience of reading is a *phenomenal worm*, a collection of conscious states that extend through time. As the name implies, phenomenal worms can wiggle and change in a way that phenomenal slices can not. It is the difference between a photo and a video; between a note and a song. While an *action-experience* concerns the defining nature of a particular experience, a *phenomenal worm* concerns the timespan we must take into account. In contrast, the perception (and perceptual experience) of color can occur in a moment, the “specious present” (James, 1890). In order to have an experience of redness, you do not need to interact with a colored object beyond a momentary glance. Nothing further must be done, and in fact, even this cursory look is more than you need. Most people are able to conjure up a mental image of redness without so much as a peep at an actual, real-life red thing. We have this power on command, the power to throw an experience into our stream of experience and hold it there, neither sinking nor floating downstream. Even when an action is necessary for the experience to occur in the first place — for a visual experience to be perceptual you must look at something — this action does not define the experience. In fact, the initial gaze shift, the constant micro eye movements, and the final look away are not thought of as part of the experience of color. What it’s like to see red does not involve a gaze fixation point that saccades

to a patch of redness and then quickly saccades away. These are essential actions for the perceptual experience, but in a contingent way, a byproduct of our particular physiology.

There is not a hierarchy between phenomenal worms and phenomenal slices. Neither is preferable in general, it all depends on the questions you are asking. If you are interested in the experience of redness, you often do not have to worry about the experience changing. A momentary snapshot tells you everything you need to know, and we can then talk about many traditional issues in philosophy: phenomenal qualities, hallucination vs. perception, and epistemic justification. The same goes with other experiences such as pain, visual imagery, and emotional experiences. By starting with an experience that has a fixed form, we can exclude information that is unnecessary for the issue at hand. As qualia-experiences, their species is fixed by definition. With a phenomenal slice of a qualia-experience, there is a measure of stability that is lacking in the experience of reading. Looking at a stop sign, it certainly seems as if we can dip our feet in the same stream of visual experience twice. It might look different in some respects on a second viewing, but there is a strong intuitive pull to the idea this leads to a representation, not a novel experience. However, when our interest is in an action-experience, the action is stipulated while the experience is not, leaving conceptual space for the experience to morph. Just as the way we read has changed since the time of the ancient Greeks, the experiences associated with reading can change,

over the course of two millennia and over the course of two pages. Looking at a phenomenal worm simply allows us to see the changes that are there already.

I have been assuming that we all share an intuitive understanding of what the experience of reading is, but we should unpack this implicit set of assumptions. Grammatically, “the experience of reading” looks similar to some of the most commonly discussed experience in philosophy, such as “the experience of redness.” As I have been arguing, it is of a different class entirely, subject to a degree of variability that makes it conceptually distinct and empirically unique. As a whole, consciousness contains a vast realm of experiences, but this is not a domain that has been evenly explored. Of the assortment of experience we have on a daily basis, perceptual and quasi-perceptual experiences (such as hallucinations) receive the bulk of the rigorous academic attention. In both philosophy and cognitive science, the vast research on vision stands in stark contrast with the dearth of attention given to non-perceptual faculties such as dreaming, the imagination, or creativity. Given the common descriptions of the experience of reading, it belongs next to these marginalized categories, and shares more in common with imagination than perception. Of course, as Illiterate-You suspects (and as you may have experienced while reading Chinese), there is a perceptual dimension to reading, but focusing on this aspect alone can lead to bizarre conclusions. Again, consider the idiosyncratic theory of Jaynes. Jaynes argues that consciousness was fundamentally different among humans in the distant past. These ancient people had a “bicameral mind,” a schizophrenic and disjointed consciousness in which

one side of the brain experienced an inner voice while the other mistook it for an auditory hallucinations of — to give his most striking example — the voice of the gods. As Jaynes says “reading in the third millennium B.C. may therefore have been a matter of hearing the cuneiform, that is, hallucinating the speech from looking at its picture-symbols, rather than visual reading of syllables in our sense” (Jaynes, 1976/2000, p. 182). This is one way to make sense of an inner voice: treat it as a literal hallucination.

No one today thinks that we are hallucinating inner speech when we read, but why not? After all, there is not a voice out there, in the world, speaking to us. Any accompanying auditory imagery exists entirely inside our heads and is not the result of auditory perception. I do not want to make an attractive view out of the idea that an inner voice is a hallucination. After all, an inner voice is not a misrepresentation of a veridical process like perception. But with hallucination out of the picture, what are we to make of introspective reports that are not grounded in an external stimulus?

There is a tradition in the philosophy of language that focuses on language use over a “sense” or “meaning” that exists independently (Austin, 1962; Korta & Perry, 2011; Wittgenstein, 1953/2009). On these views, in order to understand language, we have to treat it as a species of action, not as an isolated intellectual exercise divorced from larger social and normative constructs. We do not yell “time for dinner” to make a theoretical point about temporality or the necessity of caloric consumption, we say it because we want people to do something, to come

to the table and eat. This idea could be expanded from its roots as an analysis of meaning or information to an explanation of the experience of language. There is an intuitive pull to the idea that J. K. Rowling did not write narratives about Harry Potter because she wanted us to know abstract facts about a fictional character. She did it because she wanted us to experience something of his made-up life. If language is imbedded in a broader context, something like Wittgenstein's language game, then the experience of language might also occur in this broader context. An important aspect of a language game is the ability to play it, to make it a real and an embodied part of life. On this view, it would be fitting for the experience of reading to give rise to the experience of actions, such as the reader silently talking to herself or the narrative scene unfolding, a play with phenomenal actors. Following this thread, we write because we want others to experience the ideas conveyed, and verbal language is our earliest, most hard-wired communicative act. Similarly, we read because we want to share in these carefully crafted experiences.

Because it is defined by an action, the experience of reading could be highly variable, over time, across different people, and over different texts. The metaphor of a dynamic stream of experience is fitting. Like the river of Heraclitus, our experience of reading is constantly changing, and this flux is what makes it the experience that it is. To understand something abstract like "the experience of reading," we need to look at the actual things that people experience while reading. Although this framework carves out conceptual space for the experience of reading to be highly variegated, I should say more about the sources of this

variability. We have already discussed variability over time, the way the experience can change while reading. To this we must add variability due to individual differences. Some of the disjunct between conflicting reports of the experience of reading likely comes from substantive differences in the experiences themselves. From my vantage point, I find it unlikely that Bernard Baars is constantly experiencing an inner voice while reading. This does not mean he is wrong, but due to the sources of introspective error I will discuss shortly, it also does not mean he is right. There is variability among people in many cognitive and physical abilities, so it is entirely possible for different people to have different experiences while engaged in the same action.

Although individual variability in the experience of reading has been largely ignored in the philosophical literature, there are a few examples of philosophers grappling with the implications of individual variability. One example is the way drinking wine is experienced as someone becomes a skilled wine taster. It is nearly impossible for most people to parse out the subtle flavors of wine, and as the case is typically given, this is because us wine slob (I include myself in this category) have a fairly coarse-grained experience of drinking wine (e.g. Lycan 2004; Rosenthal, 2002). Contrast this with the experience of a master sommelier, who, during the blind taste exam administered by the Court of Master Sommeliers, must be able to identify and describe six bottles of wine from anywhere in the world. The conceptual and experiential framework of the master sommelier is considerably richer than for a wine slob, or so the argument goes. There is debate

about whether this is due to a difference in perception (Lycan, 1996) or thought (Rosenthal 1986, 1993, 2005), but there is agreement that things are radically different for the two people at the level of consciousness. In fact, the entire thought experiment depends on an asymmetry in the experience of the two wine drinkers. In thinking through this case, I find my own introspective judgments lacking. Try as I might, it is not obvious to me that these two wine drinkers have different experiences. Perhaps the change in experience is only manifest to those who have undergone the perceptual training and socialization that makes one into a wine snob. At the very least, we should not prejudge the matter.

A third source of variability in the experience of reading is the effect of passage type. The experiences we have while reading narratives could be drastically different from the experiences we have while reading philosophical treatises. This is one commonality between the experience of reading and visual experience in general. Looking at a red patch leads to a different experience than looking at a green patch, and these are the simple, bare-bones examples. Consider the visual experiences of watching a movie, playing soccer, watching a friend play soccer, gazing at the clouds, or watching a friend watch another friend play soccer. When we look at unique types of objects and scenarios we have sweepingly different visual experiences. Something similar may be true for reading: when we read distinct types of passages we have distinct experiences of the reading. Even this comparison, however, quickly reaches its limits. To state the obvious, perceptual

visual experiences must be visual in nature. There is no such constraint on the experience of reading, and it could be visual, auditory, or cross-modal.

A fourth source of variability comes from the purpose of reading. Reading is an action, but even the act of reading allows for a range of ways to engage with a text. You could be proofreading, reading for content, reading out loud, reading to learn about the world, reading to relax after a hard day, and reading to understand an abstract philosophical thesis. These different purposes and goals behind reading could lead to different experiences. For example, the visual experience of the words on the page may be more prevalent when proofreading, and an auditory experience may be more likely if you are reading lyrics to a song.

As we look harder at the experience of reading, it begins more and more to resemble Heraclitus's ever-changing stream. Instead of the simple mental picture of an inner voice or visual imagery, we have a lively phenomenal worm only loosely tethered to the pages of text. Our consciousness often flies into all manner of associations, auditory cues, images moving and still, memories, experiential blanks, and thoughts entirely unrelated. Given the diverse scope of experiences possible while reading, it is surprising that the definitive descriptions of reading have not included this sort of phenomenal flux. Perhaps we can not step into the same stream twice, but we can still measure the speed of the current, take note of animals that live nearby, and find new ways of enjoying its waters. Our collective blindness to variability in the experience of reading speaks volumes about the shortcomings of the way consciousness has been studied in philosophy and

psychology. In the next section, I look at the causes of this blindness, the sources of introspective error that we must confront if we wish to understand the experiences we have on a daily basis.

II. SOURCES OF INTROSPECTIVE ERROR

Methodology is important in the study of consciousness. The way we explore and investigate consciousness can impact our conclusions in many ways. For example, our approach could lead to distorted and intractably incoherent results, as we saw in the imageless thought debate, or it could restrict the acceptable domain of inquiry, as we saw during the height of behaviorism. As I discussed in Chapter 1, philosophers traditionally favor the armchair methodology. As a research method, it has remained relatively static since the middle of the 20th century. There has been a constant flow of arguments, of novel intuitions, of counter-examples to existing theories, and of intuition pumps that motivate a new view or nuance. The armchair methodology itself, however, is much the same today as it was decades ago. Despite its dominance in philosophy departments, armchair reasoning is not the only method of studying experience. This is not a critique of armchair reasoning per se, but given its methodological dominance over the last fifty years, I feel the importance of pushing against it to make room for alternative, parallel methods. The fact that armchair reflection has been popular historically does not mean that it can not be improved upon or augmented through information gleaned from other sources.

Introspection, in and of itself, is neither inherently reliable nor unreliable. It is one cognitive mechanism among many, and much of this dissertation is devoted to developing a method that improves the accuracy of introspection beyond naive generalizations. There is a creative aspect to experimental design that is easy to forget. Consider the extensive literature on change blindness and attention blindness. Given the amount of research on these two phenomena, it is shocking to remember that they were not widely studied until the 1990s. From this body of work, we know that there are systematic ways to fool our visual system into missing obvious and important features of a scene. Importantly, this was discovered through novel and ingenious experimental design, not through general knowledge of the visual system or through armchair introspection.

In what follows, I will make the case for a far-reaching moment of methodological reflection and self-criticism from within philosophy and psychology. I am not the first to argue for the need to improve our methods of studying consciousness (e.g. Hurlburt; 2006; Paccinini, 2003; Jack & Roepstorff, 2002). I do not have more to say against the armchair methodology specifically. Instead, I proceed on the assumption that I have already imbued you with a healthy skepticism toward unchecked armchair introspection.

In the final section, I lay out a detailed methodology for studying consciousness that can inform and check our armchair introspection and intuitions. First, however, I need to discuss the sources of introspective error in detail. I want to proceed carefully, making as few assumptions as possible about

the reliability or unreliability of introspection. Unfortunately, because introspection is our primary access to our experiences, knowing when our introspective reports are accurate is not easy, if not impossible. To up-cycle an example on calibrating intuitions from Cummins (1998), imagine Galileo peering through the first telescope. Although he expected to see a perfectly spherical celestial body when looking at the moon, instead he saw imperfections, mountains and craters that are not visible to the naked eye. From Galileo's point of view, there were two competing explanations. Of course, there could be actual mountains on the moon, however, given the novelty and untested nature of the first telescope, the apparent-mountains could also have been artifacts of this new optical tool. To settle the matter, Galileo could calibrate the telescope, pointing it at something with a known shape and seeing if the telescope introduced new geometric error. Since the telescope correctly made earth mountains look like mountains, Galileo concluded that the moon mountains were actual peaks on the lunar landscape. For the first telescope, calibration allowed Galileo to determine if an unexpected result was an instance of measurement error or simply an unexpected state of affairs. No such calibration is possible for introspection. We do not possess independent access to our experiences, so there is no way to know for certain if we are "seeing" actual mountains or introspective artifacts. We can look at the cognitive mechanisms behind perception, for example, but this does not calibrate introspection any more than knowledge about atmospheric refraction can

calibrate a telescope. They form the necessary conditions for some of our experiences but are nevertheless distinct processes.

We can not say with certainty when our introspection, and our introspective reports, correctly reflect our experiences. We can, however, look at the known causes of introspective error, the cognitive mechanisms, disturbances, and limitations that lead to inaccurate introspective reports. Once we understand the sources of introspective error, we can build a research method that takes them into account, reducing them when possible and using other means to infer the truth when not. The sources of introspective error that I will discuss are inference error, over-generalization, and memory limitations. I will address each in turn.

Inference Error: Our perception, memory, self-attributions, and even introspection are veridical, but aiming at truth is not the same as hitting it. We often see things that are not there, tell stories about our past that never occurred, and say things about ourselves that are simply false. Perhaps the clearest example is confabulation, a species of falsehood that is one part fabrication and one part ignorance. When simply mistaken, you think there are four socks instead of five; when confabulating, you devise elaborate explanations for why one of the five identical socks is preferable. Speaking of socks, the reasoning behind a subject's choice of socks in Nisbett and Wilson (1977) is one instance of inference error that I discussed in Chapter 1. While it is reasonable to assume that there are objective differences between a series of socks that someone asks you to rate, confabulation

occurs when this assumption morphs into a detailed explanation for non-existent differences.

We can even confabulate conscious intention, something as foundational to our sense of self-identity as the choices we make. Wegner (2002) gives a detailed overview of the literature on our poor knowledge of our own choices. An illustrative example comes from the I Spy Study (Wegner & Wheatley, 1999). A subject and a confederate sat down in front of a computer screen showing a jumbled picture of objects from the book *I Spy*. Both the subject and the confederate were wearing headphones, ostensibly as a distraction, and both put their hands on a mouse which could be used to jointly control the position of the cursor on the screen. The subject and the confederate were instructed to move the mouse randomly around the screen, stopping every thirty seconds. Unknown to the subject, sometimes the confederate would allow the subject to choose where the cursor stops, but other times the confederate would force the cursor to stop on a particular object. After each stop, the subjects reported whether they intended the stop to happen or merely allowed the stop to happen. Subjects tended to report the forced stops as slightly intended, but this report could be manipulated by priming the subject with an auditory cue. For example, some subjects heard “swan” through the headphones, and the closer the subject was primed to the forced stop, the more they reported intending to stop where they did (up until two seconds before, at which point the feeling of intention decreased). As experiments like this

show, even something as foundational as the sense of having chosen something can be a complete fabrication, an inference of causal control when none exists.

We are often ignorant of the extent of inference error. If you are like me, you have faith that we are not simply fabricating large chunks of daily experience, yet we know that our perceptual experiences, experiences of reasoning, and the experience of conscious will are prone to confabulation. Confabulation is difficult to control for, since we can not see it from our subjective point of view by definition. The studies cited above are specifically designed to root out confabulation, and they do this by having independent access to the facts of the situation. The researcher already knows that the socks are all the same or the true causal mechanism behind the final resting place of the cursor. The researchers know these things because they engineered the situation, custom-made to smoke out instances of confabulation. This is not possible with the experience of reading. As we found in the last section, the experience of reading is an action-experience, so we can not tinker with the situation to make one experience correct while another is not. Any experience we have while reading has an equal claim to legitimacy. Because of this, we can not know that someone is confabulating based on their introspective reports alone. However, we can corroborate introspective reports by looking for signs that they are tracking something of value in the mind. Thus, we can “triangulate” on experience by using introspective, behavioral, and physiological evidence to approach consciousness from different levels (Jack & Roepstorff, 2002). Previewing Chapter 3, reports of an inner voice can be

corroborated through psychological evidence for an auditory dimension to text processing, perhaps the phonological loop. This does not give us the certainty found in the confabulation experiments above, but in this way we can, if not calibrate introspective reports, at least find independent evidence for their truth.

Overgeneralization: Overgeneralization occurs when we take a subset of a domain and apply it universally to the category as a whole. Examples from everyday life abound, such as “my partner never does the dishes” or “I always do the cleaning.” The strong categorical statements from Chapter 1 have the hallmarks of overgeneralization. They lack nuance or an eye to variability, as if one experience persists unbroken through time. I do not know if Baars really talks to himself every moment of every day, but I do not take his claim at face value. This is exactly the sort of thing we should expect someone to say about experience given what we know about overgeneralization.

Russ Hurlburt and Christopher Heavey have found that introspective reports that generalize over experience often bear little resemblance to introspective reports on specific moments in time. For example, Hurlburt and Heavey (2006) document a subject, Donald, who was suffering from crippling anxiety. After participating in a study in which he described many of his momentary experiences over the course of a few days, the researchers noticed that Donald often described feeling angry at his kids. Surprisingly, Donald denied that he was often angry at his kids, despite the fact that he often reported as much. It was not until his reports

were read back to him that Donald acknowledged that this was a recurring and important aspect of his daily life.

One species of overgeneralization is known as the refrigerator light problem (Thomas, 1999; Block, 2007). This is a phenomenal equivalent to the Heisenberg Uncertainty Principle. The act of opening a refrigerator turns on an internal light, so if you want to know if the inside is always illuminated, cracking the door will not give you the correct answer. Similarly, the act of introspecting can alter the experiences that we have. The clearest examples come from cases of directed attention. Introspect, now, whether or not you are experiencing your left foot in your shoe (Schwitzgebel, 2011, ch. 6). Of course, once you are thinking about your foot, the experience is apparent. But I asked about your experience before the question was raised, and you can not assume that your experience remains unchanged once you direct attention to it. This example is as humdrum in its commonality as it is expansive in its application. The lesson, and a serious source of introspective error, is that we can not know our current experience without potentially altering the experience in the process.

To correct for overgeneralization, we can treat introspective reports as something akin to a scientific hypothesis. Our general reports of experience can never be proven correct, but they can be disproven through careful attention to reports on a specific experience over a short timespan. Consciousness as a whole is comprised of the experience of individual moments, a grand edifice of specious present stacked on top of specious present. Given the relationship between general

and specific reports, they should both be telling the same basic story. If Baars truly speaks to himself every moment of the day, when queried about specific experiences throughout the day, he should report inner speech in all of them. By looking at both types of reports, generalized reports and reports of short, specific momentary experiences, we can get a picture of our inner lives that balances the desire for a broad understanding with the effects of overgeneralization.

All metaphors have their limits, and there is an important disanalogy between a refrigerator and consciousness. When the refrigerator is closed, it is a black box (or rather, a white box) from our perspective. There is literally no access to the inside without opening the door and thereby turning on the light. But with consciousness, we are already inside our heads. “Opening up” is not so much a matter of removing an obstacle as paying attention to what is already present. Because of this, our memories of experience provide another route to self-knowledge, one that sidesteps the effect of attention on experience itself. Although introspection alters our current experiences, this does not necessarily apply to past experiences. To avoid the refrigerator light problem, we can give retrospective reports, passively experiencing the world and then attending to the experiences that just occurred. Of course, relying on memory comes with its own set of limitations, to which we now turn.

Memory Limitations: While the refrigerator light problem concerns the effect of introspection on our current experience, memory limitations call into question retrospective reports of experience as well. Any report of past experience

is susceptible to the shortcomings of memory. I am sure you were awake for most of yesterday, but I challenge you to make a detailed recounting of the full day's worth of experiences. It is entirely possible for experiences from the previous minute or the last few seconds to evaporate from memory the way dreams often remain stubbornly out of sight. The fact that an experience is not remembered, even seconds later, does not necessarily mean the experience did not occur.

The last fifty years of cognitive research on short-term memory shows that there is significant forgetting even after a few seconds (Baddeley, 2003; Cowan, 2000; Ericsson & Kintsch, 1995; Peterson & Peterson, 1959). Short-term memory performance is not simply a matter of time since the original presentation, it also depends on the nature of the stimulus. Here is a quick survey of some relevant experimental work. Text recollection decreases considerably when the words are phonological similar, that is, when they all sound the same (Baddeley 1966). Short-term memory decreases as the number of syllables in the words increase (Logie, Della Sala, Laiacona, Chalmers, & Wynn, 1996). Certain intervening tasks can create interference, for example, hearing irrelevant sounds or words decreases performance in word recollection tasks (Colle, 1980). Recent work on phenomenal overflow suggests that our experiences have a rich content that “overflows” or exceeds the capacity of our perceptual system (Block, 2011). This is far from comprehensive, but merely acts as a gesture toward the sizable experimental work on our short-term memory of text. I will go into more detail on memory in Chapter 3, but the thrust is that we can not assume accurate short-term

recollection of consciousness. In fact, our experiences while reading appear to be exactly the sort of thing that we would quickly forget. They are not mere single-syllable words that are rehearsed into memory and then repeated in the silence of a lab. Consciousness is complex, cross-modal, and we are continuously swimming against the current of a never ending stream of new experiences.

It makes theoretical sense for there to be hard limits on memory storage. Thinking about the computational resources required, there is little cognitive payoff for retaining a large portion of the information contained in our experiences. Perceptual experiences particularly stick out as being superfluous. Since much of the information in consciousness is extracted from the world, why remember it when you can always return to the source, the world itself? Besides, it is computationally impossible to store a continuous stream of information, suggesting that there is a great deal of forgetting of both perpetual and non-perceptual experiences. Since we are constantly having experiences, how could our brains keep up? Memory limitations can work in synergy with overgeneralization to create skewed overviews of our conscious lives. If we forget much of our experiences, then the experiences we remember could be augmented to fill the void in our inner narrative.

From the discussion of the refrigerator light problem, we are interested in retrospective reports of experience, a focus that is complicated by the limits of short-term memory. However, we can reduce memory error in a number of ways. First, because of the negative impact of intervening tasks on memory, we should

focus on reports of an immediately retrospective experience. We should be aware of the effect of similarity on memory, and account for it appropriately. Finally, we should keep retrospective reports as simple as possible to minimize the effect of complexity on short-term memory. This means focusing on uni-modal, forced choice reports. The kind of introspective report that does not require an understanding of the finer points of consciousness, but can rather be made in an instant, such as “Did you experience an inner voice a moment ago while you were reading?”

Minimizing the sources of introspective error is not nearly as grand or ambitious as creating ideal conditions for introspection. This second goal, however, is impossible given the calibration problem inherent in introspection, especially for an action-experience. All we can do is remove the known obstacles and hope that, over time, we can hone and refine our methodology, further reducing these sources of error. Incremental improvement is the best we can do. Of course, there is always the worry that a systematic understanding of experience is impossible. Perhaps introspection is simply unreliable, regardless of our best efforts and ingenious experimental design. This is one explanation for the conflicting descriptions of experience in Chapter 1, and given the lack of an independent means of calibrating introspection, it will forever remain a possibility. The success of any experimental strategy for studying consciousness is not assured, but if we think consciousness is important, we are compelled to continually try.

III. METHODS OF STUDYING CONSCIOUSNESS

We need to minimize the sources of introspective error in our study of consciousness. Allow me to quickly chart a course through the methodological concerns. First, we can ward against unrepresentative overgeneralization by focusing on reports of short, specific moments of time. As I have stressed, it is one thing to be wrong in a broad claim about consciousness, quite another to be wrong in a claim about experience of a particular moment. By doing this, we can compare reports of individual phenomenal slices with each other as well as with general claims about experience. Second, we can ward against inference error by corroborating introspective reports with objective measures of behavior. For example, an accurate recollection of visual detail does not prove a report of visual imagery correct, but it is an endorsement of the experience, independent evidence that a reader is processing visual information. Third, we can minimize the effect of the refrigerator light problem by asking subjects to report retrospectively on their experiences. However, we must be careful with retrospective reports because, fourth, we come up against the limits on short-term memory. Thus retrospective reports of experience should concern the immediate past, just a few short seconds prior to the moment of introspection.

Taking all this together, we should modify the initial question that began this thesis. We should not only ask “what’s the experience of reading,” but also “what were you experiencing a split-second ago while reading.” It is no wonder that we started this inquiry with a conflicting menagerie of descriptions. The

armchair methodology is highly prone to the sources of introspective error outlined in the previous section. The way we have gone about studying the experience of reading was doomed to fail, not for any theoretical or a priori reasons, but because of empirical facts that can not be seen from the armchair.

The Descriptive Experience Sampling (DES) method developed by Hurlburt and Heavey (2001; 2004; 2006) is an existing technique for improving the accuracy of introspective reports by reducing the sources of error. On the DES method, subjects are given a small beeper to wear throughout the day. At a random time during a predetermined interval (which occurs while the subject is engaged in daily life) the subject hears a noise from the beeper through a pair of headphones. Subjects are instructed to “freeze” their experience in the moment before the beep and write a detailed account in a notebook. Within the ensuing 24 hours, the subject participates in an expositional interview. The interviewer facilitates a close analysis of the reports to help the subject access their “pristine” experiences via a nuanced navigation through implicit presuppositions and a priori assumptions. As Hurlburt and Heavey quip, “DES is a simple method: Ask what was going on at the moment of the beep and keep out of the way of the answers” (2006, p. 92).

That is the the general approach, now let’s zoom in on the experience samples and the interview process. The experience samples occur randomly during a predetermined timespan, and further, the experiences that are chosen for further scrutiny must be randomly selected as well. This prevents implicit bias or explicit

assumptions about consciousness from skewing the results. For example, someone who holds a “thick” view of consciousness, in which they believe that everything perceptually available exists in their experiences, could favor samples that contain multiple sensory modalities. In fact, they could implicitly reject “thin” reports, such as spacing out, as being defective in some way. This unconscious self-selection could occur during the initial experience sampling or during the selection of the samples to examine during the interview. This is why it is important that random selection occurs at both levels. In addition, the experience samples must concern a specific experience that occurs at a specific moment in time, the “last undisturbed moment before the onset of the beep” (2006, p. 277). The sample and interview must remain constrained by this undisturbed moment, keeping the focus on a concrete time-slice of inner life and putting phenomenal bounds for the subsequent interview.

The interview is not merely a series of predetermined questions. It is more of a facilitated conversation between the interviewer and the subject, who Hurlburt and Heavey call “co-researchers.” The subject reports on their experiences, and the interviewer helps to guide the process. One of the main roles of the interviewer is to bracket presuppositions, that is, flagging the effects of a priori conceptions and expectations in the introspective reports. The two common types of presuppositions manifest themselves as inferences and generalizations. For example, a subject might report that they heard music in the background because the music was playing, not because they were experiencing the music at the

moment of the beep. As I mentioned before, people with a thick view of experience often assume that they are conscious of all of the perceptual stimuli available in their environment. During the Fall of 2005, I participated in a DES study on visual imagery. This style of interview is rare in psychology, and although there is no replacement for experiencing it first hand, reading about it will have to suffice. The following passage describes the nimble questioning that is involved in a DES interview.

For example, the subject 'Ahmed' said during an expositional interview, "I was saying to myself, 'my girlfriend should buy some bananas.'" The interviewer, noting that people don't generally say to themselves "My girlfriend should ..." – they say the much more natural "Jessica should..." – recognized that Ahmed was probably not quoting himself accurately and therefore asked Ahmed, "Exactly what were you saying?" Ahmed replied, "My girlfriend was on the way to the store and I thought maybe I should call her cell phone and tell her to buy bananas." The interviewer, now noting that Ahmed wasn't responding to the "Exactly what were you saying?" question, asked, "Yes, but exactly what words, if any, were you saying?" Ahmed replied, "I'd like to have bananas for a sundae that evening and Jessica could bring them." That again was not responsive to the "exactly what words?" question, so the interviewer continued to press Ahmed for the details of his experience. Ahmed said he was talking to himself, but he was unable to say exactly what the words were; that inability was frustrating to Ahmed. (Hurlburt & Heavey, 2006, p. 231)

True to the idea of subject and interviewer being co-researchers, Hurlburt and Heavey believe that subjects need introspective training to arrive at their pristine experience, just as the interviewer needs training in the DES method. Thus iteration is a core aspect of DES, both within an interview and over the course of a full study, typically 4–8 sampling days. This training requires experience sampling, feedback, new experience sampling, more feedback, etc.

As you can see, the DES method does much to minimize three of the sources of introspective error outlined above. It is specifically designed to counter the tendency to overgeneralize by emphasizing the experiences of concrete moments in time. Although the interview does not take place for up to 24 hours after the beep, the subject records their experiences immediately, thus mitigating the limitations on short-term memory. Like any methodology, DES has its shortcomings, and it does not attempt to control for confabulation. One reason is that the information gleaned from DES is not quantifiable. Hurlburt and Heavey see this as a virtue. It is designed to be idiographic, and while an emphasis on the individual experience is important, it reduces the ability to make general conclusions. The lack of quantifiable results makes it difficult to find the behavioral correlates of introspective reports, objective evidence that a report is not the result of confabulation.

I emphasize the DES method because I share the same set of theoretical concerns that it is designed to address. There are other methods of experience sampling as well. The Experience Sampling Method (ESM) is perhaps the most widely studied (Csíkszentmihályi & Csíkszentmihályi, 1988; Hektner, Schmidt, & Csíkszentmihályi, 2007), and I will return to ESM in Chapter 3. With this background, I propose a research method that combines a focus on retrospective reports of concrete experiences, quantifiable results, and behavioral corroboration of introspective reports. I call it the Subjective and Objective Measures of Experience (SOME) method. I will go into detail on the experimental design in

Chapter 4, but the basic outline is simple enough. While reading a short passage, subjects are interrupted by a beep, at which point they provide introspective reports on their experiences in the split-second before the beep. In addition to these concrete reports on the experience of reading, subjects give general reports of experience before and after the experiment. We can corroborate these reports by comparing them to objective measures such as the subsequent recollection of rhymes, performance on visual detail questions, and so on. To perform this type of analysis we need a much larger subject pool than required by DES. Depending on the effect, we are talking about hundreds or thousands of subjects instead of a dozen. The sheer number of subjects makes the introspective training of DES impossible, but it provides the statistical power to detect subtle behavioral effects that could lend credence to our fallible introspective reports.

Every method has its pros and cons, SOME method included. Hurlburt and Heavey are justified in being skeptical of untrained introspective reports that may be colored by implicit assumptions or presuppositions. In discussing the importance of iteration and training among the DES subjects, they say "...we do not trust methods (questionnaires, one-shot interviews, etc.) that gather data on a single occasion because they do not allow this iterative skill building over time" (Hurlburt & Heavey, 2006, p. 281). However, as the imageless thoughts controversy demonstrated, there are serious methodological issues with introspective training as well. It can easily introduce as much bias as it removes. For this reason, it is important for there to be parallel research designs, some that

involve introspective training and others that do not. It is an empirical matter if trained and untrained subjects will support each other, if they will access different aspects of consciousness, or if their reports will be in conflict. Trained and untrained subjects provide a necessary counterbalance for each other, a check against different sources of introspective error.

To control for the effects of bias, SOME method takes introspective reports before, during, and at the end of an experiment. In this way we can watch for changes in introspective reports as the untrained subjects think about their own experiences in a new way. Further, some reports are general (eg. “Do you experience an inner voice when you read”) while others are specific (eg. did you experience an inner voice in the split second before the beep?”). Like DES, this allows us to compare broad claims about experience with reports of short, specific moments of inner life. To further reduce the distorting effects of memory, specific reports are given within seconds of the beep in a format that allows the subject to respond immediately. Because of this, specific reports should be simple, either emphasizing broad aspects of consciousness, such as the modality of an experience, or a small number of highly specific details. As you may well guess, the set of studies in this dissertation will ask subjects to give specific reports on whether or not they were experiencing an inner voice, visual imagery, or the words on the page.

SOME method will not solve all of our introspective problems, but it does present a solution to some of them. It might seem as if it is a minor change from

the status quo, but notice the difference a small change can make. I began this inquiry by asking what you were experiencing while reading, and then asked you to compare this experience with some of the claims others have made about their own experiences. Putting your momentary introspective report in dialogue with the reports of others is already a methodological improvement over the armchair introspection that dominates the literature. Then we can fine tune this basic idea, making it more systematic by asking carefully worded questions about experience, controlling the timing and type of introspective reports that are given, the type of passage that is read, and the reasons behind reading. We can involve hundreds of subjects to give us the statistical power to test for individual and group differences. Finally, these reports can be cross checked against behavioral measures of experience to ward against inference error and over generalization. Although SOME method is new, it is built upon a solid theoretical and empirical foundation.

As it stands, this discussion is just beginning, and there are still many questions unanswered. The experimental design of SOME method needs to be spelled out in more detail, a task to which Chapter 4 is devoted. Further, given the skepticism Hurlburt and Heavey have toward untrained subjects, we have to be careful in interpreting the results of SOME method. Specifically, we need to be on the lookout for the effect of unconscious presuppositions on introspective reports. Finally, in order to corroborate introspective reports using objective measures of experience, we need a firm grasp of the psychological mechanisms behind the

processing of textual information that occurs while reading, an issue I approach in Chapter 3.

Chapter 3

PSYCHOLOGY + READING

Corroborating introspective reports using behavioral and psychological measures is a key component of the Subjective and Objective Measures of Experience (SOME) method. Because it is impossible to calibrate introspection, introspective reports form isolated islands of data. We can never know whether or not they accurately characterize our experiences. While certainty is not an option, we can still look for the tell-tale signs that our reports of experience have gone awry. A preliminary indication of introspective error is a report that does not match up with the observable world, such as a clash between an external stimulus and the experience we have of it. An example is the wildly confabulated explanations for preferring one stocking among an identical display of in Nisbett and Wilson (1977). We need to cross-reference introspective reports with behavior, cognitive processes, and the world, constantly searching for evidence of error. The grandest instantiation of this idea would involve a triangulation on consciousness, a convergence of introspective, psychological, and neurological evidence. This section attempts a more modest goal. While Chapter 2 focuses on introspective reports, now we turn to the orbiting research from psychology. This is a survey of the relevant work in psychology, and I describe a broad area of research on the cognitive mechanisms involved in the act of reading.

Unfortunately, cognitive psychology has not given widespread attention to phenomenal consciousness since the early days of introspective psychology. The behaviorist criticisms of introspective psychology brought out serious methodological issues with the experimental study of consciousness, but the problems extend beyond difficulties in method. A primary goal of psychology is to provide an explanation for objective empirical observations. For example, in the classic study on inattention blindness, subjects fail to see a man in a gorilla suit while counting the movements of a ball in a short video, a phenomenon typically explained by reference to hard limits on attentional resources (Simons & Chabris, 1999). Such explanations are reductive, explaining a behavioral response in terms of lower-level computational or neural mechanisms. As Chalmers (1996, 2004) argues, this model breaks down when applied to consciousness. According to him, the phenomenal qualities of our experiences can not be objectively observed and, arguably, are not the sort of thing that can be explained with reference to objective mechanisms. This is the “hard problem” of consciousness. There are often attempts to side step this issue by treating introspective reports as observable data, as in the case of Paivio’s (1969, 1971) early work on visual imagery. Although he was interested in the cognitive processes behind reading, not abstract metaphysics, Paivio’s work illustrates this dual rejection and embrace of experience. Writing at a time when consciousness was widely rejected in psychology, Paivio remained awkwardly silent on the existence of consciousness while simultaneously treating introspective reports as his primary independent variable. Yet insofar as

introspective reports of visual imagery assume consciousness — which they must in order to be meaningful behaviors — the hard problem persists. I am not embarking on the reductive project of explaining consciousness in terms of psychological mechanisms, so let's place metaphysics to the side. My goal, in this chapter and the dissertation as a whole, is to find independent evidence that our introspective reports are not mere figments of a collective confabulation. I am interested in the connection between our experience and observable data, a connection that is more likely to take the form of a correlation than an explanation.

There is a wealth of experimental work on factors that affect our actions, both conscious and non-conscious. This literature motivates theories about the cognitive processes that undergird the act of reading, drawing from research on memory, narrative comprehension, and the physiology of reading. This chapter is organized into four sections, each of which discusses experimental literature related to the four most commonly described experiences of reading. The first section focuses on the phonological loop, a subsystem of working memory dedicated to processing linguistic and acoustic information. The phonological characteristics of a word — that is to say, how the word sounds — is an important factor in language processing and provides empirical support for reports of an inner voice. Section two looks at the evidence for a cognitive mechanism that could underlie reports of visual imagery. Specifically, it focuses on the experimental evidence for the visuospatial sketchpad, a component of working memory dedicated to processing visual and spatial information, as well as situational

models of narrative comprehension. The third section turns to the evidence for, and against, the visual experience of the words on the page. I begin with the research on gaze and eye movement while reading, laying out the highly active and organic visual system behind text processing, before countering with a brief description of *flow*, a transcendent mental state in which one gets lost in the moment. Finally, I turn to a small but fascinating literature on mind wandering while reading, discussing the evidence and explanations behind our constantly wandering minds.

We can not directly infer that our experiences mirror the underlying cognitive processes behind reading. As I said, we are looking for correlation, not explanation. This chapter moves quickly between different levels of processing. For example, work on the phonological loop, the visuospatial sketchpad, and gaze all focus on low-level processing, while flow and mind wandering are higher on the informational hierarchy. Low-level processes are, for the most part, automatic, outside of our control, and insulated from introspective meta-awareness. High-level processes are the opposite, voluntary and knowable from a first-person perspective. It is not clear which level is most useful when looking to corroborate introspective reports. I operate on the principle that beggars can't be choosers, and when using experiments developed by researchers who are not interested in phenomenal consciousness directly, we must be opportunistic in the search for relevant research.

I. THE PHONOLOGICAL LOOP

Since the days of Plato reading out loud in the courtyard of the Academy, there has been a connection between words as they are written and words as they are spoken. Experimental psychologists studied this link since the inception of the field. In *Principles of Psychology*, William James describes the first experimental study of an inner voice while reading. True to his background as an introspective psychologist, James emphasized the phenomenology of reading, in this case, the number of distinct sounds uttered by his inner voice while reading a single page.

[Reading] is an uninterrupted and protracted recall of sounds by sights which have always been coupled with them in the past. I find that I can name six hundred letters in two minutes on a printed page. Five distinct acts of association between sight and sound (not to speak of all the other processes concerned) must then have occurred in each second in my mind. In reading entire words the speed is much more rapid. Valentin [1844] relates in his *Physiology* that the reading of a single page of the proof, containing 2629 letters, took him 1 minute and 32 seconds. In this experiment each letter was understood in 1/28 of a second, but owing to the integration of letters into entire words, forming each a single aggregate impression directly associated with a single acoustic image, we need not suppose as many as 28 separate associations in a sound. The figures, however, suffice to show with what extreme rapidity an actual sensation recalls its customary associates. Both in fact seem to our ordinary attention to come into the mind at once. (James, 1980, p. 361).

Psychology has changed drastically in the last hundred and fifty years, and along with it, the way psychologists approach reading. The emphasis on introspective reports has been marginalized in favor of third-person behavioral measures of — not an inner voice — but the effect of phonology in text comprehension, recognition, recollection, and categorization. From a body of research spanning

forty years, there is strong evidence that the way a word sounds when spoken has an effect on the way we process it while reading, even while reading silently to ourselves.

Any discussion of psychology and reading must start with the phonological loop, and in surveying decades of research I walk a well-worn path. The phonological loop is a sub-system of working memory, dedicated to processing linguistic and acoustic information. Working memory is a system responsible for briefly holding and processing information, and it accepts both novel inputs, such as perception, and those already stored, such as long-term memory (see Baddeley, 2007 for an overview of working memory). It is an online workspace, a tool bench and workflow for parsing through information using memory traces that only last a few seconds. While the phonological loop has received the most attention, there is also evidence that the working memory system has two other components, a visuospatial sketchpad for processing visual and spatial information, which we will look at in the next section, and a central executive (Baddeley & Hitch, 1974).

For my purposes, I am less interested in the theories than I am in the experimental data. However you wish to explain our capacity to manipulate and process information on the fly, it is clear that phonology matters, that the sound of a spoken word affects the way our mind interacts with it. Given the wide-ranging function of working memory, everything we read is channelled through the working memory system. I do not want to engage in the various debates between *dual-route models* (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle,

Perry, Langdon, & Ziegler, 2001; Paap & Noel, 1991) and *strong phonological theories* (Frost, 1998). Researchers disagree on the size of the role that phonology plays, and they disagree on whether phonology or visual-orthography is primary in processing text. For our purposes, these debates are tangential. There is consensus that the sound of a word matters when we are reading, and further, that it matters in a basic, automatic, low-level processing kind of way.

The evidence for the phonological loop comes from a wide array of experimental designs, and I can do little more than gesture towards a broad area of study. With that said, I will focus on evidence from phonological interference, phonology in semantic categorization, and phonetic priming.

Phonological Interference: One significant source of evidence for the phonological loop comes from experimental work on the phonological similarity effect. Conrad (1964) found that subjects made “acoustic” errors when recalling consonants that were visually presented. For example, they would mistake ‘B’ for ‘P’ and ‘F’ for ‘S’. Baddeley (1966) modified the design, showing subjects a series of words that were phonetically or semantically similar. The rate at which subjects could recall the ordered list was considerably lower for phonologically similar words (10%) than for semantically similar words (65%) or the control group (71%). This suggests that the words are encoded in short-term memory in an acoustic format. Similarity in the sound of the words leads to interference in later recollection. The phonological similarity effect, as it is been named, has been extensively studied in the subsequent decades, strengthening the case for a

fundamental phonological component to short-term memory. For example, three letter consonant-vowel-consonant non-words (e.g. baf, bor, tuh) are recalled at a higher rate when the sound sequences are common in English (Gathercole, Frankish, Pickering, & Peaker, 1999). Semantics or orthographics alone can not explain this type of result, since it is entirely rooted in the phonology of the three letter strings. Further, similarity in vowel sounds (e.g. dah, fah, gah) leads to poorer recall than similarity in consonant sounds (e.g. dih, dah, doh), a difference that can only be explained with reference to the phonology of the letters (Drewnowski, 1980).

Phonology in Semantic Categorization: A second thread of research looks at the role of phonology in semantic categorization. In these studies, subjects are given a broad semantic category, such as “flower” or “article of clothing,” and are then asked whether or not a target word is a member of the category. Intermixed in the list of words are homophonic foils, words that merely *sound* like members of the category. For example, “rows” is a homophonic foil for “a type of flower,” while “hare” is a homophonic foil for “a part of the human body.” Van Orden (1987) found that homophonic foils lead to false positives in this type of categorization task, such as subjects erroneously calling “rows” a type of flower. Follow up studies found that this effect applies to non-word foils as well, such as “sute” for “an article of clothing” (Van Orden, Johnston, & Hale, 1988). In fact, the rate of false positives was higher for words (and non-words) that sound like a prototypical member of a category, such as “sute,” than for words (and non-

words) that look like a member, such as “surt.” In both studies, subjects were instructed to read for meaning, not for sounds, yet the error rates were the same for both word and non-word foils. This suggests that, paradoxically, when reading for meaning, the sound of the words was more important than the actual meaning (or lack thereof, in the case of non-words). The phonological loop explains this effect nicely. A snap judgment in a linguistic categorization task is processed through the phonological loop which privileges the phonology of a word over the semantics.

Phonological Priming: Another series of studies involves the effects of phonological priming. These studies test if naming a target word is facilitated by priming words that are visually or phonologically similar to an “appropriate” prime. For example, “beech” is an appropriate prime for “tree,” and seeing “beech” aids in subsequently naming “tree.” While some subjects are given an appropriate prime, others are primed with orthographic or phonological variations, such as “bench” or “beach.” Lesch and Pollatsek (1993) found that the homophonic variation “beach” facilitated naming “tree” at 50 ms, but not at 200 ms, while the appropriate prime “beech” facilitated at both times. Lukatela and Turvey (1991) found similar results, including a priming effect for non-words. This study found that viewing “table” facilitated the naming of “chare,” and conversely, viewing “tayble” facilitated the naming of “chair.” It is perhaps no surprise that words can prime us in categorization tasks, but it is surprising that non-words can do the same, provided that they sound like an appropriate prime.

Together, these studies conclude that phonology plays an important role in early text comprehension, but after about 200 ms, semantics and orthography compete with the phonological processing. I am not so concerned with the relative importance of these linguistic features as much as with the deep role that phonology plays in reading.

Additional Evidence: As these examples show, the evidence for the phonological loop is broad and well established, and the effect of phonology on reading is robust across multiple experimental paradigms. I merely want to convey a taste of the breadth of this research. Treat this as a sampling platter: During silent reading, homophonic pairs, such as “sole” and “soul,” are read at identical speeds, even when the context highly favored one over the other (Folk, 1999). As far as reading speed goes, it is as if a subject does not register error from a semantically-incorrect-but-phonologically-similar word. The Stroop effect, in which subjects are slower at reading color words when the word itself is printed in a color other than the one named (e.g. ‘red’ written in green text), has a phonological relative. Subjects are slower at identifying the color of a word when the word itself is phonologically similar to an incongruent color, for example, “grean” written in red (Dennis & Newstead, 1981). Further, when a target word is briefly shown (15–30 ms), subjects are more likely to identify it if it is shortly (15–60 ms) followed by a non-word that is phonologically similar than for a non-word that is orthographically similar (Berent & Perfetti, 1995; Perfetti & Bell, 1991). Thus, the priming effect works both before and after the presentation of a word.

All of this, by the way, is in addition to a large literature in developmental psychology on the role of phonology in language learning (Baddeley, Gathercole, & Papagno, 1998), vocabulary learning (Gathercole & Baddeley, 1990; Gathercole, Willis, Emslie, & Baddeley, 1992), and in second language acquisition (Papagno, Valentine, & Baddeley, 1991).

Taken together, these studies form a compelling argument for a dedicated subsystem within working memory that is devoted to processing linguistic and acoustic information. This is precisely the sort of cognitive mechanism that can act as independent corroboration for reports of an inner voice. In this way, there is a beautiful synchronicity between first and third person evidence. Our cognitive processes and introspective reports mirror each other; they both privileging the phonology of words while reading. We mistake words that sound alike for each other in categorization tasks, in fact, we even mistake fake words as long as they are phonologically similar to real ones. We can be primed to identify “tree” by quickly seeing —not the name of a type of tree — but a word that merely *sounds* like a type of tree. We do not even slow down while reading semantically confused texts as long as the words sound as they should. Just as the sound of a word often effects the experiences people report while reading it, the sound of a word affects the way it is processed in working memory.

The reflection of consciousness in cognitive processes goes further. A common characteristic of the experience of inner voice is that it is, to some extent, voluntary. Similar to cases of directed attention, many people report being able to

summon an inner voice into existence if they choose. The phonological loop shares a similar measure of voluntariness. Although auditory information has automatic access to the phonological loop, we recruit the phonological loop to process linguistic (but non-auditory) information when we read silently. We can also consciously choose to engage the phonological loop by repeating a single word silently to ourselves, a process called articulatory suppression, and linguistic information is processed differently while performing articulatory suppression. Recall that the cause of the phonological similarity effect is interference between disparate sources of phonological information in the phonological loop. However, when participants perform articulatory suppression, they can engage and overwhelm the phonological loop, causing linguistic information to be processed through other channels. In such cases, the phonological similarity effect disappears, and short-term memory for phonologically similar words becomes the same as for phonologically dissimilar words. However, while articulatory suppression eliminates the phonological similarity effect for words presented in writing, it has no effect for words presented auditorily (Baddeley, Thomson, & Buchanan, 1975; Baddeley, Lewis, & Vallar, 1984). Thus, while auditory information is automatically routed through the phonological loop, we can also engage the phonological loop by reading or silently talking to ourselves, but not both. Since we can, at any time, choose to silently talk to ourselves, we can choose to engage the phonological loop.

While the existence of the phonological loop is not controversial, it is not easy to apply this research to our experiences. A distinct cognitive mechanism is not required for the experience of an inner voice (or for the introspective report of an inner voice), but it lends credibility to these reports. If the way a word sounds has an effect on the way we process it while reading, then the psychological prerequisites exist for the experience of an inner voice. At the very least, this shows that phonemes — and not just completed words or concepts — are meaningful units to the cognitive processing that underlies reading. If phonemes are important to the cognitive processes while reading, they could manifest themselves in our experience.

II. THE VISUOSPATIAL SKETCHPAD

Visual imagery is a polarizing topic in experimental psychology and has had a profound effect on the field as a whole. In what may be the first psychological questionnaire, Francis Galton (1880) famously asked distinguished Victorian men to describe their visual imagery of that morning's breakfast table. Their responses were a premonition of the descriptions of the experience of reading given in Chapter 1. Descriptions varied wildly, ranging from reports as detailed as visual perception to a complete denial of visual imagery whatsoever. Decades later, visual imagery was instrumental in the collapse of the first dynasty in experimental philosophy. The so-called imageless thought debate, the intractable conflict between psychology labs that reported pervasive visual imagery and those that

denied it, lead to the widespread rejection of the methods and domain of introspective psychology as a whole. Thus visual imagery was used to found the discipline, was the focus of the first experimental research methods, and was later the reason that these methods were discarded.

Even with the rejection of consciousness that occurred mid-century, conflict around visual imagery persisted, adapted to the psychological climate of the time. The “imagery debate” between Kosslyn and Pylyshyn did not concern the experience of visual imagery, but rather the cognitive architecture underpinning the way we represent and process images. Kosslyn (1980, 1994) argued that visual imagery is represented pictorially (or rather, quasi-pictorially), while Pylyshyn (1973, 2002) made the case that images have a language-like structure. The debate has lessened considerably from its height, but the issue is still not settled (Dulin, Hatwell, Pylyshyn, & Chokron, 2008; Lewis, Borst, & Kosslyn, 2011). Although this is some of the most famous experimental research on imagery, much of it has little to do with visual imagery as I have been using the term. The other name of this debate, the “analogue-propositional debate,” highlights why this is the case. At bottom, the conflict concerns the deep structure of thought, with “picture” and “language” acting as metaphors for competing functional descriptions of cognition within the context of a computational theory of mind. This structure is so far removed from measurable behavior that Anderson (1978) argued that both theories are empirically equivalent. This led to a spirited back and forth that feels uncommonly personal and emotional for a debate in psychology (Hayes-Roth,

1979; Pylyshyn, 1979; Anderson, 1979). I do not find Anderson's argument entirely persuasive, but the point remains: the issue at stake in the imagery debate remains abstract. As far as consciousness is concerned, it does not matter if our minds represent information isomorphically or through sentences in mentalese.

Attempts at combining these two historically important areas of research — reports of visual imagery from the introspective psychologists and recent experimental work on the cognitive mechanisms underlying visual imagery — have been underwhelming. Over the years, there have been a number of attempts at quantifying and operationalizing mental imagery, including the Questionnaire upon Mental Imagery (Betts, 1909) and the Test of Visual Imagery Control (Gordon, 1949), but the most influential has been Mark's (1973) Vividness of Visual Imagery Questionnaire (VVIQ). Early reviews of the literature rejected a meaningful connection between the VVIQ and objective behavioral measures (Ernest, 1977; Richardson, 1980). Consider the review by Allan Paivio who, perhaps more than anyone, was responsible for reintroducing visual imagery into a psychology culture that largely rejected consciousness. He accomplished this by operationalizing visual imagery, reducing experience to a numerical rating that allowed him to walk a fine line, simultaneously assuming consciousness while treating introspective reports like any other variable in the subsequent behavioral analysis. In a series of studies (1969, 1971) he showed that words rated as “arousing sensory images” were more likely to be remembered than those that are not. However, even Paivio (1986) concluded that there is no a meaningful

connection between the VVIQ and measurable behavior. Despite this, the idea of a relationship between visual imagery and visual processing is so intuitive that research continued, and the conclusions from recent work have been more mixed. In perhaps the rosier summary of the literature, Stuart McKelvie (1995) concluded that a definitive judgment on the VVIQ requires more research, basing this decision on a review and meta-analysis of over 200 studies. However, given that the VVIQ only comprises 16 questions that fit on single page, the failure to find compelling evidence after decades of research and hundreds of studies should properly be viewed as a failure of the research project as a whole (Schwitzgebel, 2011, p. 44–48).

It is important to remember that the VVIQ is but a single test of visual imagery. Although it is widely used, the failure to validate it with experimental evidence does not entail doom for the project of corroborating reports of visual imagery with measurable behavior. However, the history of the VVIQ should cause us to proceed with caution. Even in the absence of a significant connection between reports of visual imagery and behavior, we can still look at the role of images (or quasi-images) in our cognitive processing. With that preamble, I will now turn to research on the visuospatial sketchpad, a separate subsystem of working memory that operates in parallel with the phonological loop. I will conclude the section with an overview of situational models of text comprehension, which hold that reading comprehension involves the formation of mental models of a narrative text.

The visuospatial sketchpad is a component of working memory that operates in parallel with the phonological loop. Whereas the phonological loop is dedicated to linguistic and acoustic information, the visuospatial sketchpad operates over visual and spatial information (Baddeley, 1986). Our knowledge of the visuospatial sketchpad is considerably more limited when compared to its linguistic counterpart. One cause is the difficulty in specifying variables and properties when working with images, where it is not always clear what spatial or visual similarity even amount to. Separating out the visual component of the visuospatial sketchpad, visual memory can hold about four objects and is relatively robust, that is, intervening tasks have little effect on forgetting. For example, an intervening spatial search task does not degrade subsequent recollection of visual properties (Woodman, Vogel, & Luck, 2001). Spatial information, however, is highly affected by intervening tasks, and recall of spatial information decreases dramatically when performing a visual search (Dale, 1973; Phillips, 1974; Woodman & Luck, 2004; Oh & Kim, 2004). Although the data is less clear, there is also some evidence that the visuospatial sketchpad processes kinesthetic information as well (Smyth, Pearson, & Pendleton, 1988). For my purposes, the specific details of the visuospatial sketchpad are mostly garnish. The bulk of the dish is the evidence for a distinct cognitive mechanism dedicate to processing visual and spatial information.

Research on the visuospatial sketchpad is varied and somewhat disconnected. I begin chronologically with an early and influential paradigm from

Brooks (1967). In this study, Brooks tested short-term memory for a series of sentences. One set of sentences map onto a four-by-four matrix, and subjects were encouraged to use an imaginary spatial matrix as a mnemonic device to aid retention. A second set of sentences were the nonsense equivalents, replacing the spatial words with meaningless adjectives (Figure 3.1).

Recall was recorded for each set with the sentences presented auditorily or both auditorily and visually. For the spatial set, subject made 1.2 errors when hearing the sentences, increasing to 2.8 errors when hearing and reading them. The results were reversed for the nonsense set. Subjects made 2.3 errors when the sentences were presented auditorily compared to 1.3 errors when they were presented auditorily and visually. Brooks concluded that reading and visualizing utilize the same cognitive processes, and this creates interference in short-term memory when both are employed simultaneously. However, when spatial information is not relevant, as in the nonsense sentences, there is no interference. In this case, reading aids short-term memory.

		3	4
	1	2	5
		7	6
		8	

Spatial Set

- In the starting square put a 1.
- In the next square to the right put a 2.
- In the next square up put a 3.
- In the next square to the right put a 4.
- In the next square down put a 5.
- In the next square down put a 6.
- In the next square to the left put a 7.
- In the next square down put an 8.

Nonsense Set

- In the starting square put a 1.
- In the next square to the quick put a 2.
- In the next square to the good put a 3.
- In the next square to the quick put a 4.
- In the next square to the bud put a 5.
- In the next square to the bud put a 6.
- In the next square to the slow put a 7.
- In the next square to the bad put an 8.

Figure 3.1

The stimuli from Brooks (1967), including the spatial set of sentences, the nonsense set of sentences, and the four-by-four matrix used as a mnemonic device.

This formed the early evidence for a component of short-term memory that is responsive to visuospatial, but not auditory, interference. Later research explored the types of interference on the Brooks task. For example, spatial tapping decreased performance on the reading and listening version more than the listening only version, while articulatory suppression has the opposite effect (Salway & Logie, 1995; Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Farmer, Berman, & Fletcher, 1986). The types of interference have been expanded to include kinesthetic interference such as pointing (Brooks, 1968) and arm movement (Lawrence, Myerson, Oonk, & Abrams, 2001; Quinn & Ralston,

1986). All of this suggests two separate cognitive systems, one for processing phonological information and the other for visuospatial information. Additional evidence comes from neuropsychological patients. There is a group of patients who have a standard memory for digits but an impaired spatial memory, while another group has the opposite deficiency, impaired digit span and standard spatial memory (DeRenzi & Nichelli, 1975; Della Sala & Logie, 2002). This double dissociation strongly suggests distinct cognitive processes.

As with the phonological loop, the visuospatial sketchpad exhibits the same one-way voluntariness as the experience of visual imagery. Most people report being able to conjure visual imagery at will, requiring nothing more than the simple desire for a mental picture. In Garden, Cornoldi, and Logie (2002), subjects learned to navigate a path through the center of Padova, a medieval Italian city. Later they had to repeat this path while performing a concurrent task, either articulatory suppression or spatial tapping. Subjects who reported using a spatial map-like strategy for navigating were highly effected by spatial tapping, making twice as many errors than they did during articulator suppression. The opposite was true for the subjects who did not report using a spatial map-like navigation strategy, and these subjects committed twice as many errors while performing articulatory suppression than while engaged in spatial tapping. This is evidence that the conscious strategy used to process and retain spatial information affects the underlying cognitive mechanism, as displayed by the relationship between navigation strategy and interference from auditory or spatial tasks.

The visuospatial sketchpad comes directly out of the research on short-term and working memory. Contrast this with the research on the VVIQ, experimental work that is attempting to validate an already existing test. Put this way, the VVIQ is something of an explanandum in search of an explicans. From a subjective perspective, it feels like visual imagery is something we know while the cognitive mechanisms underlying it are a mystery, but from an experimental standpoint, the situation is reversed. We have strong evidence that our cognitive structure has a privileged, dedicated system for processing visual and spatial information, while the data on our experience of visual imagery is incomplete and inconclusive.

Another thread of research related to reports of visual imagery comes from the work on situational models of text comprehension. According to situational models, reading is primarily about the content of the text, not the words used to express the content. When we understand a text, we do not just understand the words and how they are put together, we also understand important aspects of the situation described. This might sound trivial, but prior to the 1980s, when the early experiments on the situational model were conducted, the dominant view in psychology held that reading was a matter of representing and retrieving information about the text itself. Reading was considered a purely linguistic act, similar in some ways to a game of Scrabble. Many world Scrabble champions are non-native English speakers. Because Scrabble is not a language game in the Wittgensteinian sense, it has nothing to do with language as an action or use. Scrabble is a game of vocabulary memorization. Similarly, the early psychology

theories of reading comprehension held that reading was a matter of processing the linguistic information in the text. Once the text could be read by the mind's eye, it could be unpacked into a meaningful scenario. The scenario itself, however, was secondary. The insight of the situational model is to collapse this hierarchy, arguing that reading, at least the reading of narratives, fundamentally involves creating a mental model of the situation described.

Some of the most compelling evidence for the situational model comes from research on processing the spatial content of texts while reading. Language is linear while space is not, so there is the potential for a contrast between the words describing the spatial aspect of a situation and the spatial aspect of the situation itself. The mismatch between linguistic structure and spatial organization gives us an opportunity to examine the underlying mechanisms at play (Zwaan & Radvansky, 1998).

The role of foregrounding in narrative comprehension illustrates this mismatch. A series of studies found that information that is foregrounded is more readily available for retrieval in subsequent categorization tasks. For example, subjects can access information about a car more quickly if a narrative discusses a character getting into a car versus getting out of a car. In Glenberg, Meyer, and Lindem (1987), half of the subjects read narratives that associated the protagonist with a target object ("John put on his sweatshirt before going jogging") while the other half read narratives that dissociated the protagonist from the target object ("John took off his sweatshirt before going jogging"). Afterwards, subjects

responded when they recognized the target word. Response time was longer for subjects given the dissociated narrative even though, for both groups, the target word occupied the same location in the text. A theory of reading comprehension that focuses exclusively on textual elements can not account for this. The two narratives are virtually identical from a linguistic standpoint. However, from the perspective of emphasis and action, there is a large practical differences between them.

In another series of studies, subjects were shown a narrative in which a protagonist moved around a building after previously memorizing a map of the space (Morrow, Bower, & Greenspan, 1989; Morrow, Leirer, Altieri, & Fitzsimmons, 1994; Rinck, Williams, Bower, & Becket, 1996). Subjects were given word pairs (e.g. “Mary” (the protagonist) and “lamp”) and asked if they were in the same room. The responses were fastest when the target object was in the same room as the protagonist, slower when the target object was in the room the protagonist had recently left or was about to enter, and slower still when the object was not on the protagonist’s path. In other words, response time was mediated by the spatial distance between the target object and the protagonist. The effects of foregrounding are not limited to categorization. Reading time and response time slow as the target object gets farther away from the protagonist, regardless of the locations of the target words in the text (Glenberg et al., 1987; Singer, Graesser, & Trabasso, 1994). Further, reading time is slower when information in a text is spatially inconsistent with previous information (de Vega, 1995), and reading

speed is faster when descriptions are spatially continuous than when they are spatially discontinuous (Ehrlich & Johnson-Laird, 1982).

These studies make the case that the spatial content of a narrative changes the way we process it. While reading, we do not simply put together the linguistic pieces; we construct a mental model that includes the spatial characteristics of the situation described. We have not been discussing the experience of reading directly, but nevertheless, this is evidence that spatial facts of a scene are a psychologically important component of reading comprehension. Similarly, if my eyes are open and my visual cortex is lighting up, we have good evidence that I am experiencing some sort of visual experience. But bear in mind, this is not definitive anymore than the informational suppression that occurs during a saccade entails that our visual experience blinks in and out every time we move our eyes.

Given the broad scope that I am interested in — any experiences someone has while reading — the situational model needs an important disclaimer. There is not evidence of a mental model in all instances of reading; motivation and purpose matter. For example, subjects who are instructed to proofread have a reduced ability to make causal inferences that depend on general knowledge. Subjects normally respond faster to “Do dentists require appointments?” after reading “Terry was unhappy with his dental health; he phoned the dentist," than after "Terry was unhappy with his dental bill; he phoned the dentist." When instructed to proofread, this difference disappeared (Singer & Halldorson, 1996). Zwaan and Radvansky (1998) argue that situation models are not necessary for all language

processing tasks, but they are an integral part of all language comprehension tasks. Similarly, Graesser, Singer, and Trabasso (1994) argue that situational models are only created when there is a “search (or effort) after meaning,” and that the mental representation of a text depends on the reader’s goals. Since proofreading is still a type of reading, this is a reduction of the domain that the situational model applies to, a contraction to reading for comprehension of the content. This should probably be reduced further still, since there is no evidence that a mental model is formed when reading non-narratives. Articles on abstract philosophy, newspaper headlines, restaurant menus, and movie credits all lie beyond the empirical evidence for situation models.

Stepping back a moment from the empirical literature, a limited application of the situational model is reasonable theoretically. It takes cognitive resources to construct a mental image. There are presumably times when these resources might not come into play, for reasons of attention, time, or competing loads. If a mental model would not assist in reading comprehension, it is likely that it is not invoked. This points to more than one cognitive mechanism behind reading comprehension, a pragmatic recruitment of cognitive resources that squares nicely with the varied experiences people report while reading.

III. THE VISUAL SYSTEM AND FLOW

While reading, it is necessary to see the text, at least in the sense of eyes-looking-at-words. This is a defining characteristic of reading. Given this, it is

surprising how rarely people report experiencing the words on the page. Maybe this is because the perceptual experience is simply assumed, something so obvious as to be beyond mention. Or maybe it is because subjects “take off” into the narrative, transcending the merely perceptual experience in favor of something more immersive, perhaps vivid visual imagery or an auditory enactment of the narrative. Both of these possibilities have experimental evidence in their favor, and in fact, they are not mutually exclusive. We may seamlessly roll between them without a meta-awareness of the shift. This section is devoted to evidence in favor of reports (or lack thereof) of the words on the page. First, I will describe basic details of our visual system while reading. The perceptual processes involved in reading are incredibly complex, from the physiological control of eye fixation to the complicated processing that occurs afterwards. Second, I will turn to the research on flow to outline empirical work on “taking off” into a textual narrative.

Our visual system involves an organic feedback between physiology, low-level cognitive processes, and conscious intentions. While participating in any visual perceptual task, our eyes are continuously in motion, making short, quick ballistic movements called saccades. Interspersed with the saccades are moments of fixation when our eyes remains still for 200–300 ms. Because the contents of the entire visual field is in motion during a saccade, our visual system blocks additional input, a phenomenon called saccadic suppression (Matin, 1974). Since saccades are ballistic movements, the time they take is a direct function of the distance traveled. While reading, a 2° saccade is common and lasts around 30 ms.

During scene perception, however, saccades are often 5° and require 40–50 ms (Abrams, Meyer, & Kornblum, 1989; Rayner, 1978). In addition, there is a latency period before each saccade, a time of planning and execution that lasts 150–175 ms (Abrams & Jonides, 1988; Salthouse & Ellis, 1980; Salthouse, Ellis, Diener, & Somberg, 1981). Our visual field is broken up into three main sections: foveal, parafoveal, and peripheral. Although the fovea is quite small, just the central 2° of our visual field, it is the most detailed area. Detail decreases in the parafovea, the central 10° of fixation, and decrease further still in the peripheral field. The continuous pattern of saccade, fixation, saccade is all for the purpose of getting relevant information into the detailed fovea.

Because our visual system is responsive to the scanning task and features of the visual field, there is a unique algorithm for reading. In addition to forward saccades, our eyes often perform backward saccades, called regressions, an important step in a complicated system of content extraction and re-extraction. From the center of focus, foveal detail extends about 4 spaces to the left and 15 spaces to the right, although the exact span decreases as reading difficulty increases. Fixation lasts 200–250 ms and an average saccade is 7–9 letters, about 2° of our visual field. We bounce from one side of the fovea to the other because reading is impossible without the detailed resolution it provides (Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). The visual system is also responsive to textual content, and although most words are fixated while reading, 15% of content words and 65% of function words are skipped

entirely (Carpenter & Just, 1983). The visual system does not merely translate the world into mental representations, it is constantly making inferences and filling in gaps that we are not aware exist.

A family of ingenious experiments using eye trackers provides one of the main sources of information about our visual system. The classic optical eye tracker looks like a machine you would find in an optometrist's office. While you look through the lenses, it measures the reflection of light off of various parts of your eyes, and are quite accurate at measuring the precise point of fixation. The moving window and foveal mask techniques are influential paradigms for using an eye tracker to measure visual span (Figure 3.2). In both techniques, as the subject reads, text is replaced in real time during saccades. Because of saccadic suppression, the subject is not aware of the change. The moving window technique tests visual span by changing the text outside of a window around fixation (McConkie & Rayner, 1975), and visual span is determined by counting the number of spaces at which the subject is unaware of the changes. The foveal mask technique is the opposite, changing text only within the bounds of fixation, and visual span is the number of spaces for which the subject can not read the text (Rayner & Bertera, 1979).

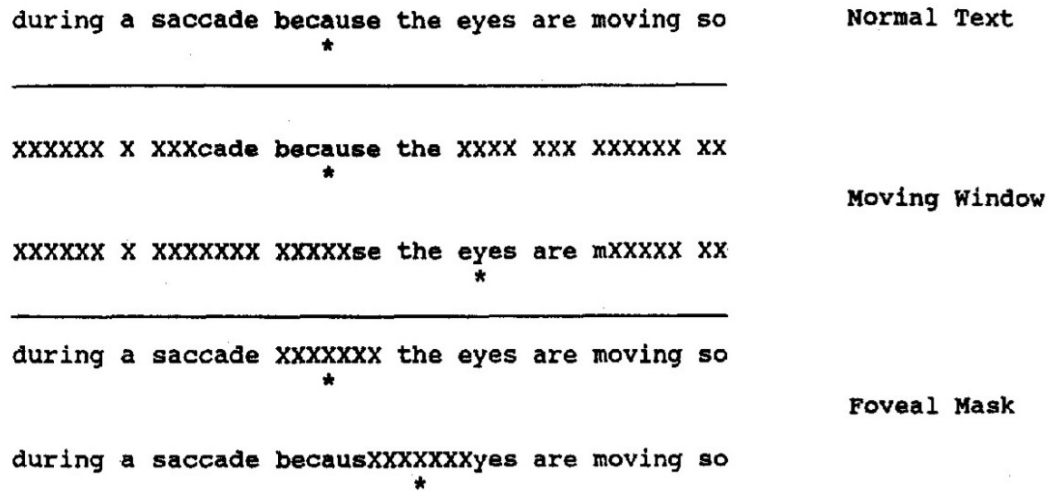


Figure 3.2
Examples of the moving window and foveal mask techniques of measuring visual span from Rayner (1998). The asterisk represents the point of fixation.

In some ways, this research only clarifies what we already know: that our eyes move across the page, as we read, left to right, line to line. Yet this detail is important, filling out the psychological foundation that underlies the experience the words on the page while reading. Our visual system is highly complex, the outputs of which are a curated version of the information our perceptual system takes in. Once again, however, we must not draw straightforward conclusions about our experiences from this data alone. Our eyes are constantly saccading back and forth across the text, and because of the results from decades of experimentation, we know the reason: we can only perceive detail within the relatively small fovea. Based on this research, Dennett (1991) argues that our visual experiences are like viewing the world through a peephole, with detail only

extending to the central area of fixation. He is correct that we do not notice changes outside of the fovea. However, this does not entail that large swaths of text lie outside of our experience. We could, for example, quickly build a mental sketch of the page, a sketch with a high degree of resolution, even if much of the detail is not responsive to the artificially changing text found in an eye tracker experiment. I am not sure how compelling this story is, positing more information in the visual field than our visual system is processing in real time. However, consider the converse, that we often do not experience the words on the page at all. This is an entirely different type of disconnect between the visual system and introspective reports, a complete denial of conscious perception while reading. This type of report is not uncommon, and demonstrates once again the problem with inferring experiences based on behavior of the visual system alone.

Can this research shed light on the vast amount of variability in the ways people describe their experience while reading? The studies on eye movement show a moderate amount of variability in the way text is processed by our visual system. The algorithm employed by our visual system is not fixed, and is responsive to high-level features of the text and our intentions behind reading. For example, our visual system changes its scanning pattern based on the difficulty of the text. As the text becomes more difficult, there is an increase in fixation times, decrease in the length of saccades, and more frequent regressions (Jacobson & Dodwell, 1979). Our visual system is also responsive to the way we are reading, and when compared to silent reading, reading out loud leads to longer fixation as the eye

stalls between saccades while the voice catches up (Lévy-Schoen, 1981). There are also interpersonal differences between readers. Compared to slow readers, fast readers make shorter fixations, longer saccades, and fewer regressions (Everatt, Bradshaw, & Hibbard, 1998; Everatt & Underwood, 1994). Even for a single reader, there is variability within a passage, and fixation ranges from 100–500 ms while saccades vary from 1 to 15 letters. Are these the sort of individual differences that can account for the differences in introspective reports? As we see throughout this discussion, the answer is not obvious. It is one thing for there to be a difference of up to 400 ms in the length of fixation, another for someone to experience the words at one time and to cease experiencing them at another.

The claim that people do not experience the words on the page seems contrary to everything we know about the visual system. Our visual system is highly active, constantly taking in and processing information, regardless of what people say they experience. There is, however, an explanation for the denial of visual experience, an explanation that comes out of work in positive psychology on flow. Up till now, this chapter has focused on low-level mental processes, such as phonological priming, an effect that disappears after a mere 200 ms. Now we turn to experimental research on experience itself. Mihály Csíkszentmihályi pioneered the work on flow through the study of artistic creation, which he then expanded to include “autotelic activity,” that is, activity that people engage in for its own sake.

Flow is the experience of getting lost in an action, complete absorption in the process, the “merging of activity and awareness” (Csíkszentmihályi, 1988, p.

32). In his work, Csíkszentmihályi is clearer when describing the experience of flow than when giving it a firm definition, saying things such as “we have all experienced times when, instead of being buffeted by anonymous forces, we do feel in control of our actions, masters of our own fate. On the rare occasions that it happens, we feel a sense of exhilaration, a deep sense of enjoyment that is long cherished and that becomes a land-mark in memory for what life should be like” (Csíkszentmihályi, 1990, p. 39). Reports of the experience of flow come from disparate sources: rock climbers, painters, sailors, composers, and parents. It might seem as if optimal experience is too grand for something as everyday as reading, yet, in his research, reading is the most commonly mentioned flow activity (Csíkszentmihályi, 1990, p. 117). Csíkszentmihályi (1988) found that 7–45% of subjects reported flow while reading, one of the four most commonly cited flow activities, along with hobbies, sports and studying.

This is all rather abstract, and it is difficult to integrate the work on flow with the rest of this chapter. Although Csíkszentmihályi is interested in the experience of reading, he is working within the larger context of optimal experience, and in doing so focuses on aspects of reading that are shared with a wide range of other activities, such as rock climbing and parenting. However, when we look at the descriptions given by individuals in his studies, subjects speak in terms that should sound familiar. One student from Turin said “I feel as if I belonged completely in the situation described in the book,” while another stated

“I identify with the characters, and take part in what I am reading” (Csíkszentmihályi, 1988, p. 73).

Regardless of the exact nature of flow — its metaphysics or its phenomenology — there is something to the idea of being “in the zone” while reading. Jennings (2015) approaches the phenomenon from a different angle, bridging the gap between this literature and the philosophical work on consciousness. While flow is inherently connected with positive affect and requires conscious attention, Jennings makes the case for a related experience called “conscious entrainment,” a form of consciousness we enter when focused on a task that does not require attention because of habituation and practice. Her examples of this experience, skilled athletes operating at peak performance, overlap with the paradigmatic examples of flow. Csíkszentmihályi and Jennings are both interested in the same basic experience, and although they emphasize different aspects, their views are compatible.

For Jennings, conscious entrainment has three defining characteristics: First, it requires total focus, a complete recruitment of cognitive resources towards a task. Second, it must be effortless and beyond conscious control, an automatic process that becomes impossible with top-down direction. Jennings cites descriptions of experience from professional athletes who report a decrease in performance when they exert conscious oversight of their actions. Finally, conscious entrainment requires the lack of a divide between subject and object. Together, the person and tools of the activity must form a coherent, unified

experience. The requirements of conscious entrainment should sound familiar, and share much in common with the experience of reading. Skilled readers are performing a task that has become habituated through practice and repetition. It requires a total focus on the text, often without meta-awareness or conscious control, and readers often become so captured by a story that they do not report an experience the object, in this case, the words on the pages of a book.

The concepts of flow and conscious entrainment are fascinating, and capture a widespread and important aspect of experience. They exist intuitively, and do not rely on complicated theoretical or conceptual frameworks. That being said, applying these concepts leads to some counterintuitive results. For example, when people are in heightened conscious states, they often do not report basic perceptual experiences. The ready explanation is an appeal to a form of inattentional blindness, but is this plausible? Do we want to say that a basketball player, in flow or conscious entrainment, is unaware (or not conscious) of the ball? This sounds more reasonable when describing a reader who reports no experience of the words on the page, and the difference could be due to the physical actions associated with dribbling and shooting. How could someone who displays such skilled movement be experiencing attentional blindness? A reader, on the other hand, is just sitting there, book or computer in hand. Unfortunately, research on this sort of perceptual phenomenology is sparse, so the experimental literature is of little help.

IV. MIND WANDERING

Mind wandering is the liminal case of the experience of reading. It does not seem particularly associated with reading per se, in fact, it seems as if it is a perversion of the experience, an example of consciousness losing its way. Saying that spacing out is a part of the experience of reading is like saying going to the bathroom is a part of the experience of a movie. But is this so wrong? As I have argued in Chapter 2, the experience of reading is an action experience, and all experiences while reading have an equal claim to legitimacy. If you are truly interested in what people experience while reading, even thoughts that are entirely unrelated are important, especially if they occupy a sizable amount of our inner lives.

Although mind wandering is rarely cited in the grand descriptions of reading, there is experimental evidence that much of the time we are ostensibly reading is actually spent with our minds somewhere else. The research on mind wandering while reading is relatively young, and only began 25 years ago with two studies from a single psychology lab (Giambra & Grodsky, 1989; Grodsky & Giambra, 1991). Schooler, Reichle, and Halpern (2004) built upon this work and designed an experimental method that combines introspective reports with objective measures of behavior. They probed subjects while reading to see if they were having task unrelated thoughts (TURT). A task unrelated thought is exactly what you imagine: a thought that is not directly connected with the task that is instructed. Our minds do not always follow orders when told, and we often have

intrusive thoughts, moments of mind wandering, and periods of spacing out. For example, if asked to read a passage, thoughts about relationships, hunger, and what is happening later in the evening are all TURT. The interesting psychological work lies in quantifying how much of our mental life is spent this way. During their first experiment, subjects reported zoning out 5.4 times while reading for 45 minutes. What kind of things were occupying their consciousness while zoning out? 27% of the time subjects reports school related topics, 19% fantasies, 11% themselves, and 18% of the time subjects reported experiencing nothing at all. Interestingly, subjects were asked if they were aware of their mind wandering, and remarkably, 13% of the time they were not. Schooler and his collaborators conclude that we are often spacing out while reading and, further, we are largely ignorant of this phenomenal absence. This sheds light on the absence of reports of mind wandering from Chapter 1. Especially if we assume that, on the occasions someone has the meta-awareness to realize they are spacing out, the experience is dismissed as a perversion of consciousness and not an attribute of it.

Subsequent research has delved deeper into the causes of mind wandering. Dixon and Bortolussi (2013) probed subjects while reading one of two texts, Rice's *Interview with a Vampire*, deemed "interesting," and the other "less interesting" text, Thackery's *The History of Pendennis*. Unsurprisingly, subjects reported more mind wandering while reading the less interesting text. Further, narrative recall was considerably lower for Thackery's boring prose than for the lively writing of Rice, even when controlling for mind wandering. Kopp, D'Mello, and Mills (2015)

found that background thoughts influence mind wandering. Subjects who were instructed to think about their short-term plans by making a to-do list were significantly more likely to report TURT and had lower narrative comprehension than the control group who were told to simply list the parts of an automobile. Other causes of spacing out are working memory capacity, topic interest, and motivation (Unsworth & McMillan, 2013). Finally, Franklin, Mooneyham, Baird, and Schooler (2014) compared mind wandering for reading silently and reading out loud. They found that although reading comprehension was the same across both conditions, subjects reported more mind wandering while reading out loud. This is somewhat paradoxical, since reading out loud would seem to require an extra step beyond reading silently. Interestingly, they detected “subtle vocal signatures” that predicted both mind wandering and reading comprehension, signifying some success in corroborating introspective reports with bona fide measurable behavior.

This last point signals a rare success for the small body of research on mind wandering. It has managed to find replicable and measurable effects that corroborate introspective reports. Schad, Nuthmann, and Engbert (2012) inserted nonsense sentences into an otherwise meaningful text and measured the error detection of readers. As reports of mind wandering increased, the rate of error detection decreased. In addition, error detection was positively correlated with a decrease in fixation duration, and they developed an algorithm that could use an eye-tracker to predict an overlooked errors 5 seconds before it occurred. Further,

there is increased pupil dilation during reports of mind wandering (Franklin, Broadway, Mrazek, Smallwood, & Schooler, 2013). Franklin, Smallwood, and Schooler (2011) had subjects advance a text one word at a time while reading. While spacing out, subjects reading speed increased slightly and they became less responsive to linguistic and orthographic properties of words, such as word length, syllable number, and familiarity. Together, these studies provide strong evidence that there are observable behavioral measures of introspective reports of mind wandering.

All of these studies reported a significant correlation between reports of mind wandering and lower reading comprehension. Mind wandering is a case of (mis)directed attention, where the complicated mental system recruited for reading comprehensions drifts off into other topics. Smallwood, McSpadden, and Schooler (2008) argue that spacing out inhibits the creation of situational models employed for narrative comprehension. Subjects read a Sherlock Holmes novella and reported on their experiences along the way. Subjects who reported spacing out did not exhibit the effects of foregrounding, suggesting that spacing out inhibits the creation of a mental model of the narrative. This was compounded when the spacing out occurred at the beginning of the text. When this occurred, subject missed out on foundational textual elements. Smallwood et al. concluded that subjects who reported spacing out experience a “double whammy.” They miss textual information as well as information conveyed through the foregrounding of certain details and not others.

V. APPLICATIONS AND OPTIMISM

Although this chapter focuses on the cognitive processes that underly aspects of memory, visuospatial processing, word categorization, and yes, reading, I have not avoided contemporary work on introspective reports. Specifically, the research on flow and the work on mind wandering take introspective reports as a central variable. None of these studies act as a phenomenal judiciary, ruling out one experience in favor of another. We can not say anyone's introspective reports are misguided based on observable behavior alone. However, they can provide corroboration for introspective reports. One promising avenue of research is the search for cognitive interference between introspective reports and reading comprehension or the recollection of textual detail. For example, if performance decreases on a task that recruits the visuospatial sketchpad when a subject is reporting visual imagery, we have evidence that the visual experience has a tangible, measurable effect on the processing of visual information. Similarly, if someone is worse at processing phonological or acoustic information while reporting an inner voice, we have evidence for interference between consciousness and the phonological loop.

I wish to end by highlighting a secondary benefit of this literature. This body of research provides examples that can be used to corroborate introspective reports. They form an extensive list of experimental designs, sources of interference, behavioral measures, and processing tasks tasks. We can cull this list by picking out effects that seem intuitively connected with introspective reports

and later check if they are empirically connected. For instance, from the classic Baddeley (1966) experiment, we know that subjects perform worse on a forced choice recall task when shown a list of words that are phonetically similar than when the words are orthographically similar. People who report an inner voice while reading might have a double interference, both at the level of the phonological loop and at the level of consciousness. We might expect a correlation between reports of an inner voice and poor performance in recollection of phonologically similar words.

Similarly, the work on mind wandering provides a welcome beacon of hope for the corroboration of introspective reports. Given the failure of finding behavioral correlations with experience, as in the case of the VVIQ, a phenomenal pessimism could set in. If we only focus on reports of inner voice and mental imagery, we might neglect the importance of their absence. The impact of mind wandering on text comprehension comes across as a robust correlation, and the effect is strong enough that it could drown out others, such as the role of visual imagery in narrative comprehension. Any attempt to look for a connection between experience and behavior must take into account the unruly nature of our minds.

Chapter 4

EXPERIMENTAL PHENOMENOLOGY

The empirical heart of this thesis is a series of five experiments I conducted under the supervision of Eric Schwitzgebel between 2009 and 2015. These experiments are designed to answer three basic questions. How do people describe their experiences while reading, can we objectively access these experiences through behavior, and what is the connection between these two measures of consciousness? This is the logical extension of the philosophical and psychological issues I have been discussing throughout this dissertation. The first three chapters established the foundation for this set of experiments, and I assume that I have already made the case for the importance of the experience of reading, that I have mapped out the theoretical and conceptual issues, and that I have surveyed the relevant experimental literature.

These experiments draw upon prior research on consciousness using experience sampling, specifically the Experience Sampling Method (ESM) and the Descriptive Experience Sampling (DES) method. In both of these experimental designs, participants engage in their daily routines while carrying a beeper that produces a short tone at random intervals. When the beep occurs, participants use a journal to write down their experiences immediately prior to the beep. In the classic paradigm, the ESM takes these long-form, written introspective reports as

its central data source (Csíkszentmihályi & Csíkszentmihályi, 1988; Hektner, Schmidt, & Csíkszentmihályi, 2007). The DES method uses the reports to structure a subsequent expositional interview, and these interviews constitute the experimental results (Hurlburt & Heavey, 2001; 2004; 2006). These two sources of data — written introspective reports and the transcripts of expositional interviews — have the advantage of being intensely detailed, but at the same time, they are exceptionally hard to quantify (Scollon, Kim-Prieto, & Diener; 2003). While recent research using the ESM focuses on quantitative data (e.g. Carstensen et al., 2011; Csíkszentmihályi & Hunter, 2003; Kane et al., 2007; Killingsworth & Gilbert, 2010), they take introspective reports at face value and do not account for the sources of introspective error outlined in Chapter 2. Research using the DES method has a healthy skepticism towards unreflective reports — one of the main goals of the DES method is to parse through presuppositions to access pristine experience — but the data are not quantifiable (Hurlburt, 2006). For these reasons, the ESM and the DES method have methodological limitations that prevent them from being used to corroborate introspective reports through objective measures of experience.

The series of experiments in this chapter build off these existing methods by including objective measures of experience in addition to introspective reports. I call it the Subjective and Objective Measures of Experience (SOME) method. In SOME experiment, participants hear a beep at a random time while reading a short text passage. Like an ESM or DES experiment, participants then report on

their experiences immediately prior to the beep. In addition, participants perform a variety of psychological tasks that act as objective measures of experience.

Although there are exceptions, the number of participants in a typical ESM or DES study is less than two dozen, while the number of participants in these five experiments ranges from 80 to 1,300.

SOME method contains three main innovations. First, in addition to recording introspective reports on specific moments in time, what I call *concrete introspective reports* (or simply *concrete reports*), it also asks participants to report on their experiences in general, what I call *general introspective reports* (*general reports* for short). This gives us introspective reports that span two scopes, one for the specious present and the other for a wider timespan. We can compare these two types of reports, testing for reliability and internal consistency. Second, these studies investigate the relationship between subjective and objective measures of experience. As I discussed in Chapter 3, past research comes to a fairly pessimistic conclusion about the connection between introspective reports and subsequent behavior, the clearest example of which is the failure to find objective validation of the VVIQ (Schwitzgebel, 2011, Chapter 3). Third, SOME method allows for a quantitative analysis with enough statistical power to detect patterns and relationships that would be overlooked in an ESM or DES study. This is instrumental for achieving the first two aims: checking for reliability and internal consistency in introspective reports and corroborating introspective reports through objective measures of experience.

For SOME method, the subjective measures of experience consist of the two types of introspective reports provided during the course of the experiment: concrete reports and general reports. Formulating objective measures of experience is more complicated. Chalmers (1996; 2004) famously argues that the “hard problem” of consciousness stems from the fact that there is no third-person access to experience. Even Chalmers, however, believes that there are regularities that govern the connection between consciousness and behavior. Unless we are all philosophical zombies or our consciousness is epiphenomenal, we should expect our experiences to manifest an effect on our behavior. Putting aside the various metaphysical issues at play, this is what I am interested in when I talk of objective measures of experience — the observable behavioral correlates of consciousness.

There is a clear intuitive pull to the idea that our experiences have an effect on, even determine, aspects of behavior. For example, consider the experience of choice, the conscious decision to act in one way and not another. It certainly feels like many of our actions are the consequence of decisions that we make and that we could measure our conscious choices by looking at the things that we do (for an argument to the contrary, see Wegner, 2002). There are also theoretical reasons to expect a connection between experience and behavior. This connection is directly built into influential psychological theories of consciousness, such as the global workspace theory. On this view, the primary function of consciousness is to act as a system of widespread information dissemination, broadcasting sensory information to specialized networks and integrating perceptual inputs (Baars,

1983, 2003b; Dehaene, Changeux, Nacchache, Sackur, & Sergent, 2006). In recording objective measures of experience, I am assuming that information in consciousness enjoys a privileged role in cognition. If you experience an inner voice while reading, the auditory information in the text should be more readily available than for someone who experiences inner silence. Thus we should expect reports of an inner voice to be correlated with performance on tasks that access this information, such as the recollection of rhymes. A similar connection should exist for reports of visual imagery and the words on the page.

I. EXPERIMENT ONE

Drawing from existing descriptions of the experience of reading in the philosophical and psychological literature, this study focused on the experience of an inner voice, visual imagery, and the visual experience of the words on the page. This study explored the variability in introspective reports while reading, including the range of experiences reported by readers as well as the way these reports change over time for a single individual. Further, it examined the causes of the variability in the experience of reading. Using self-introspection as a starting point, it seems as if the type of passage influences the experiences that I have. For example, a passage high in descriptive detail might lend itself to an experience of visual imagery, while I am more likely to have the experience of an inner voice while reading a dialogue. This hypothesis gets some tangential empirical support from the research on situational models of narrative comprehension, which posit

that the cognitive mechanism behind reading involves a mental model of the situation described in the text (see Chapter 3 for an extended discussion). Although there is strong evidence that a mental model is employed for many language comprehension tasks, there is no evidence that one is formed when reading non-narratives (Graesser et al., 1994; Zwaan & Radvansky, 1998). Thus the characteristics of the passage have an effect on the cognitive mechanism employed while reading. In addition, this study involved a general search for a connection between experience and a battery of demographic and biographical characteristics, such as English as a second language, age, and hours reading.

A final goal of this study was to examine the connection between subjective and objective measures of experiences. Because information that is available to consciousness enjoys widespread dissemination and a privileged access to cognitive subsystems, it should be possible to measure the behavioral effects of experience. For example, I expected the experience of an inner voice to improve performance on tasks that require access to phonological information from the text, such as the recollection of rhymes. Likewise, the experience of visual imagery should facilitate subsequent recollection of visual detail from the passage. Finally, participants who have a perceptual experience of the words on the page should have a higher rate of recall for the spatial position of key phrases in the text.

This experiment used a within-subjects design. Because we can not monitor participant engagement directly, I used a series of comprehension questions as a proxy and excluded participants with accuracy less than chance.

Methods

Participants

A total of 83 participants from the University of California at Riverside (UC Riverside) received course credit for participation (35 female, 48 male; mean age = 19.1 years). All participants read the same passages and answered the same set of questions.

Materials

Text: This experiment used three reading passages, all about 500 words in length. The first was the opening paragraphs of *White Noise* (DeLillo, 1985), a narrative chosen because of its rich descriptive visual detail. The second was the beginning of Act 1 of *Who's Afraid of Virginia Woolf* (Albee, 1962/1983), chosen because of its colorful dialogue. The third consisted of the first three paragraphs of *Language, Truth, and Logic* (Ayer, 1936/1971), a piece of abstract philosophy that lacks both descriptive detail and dialogue. All passages were slightly modified for length and separated into three or four pages. Each page contained two columns of text (about 75 words per column), with two buttons beneath that allowed the reader to move forward or backward between the pages.

Introspective Reports: Two types of introspective reports were recorded during the course of this experiment (Table 4.1). Participants provided general reports of their

experiences while reading for (1) an inner voice, (2) visual imagery, (3) the words on the page, and (4) an understanding of the text without an accompanying inner voice or visual imagery. (4) was included to raise the possibility that the reader could understand the passage without a modal experience and was left out of the subsequent analysis. Participants responded using a seven-point Likert scale (labeled 1 – “Never,” 4 – “Half of the Time,” 7 – “Always”). Participants provided a set of general reports at the beginning of the experiment, describing their experiences while reading, and then gave another set at the end of the experiment, reporting on their experiences while reading the passages in this experiment.

In his work using the DES method, Hurlburt and Heavey (2006) call most general claims about experience “faux generalities,” for example, “I always experience an inner voice when I read.” As Hurlburt and Heavey argue, real generalities are the result of an inductive process, such as a systematic observation of experiences over time. After such a regimented review of consciousness, you might find an inner voice in all of your observations. However, for most of the general claims people make about experience, they are relying on heuristics or salience, not a systematic inductive process. For this reason, I treated the general reports given at the beginning of the experiment as indicative of previously held beliefs, bias, or an emphasis on salient prototypical experiences rather than as an accurate representation of a participant’s experience.

Table 4.1

The question text for general and concrete introspective reports.

	Report Type	Question Text
Initial General Reports	Inner Voice	How often do you experience an inner voice while reading (i.e. you hear words in your head as if you were reading out loud, or you hear the voice of a character as if they were speaking)?
	Visual Imagery	How often do you experience visual imagery while reading (i.e. you see pictures in your mind, or you see a scene acted out in your mind)?
	Words on the Page	How often do you have a visual experience of the black words on the white background while reading?
	Understanding	How often do you understand what you are reading without also experiencing an inner voice or visual imagery?
Final General Reports	Inner Voice	While reading the three passages, how often did you experience an inner voice (i.e. you heard words in your head as if you were reading out loud, or you heard the voice of a character as if they were speaking)?
	Visual Imagery	While reading the three passages, how often did you experience visual imagery (i.e. you saw pictures in your mind, or you saw a scene acted out in your mind)?
	Words on the Page	While reading the three passages, how often did you have a visual experience of the black words on the white background?
	Understanding	While reading the three passages, how often did you understand what you were reading without also experiencing an inner voice or visual imagery?
Concrete Reports	Inner Voice	In the final split-second before the beep, did you experience an INNER VOICE?
	Visual Imagery	In the final split-second before the beep, did you experience VISUAL IMAGERY?
	Words on the Page	In the final split-second before the beep, did you have a visual experience of the WORDS ON THE SCREEN?
	Understanding	In the final split-second before the beep, did you have an UNDERSTANDING of the words you had JUST READ?

In addition to general reports, participants provided concrete reports on a specific moment in inner time. While reading a passage, participants were probed using a 1 second 500 Hz beep through their computer speakers at a random time

30–90 seconds after the page loaded. Upon completion of the beep, a new page automatically loaded and prompted participants to report on their experiences “in the final split-second before the beep” (e.g. “In the final split-second before the beep, did you experience an inner voice?”). Participants provided concrete reports using multiple choice responses (“Yes,” “No,” “Maybe / Don’t Know”) of (1) an inner voice, (2) visual imagery, (3) the words on the page, and (4) understanding the text without accompanying inner voice or visual imagery. As before, (4) was excluded from the subsequent analysis. Afterwards, participants described their experience in a free form text box.

Objective Measures of Experience: Participants answered 27 questions during the experiment, nine questions after each passage. These were chosen from a series of trial questions from a pilot study involving thirty participants. Questions were separated into groups based on the types of experience they were designed to measure. The Visual Perception Group measured the visual experience of the words on the page, and consisted of nine questions that asked participants to recall the location of a key phrase on the page. Participants chose between the upper left, lower left, upper right, and lower right quadrants using a small image of a blank page with the four quadrants labeled. The Visual Detail Group consisted of three questions that asked participants to recall visual details such as the color of an article of clothing or the type of items covering the ground. The Phonological Group consisted of six questions. Two asked participants to recall the spoken

volume of lines of dialogue, choosing the line that was said most loudly (or softly). The answer was determined by textual cues, in one case the use of all capital letters, in another the phrase was preceded by “(wearily).” Four questions tested for a phonological or orthographic mechanism behind the subsequent identification of names. Two of the passages contained a novel name (e.g. “wilfose” in a list of candies). Participants were asked to identify this novel name from a series of answer choices that were either phonologically similar (e.g. “Willphoes”) or orthographically similar (e.g. “Wilfase”). For the purposes of measuring the experience of an inner voice, the phonologically similar answer was correct. Finally, the Comprehension Group consisted of nine questions that tested basic text comprehension. For all questions, answers were selected from four choices presented in random order.

Procedure

The experiment was hosted on the Lime Survey platform, and participants took part online using Internet Explorer on their own computers. After a test to check for functioning sound and instructing participants to set the volume appropriately (“the sound should be distinct but not startling”), participants provided general reports on their experience while reading. Afterward, they read three passages presented at random. While reading each passage, participants heard a beep which prompted a new page to load and instructed the participants to provide concrete reports. Participants then returned to the passage and continued reading. After

each passage, participants answered objective measures questions. The experiment concluded with each participant providing a final set of general introspective reports. The entire study took about 15 minutes.

Results and Discussion

Responses from three of the comprehension questions, one for each passage, were excluded from the analysis. On a subsequent reading of the questions, I decided post hoc that the correct answer was unclear. Once these questions were excluded, data from 26 participants were excluded because their performance on the Comprehension Group was less than chance, that is, they answered less than two out of six questions correctly. Because of this, the analysis includes data from 57 participants. If these three comprehension questions were to be included, the exclusion criteria for participants would remain the same — participants must performed greater than chance on the comprehension questions — but participants would be excluded if they answered less than three out of nine comprehension questions correctly. Although the reported results to not include these three ambiguous questions, I performed the analysis on both sets of data and note when they diverge.

Introspective Reports: The first set of analyses looks at introspective reports, beginning with the final set of general reports. The experiences reported during this experiment spanned the extremes. While some participants provided general

reports of a constant inner voice, others reported never experiencing one at all. This broad range is also found in general reports of visual detail and reports of the words on the page (Figure 4.1). For all three types of reports, the mean was ± 1 from the median of 4 on the Likert scale, explicitly labeled “Half of the Time.” The experience of an inner voice ($M = 4.54$, $SD = 1.90$) was reported more often than visual imagery ($M = 4.35$, $SD = 1.63$) or the words on the page ($M = 3.02$, $SD = 1.73$). There was a similar distribution in concrete reports (Figure 4.2). Participants reported an inner voice in 53% of reports, visual imagery in 46% of reports, and words on the page in 31% of reports. In both general and concrete reports, participants expressed the same ranking, and an inner voice was the most common report while the perception of the words on the page the least prevalent.

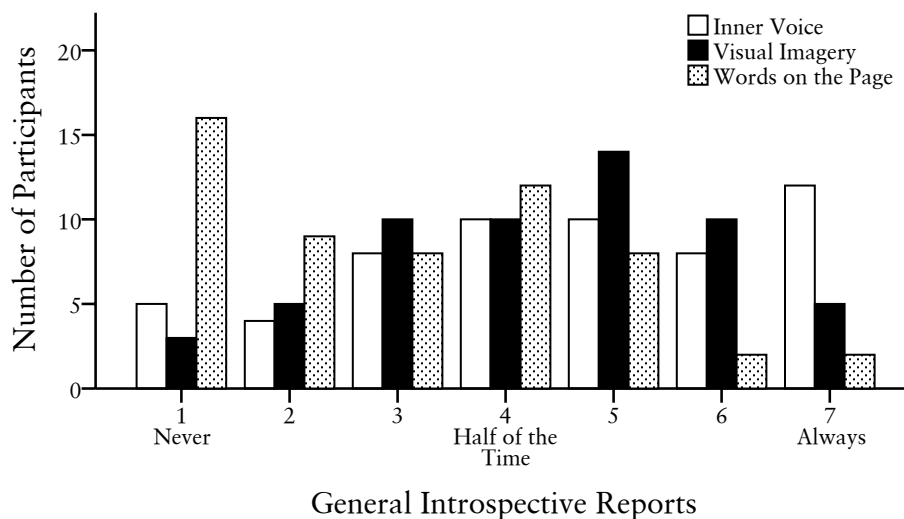


Figure 4.1
Histogram of general reports for an inner voice, visual imagery, and the words on the page ($N = 57$).

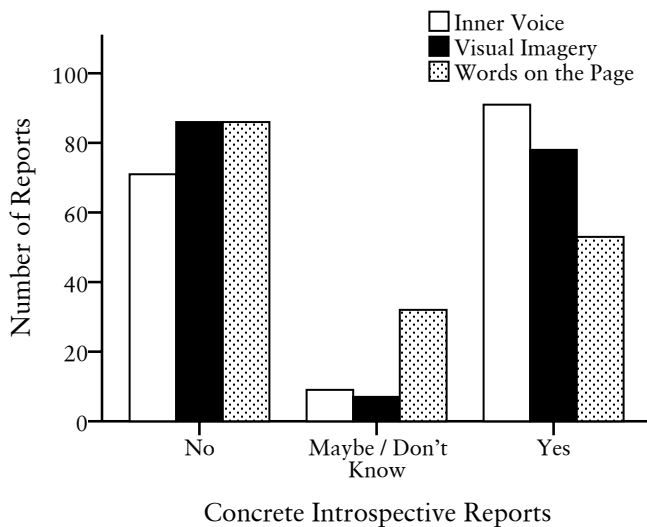


Figure 4.2

The sum of the three concrete introspective reports for an inner voice, visual imagery, and the words on the page (N = 171).

Participants provided a written description of their experiences in 88% of the concrete reports. The diverse array of experiences described by participants mirrors the descriptions provided by writers in the psychological and philosophical literature. One participant described their experience of an inner voice, saying “it was as if there was a narrator speaking to me of what the words on the page were,” while another wrote “the inner voices i heard were reflecting of the characters, one was male and one was female.” One participant gave a detailed description of their visual experience before the beep and wrote “I was visualizing a church, one of those large, gothic styled churches, with a black roof, and grey walls. There was a yellow stained glass clock on the front, with a bell at the top. The church was behind a black gate, with a path leading towards it, winding through well maintained grass.” Descriptions of the words on the page were less

common and typically also involved a modal experience, such as “this passage was the most difficult to visualize, but certain words like ‘philosophy’ and ‘metaphysics’ served almost as picture-words on the page to make the passage as interesting as if visualization was easy.” Finally, there were some participants who reported no experience at all or experiences that were entirely unrelated to the passage they were reading, such as “I experienced a scene from the movie, *The Break Up* with Jennifer Aniston and Vince Vaughn. I remember the scene when Richard is talking about the feeling he gets when he sings. He describes it as something ‘transcending.’”

An analysis of introspective reports at the extremes illustrates the large interpersonal variability in responses. For example, through their general reports, five participants (9%) claim they never experience an inner voice, while twelve (21%) say that they always do. There is also a large variability in concrete reports. Eleven participants (19%) report no experience of an inner voice during the three probes, while 19 (33%) report an inner voice in all of them. This wide range is also found in general and concrete reports of visual imagery and words on the page (Table 4.2).

Table 4.2

The number of participants who reported experiences at the extremes. For general introspective reports, these are the participants who responded that they always, or never, have a certain experience. For concrete reports, these are the participants who reported the same experience in all three probes (N = 57).

	Inner Voice		Visual Imagery		Words on the Page	
	General	Concrete	General	Concrete	General	Concrete
Never Experience	5 (9%)	11 (19%)	3 (5%)	9 (16%)	16 (28%)	12 (21%)
Always Experience	12 (21%)	19 (33%)	5 (9%)	9 (16%)	2 (4%)	6 (11%)

As stated earlier, there are reasons to believe that the initial set of general reports are faux-generalizations and should be treated with skepticism, but let us take a closer look (Figure 4.3). For all three types of experiences, participants reported a higher level of experience in the initial general reports than in the final general reports, although the only significant difference is between initial reports of visual imagery ($M = 5.26$, $SD = 1.29$) and final reports of visual imagery ($M = 4.35$, $SD = 1.63$), $t(56) = 3.99$, $p < .000$. Interestingly, not a single participant reported low levels of visual imagery, a Likert report of 1 or 2, in the initial set of reports, although 8 (14%) did so in the final set.

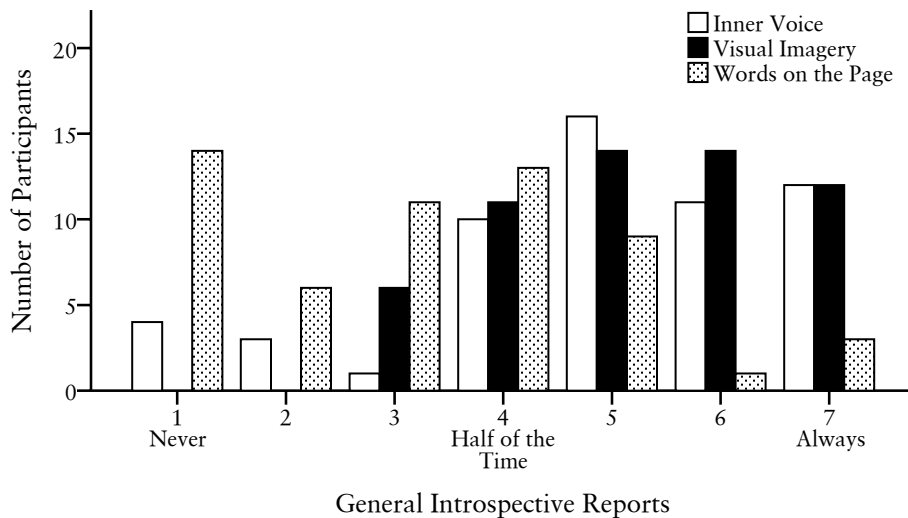


Figure 4.3

Histogram of the initial set of general reports for the experience of an inner voice, visual imagery, and the words on the page (N = 57).

In addition to these interpersonal differences in introspective reports, there was also a high degree of variability in reports for a single reader over time. As already stated, the mean Likert report for the experience of an inner voice, visual imagery, and the words on the page were all ± 1 from the median of 4, implying that the experience while reading is frequently morphing, for example, an inner voice at one time but not another. In fact, a Likert report other than 1 or 7 implies an experience that has some degree of variability. 40 participants (70%) reported this variability in their experience of an inner voice, 49 (86%) for visual imagery, and 39 (68%) for the perceptual experience of the words on the page. The frequency of concrete reports is consistent with these results. In the three concrete reports given throughout the experiment, 27 (47%) participants reported an inner voice at one time and not another, while 39 (68%) reported as much for visual

imagery, and similarly, 39 (68%) did the same for words on the page. These concrete reports alone imply that 47–68% of the participants had an experience that changed during the course of this short experiment.

To examine the way that the experience of reading changed over time, I performed a two-tailed Pearson correlation between the two sets of general reports and the aggregate concrete introspective reports (Table 4.3). First, the correlation between the general introspective reports given at the beginning and end of the experiment is relatively large ($r = .32-.55$). One way to read these results is as a measure of test-retest reliability, asking the same question at the beginning and end of the experiment. If this is the case, the correlation is considerably smaller than we should expect. This interpretation is called into question by the wording of the questions themselves. The initial set of general reports asked about the regular experience of reading, while the final set of reports asked about the experience of reading during this experiment. Perhaps the different reports are due to participants answering two distinct questions. Given the context, I find it unlikely that participants applied two distinct interpretations to these questions. Another interpretation is that participants treated the final set of reports as a compromise between their initial faux-generalizations and their experiences during the experiment. A third interpretation is that the participants are moving beyond their prototypical conception of reading. It is possible that participants had a specific type of reading passage in mind, such as a descriptive narrative, when providing the initial set of reports and were not thinking about their experiences while

reading abstract philosophy. At the end of the experiment, they had a broader range of passage types in mind, thus their reports were more representative of their experiences as a whole. I will return to these competing interpretations in the general discussion in Chapter 5.

Table 4.3

Table of two-tailed Pearson correlations between the three sets of introspective reports (N = 57).

	Initial and Final General Reports	Initial General Reports and Concrete Reports	Final General Report and Concrete Reports
Inner Voice	.43 (p = .001)	.30 (p = .024)	.71 (p < .000)
Visual Imagery	.32 (p = .015)	.18 (p = .172)	.68 (p < .000)
Words on the Page	.55 (p < .000)	.32 (p = .015)	.78 (p < .000)

The correlation between the initial set of general reports and the concrete reports is small ($r = .18-.32$), especially given that general reports are supposed to be constituted by concrete reports. Although the correlation between the final set of general reports and the aggregated concrete reports is considerably higher ($r = .68-.78$), it is still low given the variables we are measuring. Once again, general reports are constituted by moments in time, precisely the thing that concrete reports are supposed to be measuring. This could be due to faux-generalizations, overgeneralization in general reports, or the small sample size of concrete reports.

Next I looked at the effect of passage characteristics on concrete reports (Figure 4.4). Recall that participants read a descriptive passage, a dialogue, and an

abstract passage, providing a concrete report for each. I performed three repeated measures ANOVAs to compare the effect of the passages on the three types of concrete reports. For reports of an inner voice, the passage had a significant effect, ($F(2, 112) = 5.5, p = .005$), and a post hoc test using the Bonferroni correction revealed that there was a higher level of inner voice reported in the dialogue ($M = .33, SD = .93$) than the abstract passage ($M = -.11, SD = .99$). The descriptive passage ($M = .12, SD = .95$) did not significantly differ from the other two. For reports of a visual imagery, the passage again had a significant effect ($F(2, 112) = 19.2, p < .000$), and a post hoc test using the Bonferroni correction showed a significantly lower level of visual imagery reported in the abstract passage ($M = -.58, SD = .80$) than in the descriptive passage ($M = .28, SD = .96$) or the dialogue ($M = .16, SD = .99$), while the the descriptive passage and the dialogue did not differ. Finally, the passage had a significant effect for reports of words on the page, ($F(2, 112) = 4.85, p = .010$). Another post hoc tests using the Bonferroni correction showed a higher level of reports of words on the page in the dialogue ($M = .00, SD = .93$) when compared to the descriptive passage $M = -.16, SD = .84$) or the abstract passage ($M = -.42, SD = .84$). However, the descriptive passage and the abstract passage did not significantly differ from each other. (The effect of passage type on reports of words on the page disappears when including the three comprehension questions that were excluded post hoc).

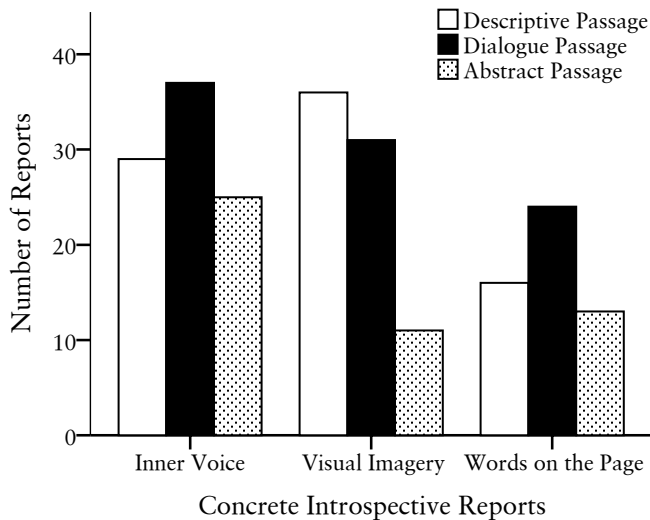


Figure 4.4
Histogram of positive concrete reports for the descriptive passage, the dialogue, and the abstract passage.

This suggests that the characteristics of the text affect the way people experience reading it. Specifically, dialogue has a significant positive effect on reports of an inner voice, while abstract philosophy has a negative effect on the experience of visual imagery or the words on the page. This is consistent with exiting literature on the situational model of text comprehension, which finds that different cognitive mechanisms are employed while reading narratives when compared to other types of text (Graesser et al., 1994; Zwaan & Radvansky, 1998). One possible explanation is that participants are spacing out or their minds are wandering while reading the abstract philosophy, which could explain the overall lower levels of experience reported for that passage. I look further into the effects of mind wandering in Experiments 3–5.

An independent-samples t-test for equality of means found no difference between native english speakers and non-native english speakers for general or concrete introspective reports of an inner voice, visual imagery, or the words on the page. (If I include the three comprehension questions excluded post hoc, there is a significant difference in concrete reports of an inner voice between native ($M = .77$, $SD = 2.32$) and non-native ($M = .36$, $SD = 1.98$) english speakers, ($t(62) = 2.01$, $p = 0.049$), as well as concrete reports of visual detail between native ($M = .23$, $SD = 1.86$) and non-native ($M = -.72$, $SD = 1.82$) english speakers, ($t(38) = 2.02$, $p = 0.048$). An independent-samples t-tests for equality of means found that gender did not have a significant effect on introspective reports. A two-tailed Pearson correlation found that age and weekly hours reading did not significantly correlate with general or concrete reports.

Objective Measures of Experience: All objective measures of experience were multiple choice questions with four answer options, so chance performance was 25%. Accuracy was significantly above chance on all question groups and subgroups except for the spoken volume subgroup (Table 4.8).

Table 4.4
Mean accuracy for the objective measure of experience.

Question Group	Question Subgroup
Phonological (33%)	Name identification (42%) Spoken volume (15%)
Visual Detail (41%)	—
Visual Perception (33%)	—
Comprehension (49%)	—

I performed a series of two-tailed Pearson correlations to examine the covariance between subjective and objective measures of experience. Unexpectedly, there was no significant correlation between any of the six hypothesized relationships (Table 4.5). This either suggests a skeptical view toward the reliability of introspective reports or a skeptical view towards the causal efficacy of our experiences. Of course, it is always possible that this study employed poor objective measures of experience. Although there are theoretical reasons to expect a connection between experience and performance on these questions, designing objective measures of experience was considerably more difficult than expected. When discussing “triangulating on experience” in Chapter 2, I was talking about honing in on consciousness using introspective, behavioral, and neurological measures. However, we may need to do a sizable amount of refining on the behavioral measures of experience alone.

Table 4.5

Table of two-tailed Pearson correlations between subjective and objective measures of experience. None of the results were significant ($p < .05$).

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Phonological Group	Inner Voice Reports	General Concrete	.171 ($p = .205$) .101 ($p = .453$)
Visual Group	Visual Imagery Reports	General Concrete	.154 ($p = .254$) .041 ($p = .764$)
Perceptual Group	Words on the Page Reports	General Concrete	-.067 ($p = .618$) -.070 ($p = .606$)

There was a significant negative correlation between performance on the comprehension group and reports of words on the page, both general ($r = -.328$, $p = .013$) and concrete ($r = -.261$, $p = .050$). (With the three comprehension questions that were excluded post hoc, the correlation between comprehension questions and concrete reports of words on the page is not significant). It is plausible that the experience of the words on the page is detrimental to reading. This gets minor experimental support from Singer and Halldorson (1996, Experiment 4) who found that instructing participants to proofread reduced their ability to make spatial textual inferences. Perhaps the perceptual experience of the words on the page interferes with, or is caused by, a failure in the cognitive mechanisms behind reading comprehension.

II. EXPERIMENT TWO

The present study further explores the relationship between subjective and objective measures of experience. To detect smaller effects, the number of participants was increased, and I improved upon the objective measures of experience. This experiment used a within-subjects design. As in Experiment 1, participants with below chance performance on the comprehension questions were excluded.

Methods

Participants

A total of 148 participants from UC Riverside received course credit for participation (95 female, 53 male; mean age = 19.7 years). All participants read the same passages and answered the same set of questions.

Material

Text: This experiment used two poems and two short narratives. The poems were chosen for their fixed rhyming patterns and were 100–150 words long. The first was a combination of two poems by E. E. Cummings (1994), beginning with “The Noster Was a Ship of Swank” and ending with a modified version of “Here is Effie’s Head.” The second poem was the original text of “A Dream Within a Dream” by Edgar Allen Poe (1849/2008). The first narrative was a shortened

version of *The Disk* by Jorge Luis Borges (1975/1999), chosen for its unique voice and mixture of descriptive visual detail and dialogue. The second narrative was an account of being shrunken down to investigate a clover, a story rich in visual detail. The full text of the poems were displayed on a single screen, while the narratives were separated into three pages as in Experiment 1.

Introspective Reports: This experiment recorded general and concrete reports using the same measures as Experiment 1.

Objective Measures of Experience: Participants were asked five to eight questions after reading each of the four passages. Six questions formed the Phonological Group. Two tested the pronunciation of novel names with ambiguous phonemes (e.g. “Pater”), with the answer determined by the rhyming scheme of the poems (e.g. “That dread day a craft of cost / The whole ship perished later / All hands (you may recall) being lost / Including captain Pater”). One had participants identify a line written in the same cadence as the poem, in this case, iambic feet. Two asked participants to recall rhymes from the poem. One asked participants to recall the tone of a line of dialogue. Six questions formed the Visual Detail Group, and all asked participants to recall visual detail from the passages. Seven questions formed the Visual Perception Group. Three asked participants to identify the location of phrases on the page or in a verse. One had participants identify the font that the passage was written in. One asked participants to identify punctuation

used in the poem. One asked the number of times a line was repeated in the poem. One asked participants to identify the first line of the poem. Finally, five questions formed the Comprehension Group and tested basic reading comprehension.

Procedure

This experiment used a similar procedure as Experiment 1. The experiment was hosted on the Lime Survey platform, and participants took part online using Internet Explorer on their own computers. After a test to check for functioning sound and instructing participants to set the volume appropriately (“the sound should be distinct but not startling”), participants provided general reports on their experience while reading. Participants then read all four passages randomly displayed in the order of poem–narrative–poem–narrative. While reading the two narratives, participants heard a beep which prompted a new page to load, prompting them to provide concrete reports. Participants then returned to the passage and continued reading. After each of the four passages, participants answered a series of questions designed to be objective measures of experience. The experiment concluded with each participant providing another set of general reports. The entire study took about 15 minutes.

Results and Discussion

8 participants were excluded for performing less than chance on the Comprehension Group, that is, they answered less than two questions correctly.

After this exclusion, 140 participants were included in the analysis. Two questions from the Visual Detail Group and two questions from the Phonological Group were excluded post hoc because the answers were ambiguous. Although the reported results do not include these questions, I performed the analysis on both sets of data and note when they diverge.

Introspective Reports: An independent-samples t-test was used to compare introspective reports for Experiment 1 and Experiment 2. There was no significant difference in means between any of the three general reports given at the beginning of the two experiments, confirming that participants in the two experiments had the same initial views about their experience while reading.

I conducted a chi-squared test to examine the effect of the passage on concrete introspective reports in this experiment. Reports of visual imagery were significantly higher for The Disk passage ($M = .66$, $SD = .74$) than the Clover passage ($M = .37$, $SD = .88$); $\chi^2(1) = 6.85$, $p = .009$). There were no other significant differences between passage and reports. An independent samples t-test for equality of means found no significant difference between reports of male and female participants. An independent samples t-test found that general reports of an inner voice are higher for native english speakers ($M = 5.49$, $SD = 1.61$) than for non-native english speakers ($M = 4.79$, $SD = 1.79$), $t(138) = 2.45$, $p = .015$).

Objective Measures of Experience: Except for the punctuation subgroup, mean accuracy for all question groups and subgroups was significantly above chance ($p < .05$) (Table 4.6).

Table 4.6
Mean accuracy for the objective measure of experience.

Question Group	Question Subgroup
Phonological (43%)	Pronunciation (38%) Cadence (38%) Rhyme (58%) Tone (excluded)
Visual Detail (62%)	—
Visual Perception (45%)	Location of phrases (44%) Font (77%) Punctuation (30%) Repeated lines (69%) First line (85%)
Comprehension (51%)	—

To examine the relation between subjective and objective measure of experience, I performed a series of two-tailed Pearson correlations (Table 4.7). There was a significant correlation between performance on the Phonological Group and reports of an inner voice, both general and concrete. Further, as in Experiment 1, there was a negative correlation between general reports of words on the page and performance on the Comprehension Group ($r = -.231, p = .006$).

Table 4.7

Table of Pearson correlations between introspective reports and performance on objective measures of experience.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Phonological Group	Inner Voice Reports	General	.169 (p = .045)
		Concrete	.197 (p = .019)
Visual Detail Group	Visual Imagery Reports	General	.102 (p = .231)
		Concrete	.106 (p = .212)
Visual Perception Group	Words on the Page Reports	General	-.034 (p = .138)
		Concrete	-.009 (p = .918)

There was a significant positive correlation between performance on the Comprehension Group and reports of an inner voice, both general ($r = .212$, $p = .012$) and concrete ($r = .239$, $p = .004$). Because of the shared variance between introspective reports and reading comprehension, it is plausible that performance on the objective measures of experience is simply the result of improved reading comprehension. To test for this, I performed three sets of stepwise regressions (Table 4.8). It is possible that any connection between subjective and objective measures of experience would depend on reading comprehension or participant engagement. For example, for a reader who does not understand the text, an inner voice might bear no relation to performance on the Phonological Group. However, as engagement with the text increases, so might the strength of a relationship between these two measures of experience. Because of this, the regression models include interaction variables between introspective reports and performance on the Comprehension Group. For these regressions, the general introspective reports are

centered on a Likert response of “0.” Positive numbers imply an experience more than half the time and negative numbers less than half the time. Concrete reports are also centered on “0”. Positive numbers indicate an affirmative report of experience, and negative numbers indicate a negative report of experience.

Table 4.8

Stepwise regression models for predicting objective measures of experience using the corresponding introspective reports and reading comprehension.

Dependent Variable	Independent Variable	Model	Predictor
Phonological Group	(1) Comprehension questions (2) General reports of inner voice (3) Concrete reports of inner voice (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .039$ $F(1, 138) = 5.59$ $p = .019$	IV (3) $b = .125$, $SE = .053$ $p = .019$
Visual Detail Group	(1) Comprehension questions (2) General reports of visual imagery (3) Concrete reports of visual imagery (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .053$ $F(1, 138) = 7.70$ $p = .006$	IV (1) $b = .214$, $SE = .077$ $p = .006$
Visual Perception Group	(1) Comprehension questions (2) General reports of words on the page (3) Concrete reports of words on the page (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	Not significant	

The first model shows that performance on the Phonological Group is predicted by the concrete reports of an inner voice, confirming that the correlation between the Phonological Group and concrete reports of an inner voice is not merely the result of reading comprehension or participant engagement. Given the lack of a significant correlation between reports of visual imagery or the words on

the page and their objective counterparts, it is not surprising that introspective reports are not predictive on these tasks. (If I include the questions that were removed post hoc, the first model remains predictive of performance on the Phonological Group [$R^2 = .055$, $F(1, 138) = 8.03$, $p = .005$], however the only significant predictor is IV (5), the interaction variable between concrete reports of an inner voice and the Comprehension Group ($b = .043$, $SE = .015$, $p = .005$).

III. EXPERIMENT THREE

The next study built upon the results from Experiments 1 and 2 through refinements. Although everyone is familiar with the experience of mind wandering while reading, by and large this everyday experience is absent from the descriptions of experience in the philosophical and psychological literature. There is, however, a small body of quantitative research on mind wandering. Unlike the experience of an inner voice or visual imagery, research on mind wandering has been successful in finding behavioral correlations between subjective and objective measures of experience. Studies have repeatedly found that reports of mind wandering are negatively correlated with reading comprehension (Schooler et al., 2004; Smallwood, Fishman, & Schooler, 2007; Smallwood, McSpadden, & Schooler, 2008). This relationship is robust across introspective reports that are self-caught and those that are caught by introspective probes, a distinction that maps onto the general and concrete reports used in this set of studies. If the effect is large enough, mind wandering could be the dominant factor in explaining performance on the

objective measures of experiences. Perhaps, for all the experiences that people report while reading, the most important factor is not the modality of the experience, but the content. To test for this, the present study asked participants to report on their experience of mind wandering.

This study also explored the effects of directed attention. People often report having a degree of control over their experiences. For example, merely thinking about whether or not you have a tactile experience of your foot in your shoe can cause you to have that very experience (Schwitzgebel, 2011, ch. 6). Casual introspection of my own experience suggested that directed attention can affect the experiences that I have while reading. This study investigated the effects of priming and directed attention by instructing participants to focus on a single type of experience, in this case, the experience of an inner voice, visual imagery, or the words on the page. Since it is confusing to ask participants to focus on the experience of mind wandering, a control group was informed about the possibility of mind wandering and instructed to stay focused on the text.

Third, I explored the effects of probing for concrete reports on the experience of reading. In Experiments 1 and 2, some participants reported being anxious about the beep, expecting the beep, or awaiting the beep. As I argue in Chapter 1, there are numerous ways that we read, all of which are equally valid contexts for studying the experience of reading. Notably absent from that list, however, is a common reading task found in the experimental literature on mind wandering. Participants in these studies often have a single word presented at a

time when they read and must manually advance the text with a button press (Smallwood et al., 2007; Smallwood et al., 2008). It is unclear how this affects our experiences while reading. Imagine pressing a button to view the next word of this thesis, sentence after sentence, page after page. It is a radical departure from the rapid succession of words that is normal while reading, where our eyes saccade back and forth across the text, often skipping entire words (see Chapter 3 for an extended discussion). Because of this, I did not probe participants for concrete reports in this experiment, and used general reports as the sole subjective measure of experience.

In addition, I added two new objective measures of experience, a Stroop Task and a memory task. Over decades of research, the Stroop Task has demonstrated a robust interference between irrelevant information —the meaning of the word — and relevant information — the color of the word (Stroop, 1935). I expected reports of visual experience to negatively correlate with performance on the Stroop Task as the experience itself interfered with categorization. A memory task was chosen because of a general interest in the effect of short term memory on the accuracy of introspection. This experiment used a within-subjects design, and participants were excluded if they answered the single comprehension question incorrectly.

Methods

Participants

A total of 243 participants from the United States were paid to take part in the experiment through Amazon MTurk (148 female, 95 male; mean age = 38.6 years). Participants were randomly assigned to one of four conditions: 57 were in the inner voice condition, 59 were in the visual imagery condition, 51 were in the words on the page condition, and 76 were in the control condition. All participants read the same passage and answered the same set of questions.

Material

Text: Subjects read a modified version of “The Egg And The Machine” by Robert Frost (1927/1969), about 250 words long. The poem was separated into five stanzas. Two words were repeated twice in a row (“had had” and “was was”) for use in a word recognition task.

Introspective Reports: As in Experiment 1, participants provided general reports during the experiment (Table 4.9). This experiment added a report for mind wandering and removed the report for understanding without an accompanying auditory or visual experience. Thus, participants reported on their experiences of (1) an inner voice, (2) visual imagery, (3) the words on the page, and (4) mind wandering. Participants responded using a seven-point Likert scale. This

experiment did not ask participants to provide concrete introspective reports. In the previous experiments, participants were asked how often they experience the words on the page while reading. However, the present study changes the question, and asks how often they do not experience the words on the page. Because there is a sense in which we are always aware of the text while reading, this change makes the question easier to interpret. Reports of words on the page were inverted in the subsequent analysis, so like the other reports, a higher number indicates a more frequent experience.

Table 4.9

The question text for the two sets of general introspective reports.

	Report Type	Question Text
Initial General Reports	Inner Voice	How often do you experience an inner voice when you read? Examples: you hear a voice reading in your head, you hear the characters speaking in your head.
	Visual Imagery	How often do you experience visual imagery when you read? Examples: you see pictures from the story in your mind, you see a scene acted out like a movie in your head.
	Words on the Page	How often do you NOT experience the words on the page when you read Examples: you're so absorbed in a story that it almost seems like you're there, your mind is filled with the ideas in the story and not the actual black letters against the white background.
	Mind Wandering	How often do you find your mind wandering when you read? Examples: you are thinking entirely unrelated thoughts, you are spacing out and can't focus on the text.
Final General Reports	Inner Voice	While reading the poem, how often did you experience an inner voice? Examples: you heard a voice reading in your head, you heard the characters speaking in your mind.
	Visual Imagery	While reading the poem, how often did you experience visual imagery? Examples: you saw pictures from the story in your mind, you saw a scene acted out like a movie in your head.
	Words on the Page	While reading the poem, how often did you NOT experience the words on the screen? Examples: you were so absorbed in the story that it almost seems like you were there, your mind was filled with the ideas in the story and not the actual black letters against the white background.
	Mind Wandering	While reading this poem, how often did you find your mind wandering? Examples: you were thinking entirely unrelated thoughts, you were spacing out and couldn't focus on the text.

Objective Measures of Experience: After reading the poem, participants answered six questions presented at random. One was a multiple choice reading comprehension question with four answer options. One formed the Phonological Group and asked participants to recall ten rhyming pairs from the poem out of a list of twenty rhyming pairs. Three questions formed the Visual Perception Group.

One asked participants to recall the font used in the poem from a list of four possible fonts. One asked participants to recall the location on the line (beginning or end) of 10 two-word phrases. One question asked participants to recall the two words that were repeated twice in a row (“was was” and “had had”). Participants also performed a Stroop Task with 16 trials. The first four trials were excluded as training. Finally, participants performed a memory task. They were shown a series of 15 random 3-digit numbers for three seconds each and then given a forced choice recall task with 30 options.

Procedure

As in Experiment 1, participants took part in the experiment on their own computers using the Lime Survey platform. The list of browsers was expanded to include Firefox, Chrome, and Safari. Participants provided demographic information and then gave general reports on their experience while reading. Participants were then separated into four conditions, each instructing participants to focus on the experience of either (1) an inner voice, (2) visual imagery, or (3) the words on the page. A control group was instructed (4) not to let their minds wander. After reading the passage, participants provided another set of general reports and then answered a series of questions that act as objective measures of experience. The experiment took about 15 minutes to complete.

Results and Discussion

15 participants answered the comprehension question incorrectly and were excluded. After this exclusion, 228 participants were included in this analysis. A post hoc analysis of the Stroop Task results indicated that some of the participants did not understand the task. Because of this, an additional 18 participants were excluded from the analysis of the Stroop Task because they answered less than six of the sixteen questions correctly.

While Experiments 1 and 2 recruited participants through psychology courses at UC Riverside, this experiment recruited participants through Amazon MTurk. There are some clear differences between these two samples, the most striking of which is age (Experiment 1: $M = 19.1$, Experiment 3: $M = 38.6$). However, recent studies that compare the results given by university and MTurk participants have found no significant differences between these two populations (Horton, Rand, & Zeckhauser, 2011; Mason & Suri, 2011). I assume that these two populations are identical in this analysis.

Introspective Reports: In the final set of general reports, reports of mind wandering had a positive skew ($M = 2.89$, $SD = 1.62$) (Figure 4.5). Mind wandering was not uncommon yet was considerably less prevalent than the experience of inner voice ($M = 4.25$, $SD = 1.97$), visual imagery ($M = 4.77$, $SD = 1.55$), or the experience of the words on the page ($M = 3.36$, $SD = 1.55$). It should be noted, however, that

Schooler, Reichle, and Halpern (2004) found that people underreport mind wandering, so the actual amount could be significantly higher.

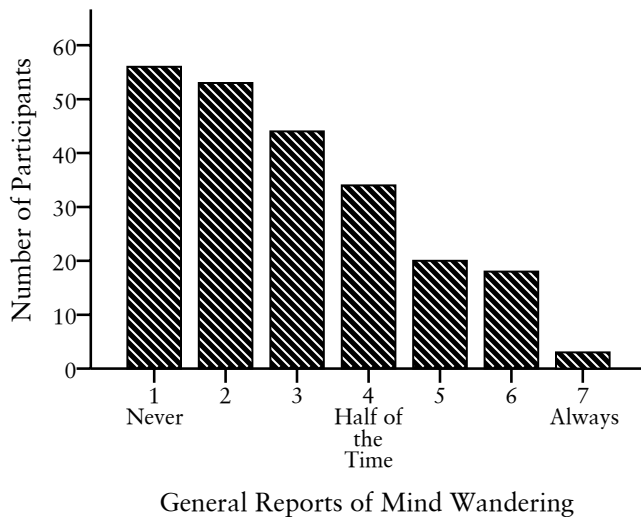


Figure 4.5
Histogram of general reports of mind wandering (N = 228).

I performed a series of independent samples t-tests to analyze the effect of directed attention on introspective reports (Table 4.10). While the mean report of an inner voice and visual imagery was higher when instructed to attend to these experiences, the difference was not significant. A power analysis shows that the 95% confidence interval on the difference is rather large, and at its most precise can only detect a difference of 1.15 on a Likert scale from 1 to 7. Because of this power limitation, these results are inconclusive.

Table 4.10

Table of independent samples t-tests on the difference in mean reports for the directed attention conditions and the control group.

Directed Attention Condition	Dependent Group Mean	Control Group Mean	Sig.	Mean Difference	95% C.I. Lower	95% C.I. Upper
Inner Voice	4.64 (N = 55)	4.14 (N = 71)	0.19	0.50	-0.25	1.24
Visual Detail	4.23 (N = 56)	4.14 (N = 71)	0.49	0.20	-0.37	0.78
Words on the Page	3.98 (N = 46)	4.14 (N = 71)	0.22	0.16	-0.96	0.22

Objective Measure of Experience: Mean accuracy for all questions was significantly above chance except for the location of phrase subgroup ($p < .05$) (Table 4.11). Performance on the visual Stroop Task exhibited a ceiling, with accuracy at 98%.

Table 4.11

Mean accuracy for the objective measure of experience.

Question Group	Question Subgroup
Phonological (68%)	Pronunciation (38%) Cadence (38%) Rhyme (57%)
Stoop Task (98%)	—
Visual Perception (58%)	Identification of font (57%) Location of phrase (29%) Repetition of words (61%)
Comprehension (94%)	—
Memory (72%)	—

A two-tailed Pearson correlation was used to assess the relationship between introspective reports and objective measures of experience (Table 4.12). Although reading was not artificially interrupted by a prompt to provide a concrete report of experience, the connection between objective and subjective measures of experience remained weak. There was a negative correlation between reports of visual imagery and performance on the Stroop Task. This fits the hypothesized relationship that increased visual imagery causes visual interference between the meaning of the color word and the color of the word itself. There was also a small negative correlation between reports of mind wandering and performance on the memory task ($n = 228$, $r = -.146$, $p = .027$). Reports of mind wandering could track participant engagement in the memory task. There was no significant relationship between reports of an inner voice and performance on the Phonological Group, nor between reports of words on the page and performance on the Visual Perception Group. Participants who answered the single comprehension question incorrectly were excluded, so unlike the previous experiments, was no reading comprehension results to analyze.

Table 4.12

Table of Pearson correlations between introspective reports and performance on objective measures of experience. The negative correlation between reports of visual imagery and the Stroop Task was expected.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Phonological Group	Inner Voice Reports	General	$r = .106$ ($p = .109$)
Stroop Task	Visual Imagery Reports	General	$r = -.142$ ($p = .040$)
Visual Perception Group	Words on the Page Reports	General	$r = -.076$ ($p = .251$)

IV. EXPERIMENT FOUR

This study has a higher statistical power by drastically increasing the number of participants. Further, this study increased the number of passages from 4 to 24 to assess the robustness of the procedure to minor variations. As in the previous three experiments, this experiment used a within-subjects design, and participants who performed worse than chance on the comprehension questions were excluded. In addition, there were two new exclusions. In four of the concrete reports given in Experiment 2, participants commented that they had already finished reading the passage when the beep occurred. Because of this, I excluded participants who explicitly stated that they had finished reading at the time of the beep. Further, I excluded participants who spent less than ten seconds reading the passage after providing concrete reports. This is an arbitrary cutoff point to ensure that participants finished reading the passage.

Methods

Participants

1,380 participants were recruited for this experiment (760 female, 620 male; mean age = 26.1 years). 787 received course credit at UC Riverside, and 593 received a payment of \$0.80 through Amazon MTurk. All MTurk participants were from the United States. Participants were randomly sorted into 24 conditions, each with its own reading passage. Conditions 1–8 explored the experience of an inner voice, conditions 9–16 explored the experience of visual imagery, and conditions 17–24 explored the experience of words on the page.

Materials

Text: This experiment used 12 reading passages, each with two slight variations, for a total of 24 unique passages in 24 conditions. Passages 1–8 were poems with a fixed rhyming scheme. Passages 9–16 were narratives high in descriptive detail and dialogue. Passages 17–24 were narratives broken up into 3 or 4 separate pages as in Experiment 1. All were about 500 words long.

Introspective Reports: As in Experiment 1, participants provided general and concrete introspective reports. As in Experiment 3, participants reported on their experiences of (1) an inner voice, (2) visual imagery, (3) the words on the page,

and (4) mind wandering (Table 4.13). Participants provided general reports before and after reading the passage.

Objective Measures of Experience: After reading the passage, all participants answered two reading comprehension questions. Additionally, conditions 1–8 had a phonological question, conditions 9–16 had two visual detail questions, and conditions 17–24 had three visual perception questions. The phonological question asked participants to identify a word that rhymed with a novel name in the text (e.g. “Tenaïsse”, answer “vice”), the pronunciation of which was fixed by the rhyming scheme of the poem (e.g. “And their gallant moves precise / Sailing safely into port / Chased by beautiful Tenaïsse?”). The Visual Detail Group consisted of two questions asking participants to identify the colors of two nouns (determined by an adjective preceding the nouns in the text). The answers were semantically similar but orthographically distinct (“golden” in the text, “yellow” in the answer). The Visual Perception Group asked participants two phrasal location questions, as in Experiment 1, and to identify the font used in the passage. All questions were multiple choice with four answer options.

Table 4.13

The question text for general and concrete introspective reports.

Report Type		Question Text
Initial General Reports	Inner Voice	How often do you experience an inner voice when you read? Examples: you hear a voice reading in your head, you hear the characters speaking in your head.
	Visual Imagery	How often do you experience visual imagery when you read? Examples: you see pictures from the story in your mind, you see a scene acted out like a movie in your head.
	Words on the Page	How often do you NOT experience the words on the page when you read?Example: Your mind is filled with the ideas in the story and not the actual black letters against the white background
	Mind Wandering	How often do you find your mind wandering when you read? Examples: you are thinking entirely unrelated thoughts, you are spacing out and can't focus on the text.
Final General Reports	Inner Voice	While reading the passage, how often did you experience an inner voice? Examples: you heard a voice reading in your head, you heard the characters speaking in your mind.
	Visual Imagery	While reading the passage, how often did you experience visual imagery? Examples: you saw pictures from the story in your mind, you saw a scene acted out like a movie in your head.
	Words on the Page	While reading the passage, how often did you NOT experience the words on the screen? Example: your mind was filled with the ideas in the story and not the actual black letters against the white background.
	Mind Wandering	While reading the passage, how often did you find your mind wandering?Examples: you were thinking thoughts entirely unrelated to the reading, you were spacing out and not focusing on the text.
Concrete Reports	Inner Voice	In the final split-second before the beep, did you experience an INNER VOICE?
	Visual Imagery	In the final split-second before the beep, did you experience VISUAL IMAGERY?
	Words on the Page	In the final split-second before the beep, did you have a visual experience of the WORDS ON THE SCREEN?
	Mind Wandering	In the final split-second before the beep, was your MIND WANDERING?

Procedure

This experiment used a similar procedure as Experiment 1. Participants took part in the experiment on their own computers using a web browser and the Lime Survey platform. After a sound check, participants were instructed to adjust the volume so that the sound was distinct but not startling. Participants provided general reports on their experience while reading and then read 1 of the 24 passages determined at random. While reading the passage, participants heard a 500 Hz tone at a random time 30-90 seconds after the page loaded. This prompted a new page to load, and instructed participants to provide concrete reports on their experiences. Participants then returned to the narrative and finished reading. Afterwards, participants answered two comprehension questions and 1–3 questions designed to be objective measures of experience. Finally, participants provided another set of general reports on their experience while reading. The entire study took about 15 minutes.

Results and Discussion

122 participants were excluded because they answered both comprehension questions incorrectly. 8 participants were excluded because they spent less than ten seconds reading the passage after providing concrete reports. 9 participants were excluded because they explicitly stated that they had finished reading when prompted to provide concrete reports. After these exclusions, 1,236 participants were included in the analysis.

Introspective Reports: This study is the first to include concrete reports mind wandering, so it is worthwhile to survey all of the introspective reports provided. As in Experiment 1, general reports of an inner voice ($M = 4.62$, $SD = 1.86$), visual imagery ($M = 4.99$, $SD = 1.51$), words on the page ($M = 4.27$, $SD = 1.59$), and mind wandering ($M = 3.50$, $SD = 1.70$) all fall ± 1 from the median Likert response of 4, explicitly labeled “Half of the Time” (Figure 4.6). An interesting replication of reports from Experiment 1 is that only 2% of participants reported never experiencing visual imagery. Turning to concrete reports, 732 (59%) participants reported an inner voice, 865 (70%) reported visual imagery, 631 (56%) reported words on the page, and 353 (29%) reported mind wandering (Figure 4.7). This final result is consistent with Schooler et al. (2004) who found that participants are “caught” mind wandering while reading in 23% of probes.

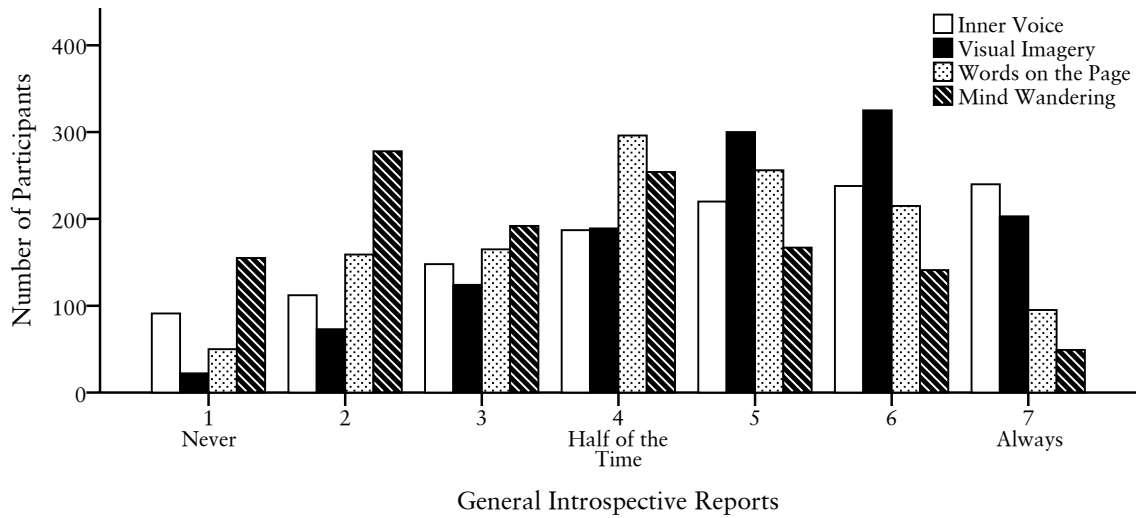


Figure 4.6
Histogram of general reports of an inner voice, visual imagery, words on the page, and mind wandering (N = 1,236).

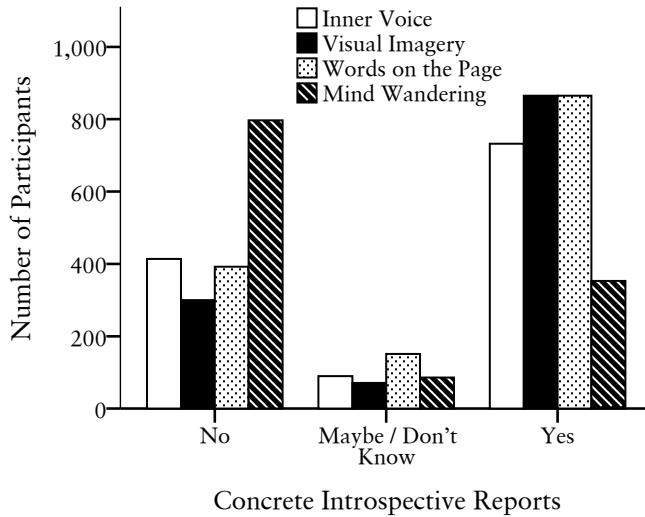


Figure 4.7
Histogram of concrete reports of an inner voice, visual imagery, words on the page, and mind wandering.

Objective Measures of Experience: Mean accuracy for all question groups and subgroups was above chance (Table 4.14). Participants recruited through Amazon MTurk performed significantly better on all groups than those recruited through UC Riverside ($p < .05$).

Table 4.14
Mean accuracy for the objective measure of experience.

Question Group	Question Subgroup	MTurk Accuracy	UC Riverside Accuracy
Phonological (39%)	—	49%	31%
Visual Detail (59%)	—	68%	53%
Visual Perception (50%)	Location of phrase (45%) Identification of font (58%)	53%	47%
Comprehension (83%)	—	88%	79%

A two-tailed Pearson correlation was used to assess the relationship between introspective reports and objective measures of experience (Table 4.15). As in Experiment 1, none of the results were significant.

Table 4.15

Table of Pearson correlations between introspective reports and performance on objective measures of experience. None of the results were significant ($p < .05$).

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Phonological Group	Inner Voice Reports	General Concrete	.040 ($p = .408$) .083 ($p = .084$)
Visual Detail Group	Visual Imagery Reports	General Concrete	.092 ($p = .061$) .005 ($p = .923$)
Visual Perception Group	Words on the Page Reports	General Concrete	-.003 ($p = .949$) -.022 ($p = .662$)

As in Experiment 2, there is a robust relationships between introspective reports and reading comprehension (Table 4.16). Because of the shared variance between introspective reports and reading comprehension, it is plausible that performance on the objective measures of experience is partially due to reading comprehension or participant engagement. As in Experiment 2, I tested this using three stepwise regressions (Table 4.17). The first model was not predictive of performance on the Phonological Group. The second and third models were predictive of performance on the Visual Detail Group and the Visual Perception Group, but the only significant predictor for both was accuracy on comprehension questions.

Table 4.16

Table of Pearson correlations between introspective reports and performance on comprehension questions. The negative correlations for reports of mind wandering are expected.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Comprehension Questions	Inner Voice Reports	General Concrete	.059 (p = .030) .056 (p = .051)
Comprehension Questions	Visual Imagery Reports	General Concrete	.191 (p < .000) .116 (p < .000)
Comprehension Questions	Words on the Page Reports	General Concrete	.098 (p < .000) .102 (p < .000)
Comprehension Questions	Mind Wandering Reports	General Concrete	-.337 (p < .000) -.211 (p < .000)

Table 4.17

Stepwise regression models for predicting objective measures of experience using the corresponding reports and reading comprehension.

Dependent Variable	Independent Variable	Model	Predictor
Phonological Group	(1) Comprehension questions (2) General reports of inner voice (3) Concrete reports of inner voice (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	Not significant	
Visual Detail Group	(1) Comprehension questions (2) General reports of visual imagery (3) Concrete reports of visual imagery (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	R ² = .018 F(1, 410) = 7.58 p = .006	IV (1) b = .207 SE = .075 p = .006
Visual Perception Group	(1) Comprehension questions (2) General reports of words on the page (3) Concrete reports of words on the page (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	R ² = .044 F(1, 387) = 17.90 p = .000	IV (1) b = .387 SE = .091 p = .000

Are introspective reports and behavior really this disconnected? Perhaps the modality of an experience is less important than whether or not our minds are wandering. To test this, I performed a series of two-tailed Pearson correlations between reports of mind wandering and performance on the objective measures of experience (Table 4.18). Reports of mind wandering were negatively correlated with almost all behavioral measures. Because mind wandering covaries with performance on comprehension questions, again it is possible that these results are due to reading comprehension or participant engagement. To test this, I performed three stepwise regressions (Table 4.19). For the same reasons as stated in Experiment 2, these regressions include an interaction variable between reports of mind wandering and performance on the comprehension questions.

Despite the robust correlations between reports of mind wandering and performance on objective measures, these stepwise regressions show that the relationship was nuanced. General reports of mind wandering were predictive of performance on the Visual Detail Group, however, only comprehension questions were predictive on the Visual Perception Group. For the Visual Detail Group, both general reports of mind wandering and accuracy on comprehension questions were predictive. This is consistent with results from the previous experiments, and the correlation between subjective and objective measures of the experience of the words on the page has consistently been the weakest.

Table 4.18

Table of Pearson correlations between reports of mind wandering and performance on objective measures of experience.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Phonological Group	Mind Wandering Reports	General Concrete	-.117 (p = .015) .013 (p = .792)
Visual Detail Group	Mind Wandering Reports	General Concrete	-.144 (p = .003) -.013 (p = .037)
Visual Perception Group	Mind Wandering Reports	General Concrete	-.103 (p = .043) .021 (p = .682)
Comprehension Group	Mind Wandering Reports	General Concrete	-.283 (p < .000) -.144 (p < .000)

Table 4.19

Stepwise regression models for predicting objective measures of experience using reports of mind wandering and accuracy on reading comprehension questions.

Dependent Variable	Independent Variable	Model	Predictor
Phonological Group	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .014$ $F(1, 433) = 5.96$ $p = .015$	IV (2) $b = -.033$ $SE = .014$ $p = .015$
Visual Detail Group	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .030$ $F(2, 409) = 6.34$ $p = .002$	IV (2) $b = -.050, SE = .022$ $p = .025$ IV (1) $b = .156, SE = .078$ $p = .048$
Visual Perception Group	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .044$ $F(1, 387) = 17.90$ $p < .000$	IV (1) $b = .387$ $SE = .091$ $p < .000$

V. EXPERIMENT FIVE

Because the correlations between subjective and objective measures of the experience of the words on the page have consistently been the weakest, this experiment does not include these reports. In addition to a series of visual detail questions, I added a mental folding task to the objective measures of visual imagery. The psychological mechanism behind mental folding has been central to the modern “imagery debate” between Kosslyn and Pylyshyn (Kosslyn, 1994; Pylyshyn, 2002; Shepard & Metzler, 1971; Shepard & Feng, 1972). I hypothesized that the experience of visual imagery facilitates performance on a mental folding task. Further, drawing on a classic study on the phonological loop, I added a novel objective measure of the experience of an inner voice. Baddeley (1966) found that short term memory is impaired when a list of words is phonologically similar (e.g. mad, man, map) when compared to a list of words that is semantically or orthographically similar (e.g. pen, rig, day). He argued that this is the result of phonological interference in short-term memory. In the present experiment, participants performed this sort of Baddeley Task, and I hypothesized that performance will decrease with increased reports of an inner voice because of phonological interference between experience and short-term memory.

This experiment used a within-subjects design. As in Experiment 4, participants were excluded if they answered all of the comprehension questions incorrectly, spent less than ten seconds reading the passage after providing concrete reports, or explicitly stated that they had finished reading at the time of the beep.

Methods

Participants

A total of 616 participants from the United States were recruited through Amazon MTurk (301 female, 315 male; mean age = 34.2 years). Participants were paid \$3.00 and were randomly assigned to three groups, each of which read a different passage: 192 read the White Noise passage, 180 read the Clover passage, and 209 read The Disk passage.

Materials

Text: The passages used in this experiment were slightly modified versions of passages from past experiments. The first passage was the opening paragraphs of White Noise from Experiment 1. The second was the Clover passage from Experiment 2. The third was the shortened version of The Disk from Experiment 2. All were about 500 words long and were presented in a single page of text.

Introspective Reports: This experiment recorded general and concrete reports using the same method as Experiment 4. To test for the robustness of the introspective reports over minor textual changes, I changed the final set of general reports to exactly mirror the initial set of general reports. In this way, the two sets of general reports are a measure of test-retest reliability (Table 4.20). Participants were not asked to report on their experience of the words on the page.

Objective Measure of Experience: In all three passages, a color adjective proceeded a noun in the first sentence. In the middle third of the passage, a different color adjective proceeded a different noun. The Visual Detail Group consisted of two questions, each asking the participants to identify the color of one of the objects. The correct answer was semantically similar to the adjective in the text but orthographically distinct (e.g. “golden” in the text, “yellow” in the answer). Two general comprehension questions formed the Comprehension Group. These questions all had four possible answers randomly displayed. The Folding Task consisted of ten questions that tested the ability to mentally fold a piece of paper (Ekstrom et al., 1976). In a detailed page of instructions, participants were shown how to mentally fold a piece of paper that was then punctured (Figure 4.8). The correct answer was the image of the unfolded paper with puncture marks in the appropriate places. Participants chose from five answer options. The Phonological Group consisted of a Baddeley Task. Participants were shown 5 words in series for 1 second each, drawn at random from three sets of ten words. They were then instructed to recall the words in order from a list of the entire set. Set A was used for training, and results were not included in the analysis (old, deep, foul, late, safe, great, strong, thin, long, broad). Set B contained ten acoustically similar words (mad, man, map, mat, max, can, cad, cap, cat, cab), and set C contained a control set of acoustically different words with an equal Thorndike-Lorge frequency (pen, rig, day, bar, cow, sup, pit, hot, few, bun) (Thorndike & Lorge, 1944). After participants performed a training round on set A, they performed the

task six times, alternating between set B and set C. The experimental design and the word sets come from Baddeley (1966).

Table 4.20

The question text of general and concrete introspective reports. Unlike the previous experiments, the initial and final set of general reports use the same wording.

Report Type		Question Text
Initial General Reports	Inner Voice	How often do you experience an inner voice when you read? Examples: You hear a voice reading in your head, you hear the characters speaking in your mind.
	Visual Imagery	How often do you experience visual imagery when you read? Examples: You see pictures from the story in your mind, you see a scene acted out like a movie in your head.
	Mind Wandering	How often do you find your mind wandering when you read? Examples: You are thinking thoughts entirely unrelated to the reading, you are spacing out and aren't focusing on the text.
Final General Reports	Inner Voice	How often do you experience an inner voice when you read? Examples: You hear a voice reading in your head, you hear the characters speaking in your mind.
	Visual Imagery	How often do you experience visual imagery when you read? Examples: You see pictures from the story in your mind, you see a scene acted out like a movie in your head.
	Mind Wandering	How often do you find your mind wandering when you read? Examples: You are thinking thoughts entirely unrelated to the reading, you are spacing out and aren't focusing on the text.
Concrete Reports	Inner Voice	In the final split-second before the beep, did you experience an INNER VOICE?
	Visual Imagery	In the final split-second before the beep, did you experience VISUAL IMAGERY?
	Mind Wandering	In the final split-second before the beep, was your MIND WANDERING?

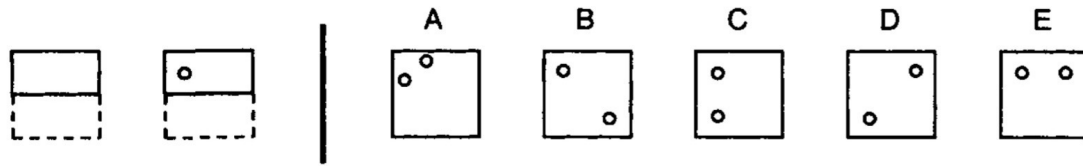


Figure 4.8

An example of a mental folding question. Although this one is quite easy, the most difficult questions required up to three non-symmetrical folds. The correct answer is “A”.

Procedure

This study used a similar procedure as Experiment 1. Participants took part in the experiment on their own computers using a web browser and the Lime Survey platform. After a sound check, participants were instructed to adjust the volume so that the sound was distinct but not startling. Participants provided general reports on their experiences while reading and then read one of the three passages determined at random. While reading the passage, participants heard a 500 Hz tone at a random time 30-90 seconds after the page loaded. This prompted a new page to load, and instructed participants to provide concrete reports of their experiences. Participants then returned to the narrative and finished reading. Afterwards, participants answered questions from the Visual Detail Group and Comprehension Group presented in random order. Next the participants provided another set of general reports on their experience while reading. Finally, participants performed the Mental Folding Task and the Baddeley Task, given in random order. The experiment took about 20 minutes to complete.

Results and Discussion

Data for 80 participants were excluded because they spent less than ten seconds reading the passage after providing concrete reports. Data from 31 more were excluded because they missed both comprehension questions. After these exclusions, data from 550 participants were included in the analysis.

Introspective Reports: I performed a one-way ANOVA to examine the relationship between introspective reports and the three passages. There was no significant difference in reports between the three conditions.

Objective Measures of Experience: Mean accuracy for all questions was significantly above chance ($p < .05$) (Table 4.21).

Table 4.21
Mean accuracy for the objective measure of experiences.

Question Group	Question Subgroup
Visual Detail (73%)	—
Folding Task (59%)	—
Baddeley Task (61%)	—
Comprehension (92%)	—

To analyze the relationship between objective and subjective measures of experience, I performed a series of two-tailed Pearson correlations (Table 4.22). As in the other experiments, the connection between introspective reports and performance on objective measures of experience is not easy to interpret. Although concrete reports of visual imagery are significantly correlated with performance on the Visual Group, general reports are not. Unexpectedly, the correlation between reports of visual imagery and performance on the Folding Task is negative. I hypothesized a positive correlation on the assumption that the experience of visual imagery would covarying with the type of visuospatial processing required to mentally fold a piece of paper. These results suggest that visual imagery could be a detriment to mental visuospatial manipulation.

Table 4.22

Table of Pearson correlations between introspective reports and performance on objective measures of experience. Unexpectedly, there is a negative correlation between general reports of visual imagery and performance on the Folding Task.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Visual Group	Visual Imagery Reports	General Concrete	.071 (p = .097) .105 (p = .014)
Folding Task	Visual Imagery Reports	General Concrete	-.087 (p = .042) .056 (p = .193)
Baddeley Task	Inner Voice Reports	General Concrete	-.028 (p = .508) -.008 (p = .852)

As in Experiments 2 and 4, there is a robust relationships between introspective reports and reading comprehension (Table 4.23). As before, I investigated the connection between introspective reports, reading comprehension, and objective measures of experience using three stepwise regressions (Table 4.24). The first model found that concrete reports of visual imagery were predictive of performance on the Visual Detail Group. The second model found that performance on the Folding Task was predicted by both general and concrete reports of visual imagery. As expected given the negative correlation found above, the beta for general and concrete reports of visual imagery have opposite signs. This makes the results difficult to interpret. According to this model, increased mental folding performance is predicted by high concrete reports of visual imagery and low general reports of visual imagery. One possible explanation is that general reports are merely faux-generalizations based on salient or prototypical experiences. When it comes to visual imagery, the self-narrative that a reader has pervasive experiences does not mean that they actually do. If true, this issue extends throughout the entire paper and warrants further discussion.

Table 4.23

Table of Pearson correlations between introspective reports and performance on comprehension questions.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Comprehension Questions	Inner Voice Reports	General Concrete	.028 (p = .515) .085 (p = .047)
Comprehension Questions	Visual Imagery Reports	General Concrete	.030 (p = .481) .087 (p = .041)
Comprehension Questions	Mind Wandering Reports	General Concrete	-.110 (p = .010) -.117 (p = .006)

Table 4.24

Stepwise regression models for predicting objective measures of experience using the corresponding introspective report and reading comprehension.

Dependent Variable	Independent Variable	Model	Predictor
Visual Detail Group	(1) Comprehension questions (2) General reports of visual imagery (3) Concrete reports of visual imagery (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .011$ $F(1, 548) = 6.11$ $p = .014$	IV (3) $b = .100$ $SE = .041$ $p = .014$
Mental Folding Task	(1) Comprehension questions (2) General reports of visual imagery (3) Concrete reports of visual imagery (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .016$ $F(2, 547) = 4.51$ $p = .011$	IV (2) $b = -.223, SE = .082$ $p = .007$ IV (3) $b = .354, SE = .161$ $p = .028$
Baddeley Task	(1) Comprehension questions (2) General reports of inner voice (3) Concrete reports of inner voice (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	Not significant	

To investigate the effects of mind wandering, I performed a series of two-tailed Pearson correlations between reports of mind wandering and performance

on the objective measures of experience (Table 4.25). Reports of mind wandering are negatively correlated with performance on the Mental Folding Task and the Baddeley Task. As above, I performed three sets of stepwise regressions to test if this effect is predicted by reading comprehension alone (Table 4.26). The first model was not predictive of performance on the Visual Detail Group. In the second model, concrete reports of mind wandering are predictive of performance on the Mental Folding Task. Finally, the third model found that the relationship between reports of mind wandering and performance on the Baddeley Task differs according to reading comprehension. Although the results are merely suggestive, it could be that as reading comprehension increases, so does the negative effect of mind wandering. Alternatively, as reading comprehension increases, mind wandering does not have such a detrimental effect. Unfortunately, the data is little more than suggestive of these two interpretations.

Table 4.25

Table of Pearson correlations between performance on objective measures of experience and introspective reports of mind wandering.

Objective Measure of Experience	Subjective Measure of Experience		Pearson Correlation
Visual Detail	Mind Wandering Reports	General Concrete	-.066 (p = .123) -.069 (p = .108)
Mental Folding	Mind Wandering Reports	General Concrete	-.055 (p = .202) -.126 (p = .003)
Baddeley Task	Mind Wandering Reports	General Concrete	-.099 (p = .020) -.124 (p = .004)

Table 4.26

Stepwise regression models for predicting performance on objective measures of experience using reports of mind wandering and reading comprehension.

Dependent Variable	Independent Variable	Model	Predictor
Visual Detail Group	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	Not significant	
Mental Folding Task	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .016$ $F(1, 548) = 8.822$ $p = .003$	IV (3) $b = -.437$ $SE = .147$ $p = .003$
Baddeley Task	(1) Comprehension questions (2) General reports of mind wandering (3) Concrete reports of mind wandering (4) Interaction variable between (1) and (2) (5) Interaction variable between (1) and (3)	$R^2 = .017$ $F(1, 548) = 9.393$ $p = .002$	IV (5) $b = -.334$ $SE = .109$ $p = .002$

VI. CONCLUSION

These are the results from five experiments involving two thousand participants, by far the largest experimental study of the phenomenology of reading. Each experiment is difficult to interpret individually, and viewing each study in isolation obscures the larger trends. Chapter 5 is devoted to a general discussion of the results, their interpretations, and their implications for the study of consciousness.

Chapter 5

PHENOMENAL VARIABILITY + FUNCTIONAL ISOLATION

During the Fall of 2014, The Guardian, a newspaper based out of the United Kingdom, ran a collection of articles on topics ranging from the central role of hearing voices in the writing of Charles Dickens to the stigma that goes along with “hearing voices” in your head. Together, these articles formed the Inner Voice Series. For the final entry, the Guardian started an open thread online, asking people to write in and describe their experiences while reading, specifically if they hear an inner voice, what this voice is like, and if they have visual or other types of experiences as well. The short article summarizing these comments forms the most comprehensive written account of the experiences people have while reading. At least, if better exists, I have yet to find it. In their own words, readers recount inner voices, character accents, visual imagery, and the connection between an inner voice and visual imagery (or lack thereof). One person even reports on their experiences before and after becoming deaf.

How could a newspaper, asking for comments online, provide a more complete inventory of the phenomenology of reading than the academic literature on consciousness? As you read this final chapter, keep this question in mind. Introspective reports from the five studies in this dissertation have more in

common with this unscientific, curated collection of experiences than with the published content of philosophy and psychology journals. What the Inner Voice Series and these studies have in common is that they both asked people what they were experiencing instead of focusing on the experiences of the author.

In this chapter, I discuss general conclusions from this set of experiments. First, I look at introspective reports, providing a combined overview of the results and contrasting them with the descriptions of experience in the academic literature. I argue that philosophers and psychologists are beset by introspective error which leads to systematic biases in the ways they describe their experiences. Second, I look at the connection between introspective reports and objective measures of experience. The data suggests that our modal experiences are functionally isolated from the cognitive mechanisms behind the processing of textual information. In contrast, mind wandering has a robust negative correlation with objective measures of experience and reading comprehension. This suggests that widespread aspects of experience are *paraphenomena*, unrelated to the functional aspects of consciousness. When our experiences are not functionally isolated, the relationship between experience and behavior is often one of interference instead of facilitation. I finish by laying out the implications for the relationship between phenomenal and functional notions of consciousness.

I. INTROSPECTIVE REPORTS

Hitherto I have been speaking of “the experience of reading,” as if there were a single phenomenal entity that we all share. In retrospect, this phrase is highly misleading, collapsing a phenomenon with a high degree of variability into a single category. It would be more accurate to speak of *experiences while reading*. Expanding on the critiques from Chapter 2, instead of conceptualizing of the experiences while reading along the lines of perceptual visual experiences, we need a new approach. To repurpose a battered metaphor, our experiences while reading are like phenomenal performances in, you guessed it, a Cartesian theater. There are centuries of philosophical baggage that come along with the idea of a Cartesian theater, and I do not want to bring all of that into this conversation. However, the comparison to a performance corrects for biases found in the existing comparison to perceptual experiences. A theatrical performance captures the dynamic variations in consciousness as well as the prevalence of modal experience, the sound of the performers and the visual images of them. All of experience is a stage, and we are our own audience and actor (see Kivy, 2006 for more on reading as a performance).

The results from these studies are unequivocal. There is a tremendous amount of variability in the experiences people have while reading. This variability comes in two forms: interpersonal variability and variability over time. To get a macro view of the range of experiences, let’s begin by looking at the combined introspective reports from this set of studies. Not all experiments asked

participants to provide all four types of reports, so taken together, these experiments form a bricolage of data. To account for this and allow us to make comparisons between different types of reports, we can look at the distribution of reports for each of the four types of experiences: an inner voice, visual imagery, words on the page, and mind wandering. Figure 5.1 shows the percentage of participants who provided each type of general reports. Figure 5.2 shows the combined concrete reports from Experiments 1, 2, 4, and 5 (Experiment 3 did not ask participants to provide concrete reports). Finally, Figure 5.3 combines the responses from Experiments 1 and 2 where participants provided more than one concrete report. This allows us to see the changes in concrete reports over time for a single individual.

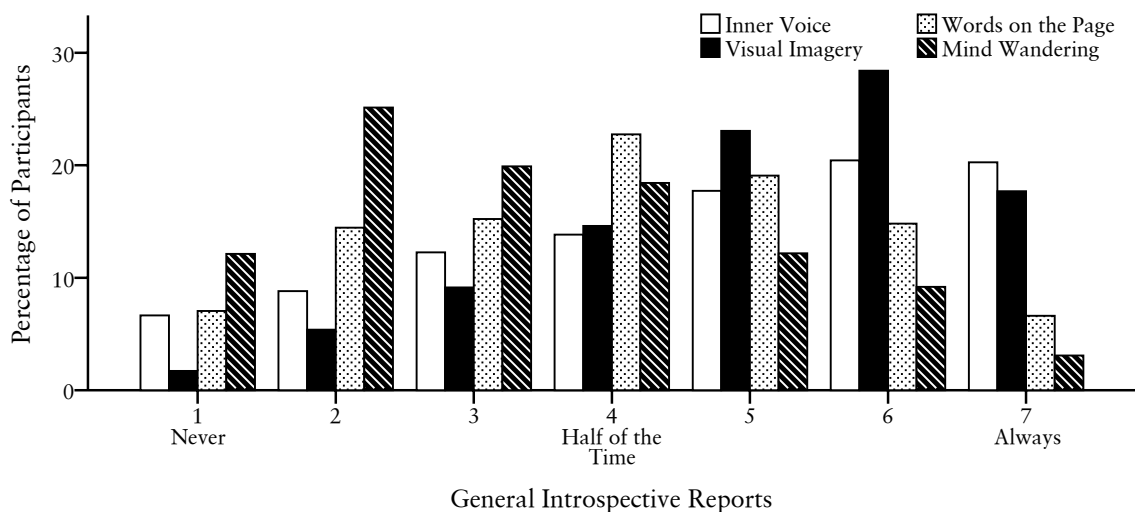


Figure 5.1 Percentage of participants who provided general reports of inner voice, visual imagery, words on the page, and mind wandering from all five experiments.

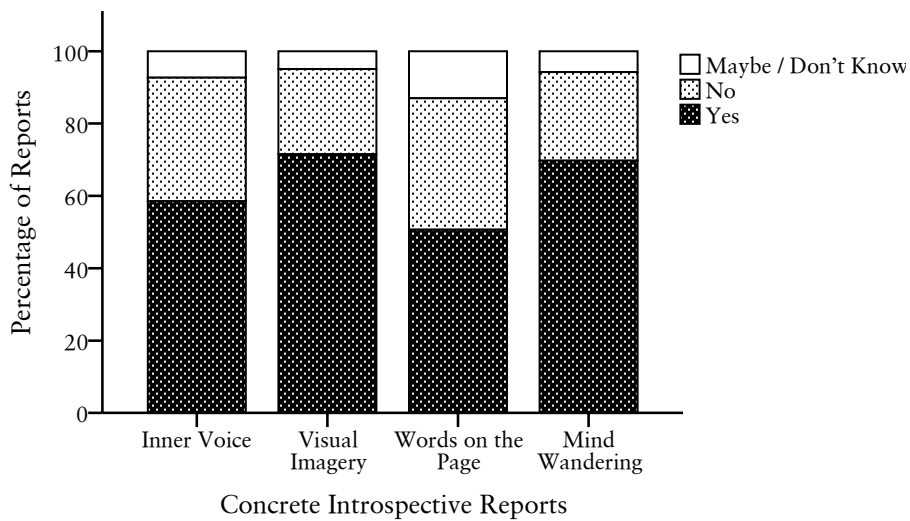


Figure 5.2
The combined concrete reports for Experiments 1, 2, 3, and 5, presented as a percentage of each type of report.

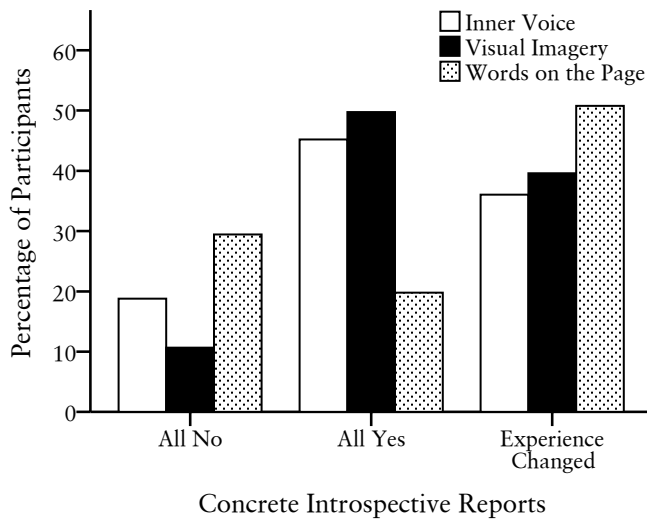


Figure 5.3
The combined concrete reports from Experiments 1 and 2, separated into three categories: the percentage of participants who consistently reported an experience, who consistently denied and experience, and who reported a change in experience.

The colorful variability in reports between readers is manifest. Barring a profound skepticism towards introspective reports of such coarse-grained features of consciousness, it is hard to escape the conclusion that some people experience an inner voice while others do not, and likewise for visual imagery and the perceptual experience of the words on the page. Similarly, the evidence that experiences change over time is equally compelling. As the general reports show, at most 27% of the participants reported stable experiences, a Likert report at the extremes, never or always. The overwhelming majority of readers reported a phenomenal flux, an ever changing inner performance, sometimes exhibiting an inner voice, while at other times, they are confronted with inner silence. Figure 5.3 provides another perspective on the protean nature of consciousness. About half of the participants provided different concrete reports during the two or three probes in these short experiment. Remember, we are talking about salient, macro aspects of experience, reported over the course about five minutes of reading. The variability in experience is not hiding, it is right there, plain for all to see.

One striking feature of these results is the vast distance between the introspective reports provided in this set of experiments and those found in the academic literature. Let's look at each, starting with the experience of the words on the page. Charles Siewert (1998) argues that we always experience the words on the page when we read. According to Siewert, we are not conscious of every letter on the page, but rather "vaguely identifiable groups of characters" (1998, p. 249). However, out of 1,661 participants, only 110 (7%) reported always

experiencing the words on the page. Similarly, out of 1,687 concrete reports, only 855 (51%) were positive reports of this experience. While Siewert expresses a nuanced and well-argued position on the experiences people have when they read, the reports of actual readers suggests that he is mistaken.

When I began this set of experiments, I expected reports of words on the page to be considerably more common than the data indicate. I was convinced by Daniel Dennett (1996), who famously argues that many of the details people claim they experience are not actually represented in consciousness. In his own words “when we marvel, in those moments of heightened self-consciousness, at the glorious richness of our conscious experience, the richness we marvel at is actually the richness of the world outside, in all its ravishing detail. It does not ‘enter’ our conscious minds, but is simply available” (1996, 407–8). On his view, it is a fallacy to infer facts about consciousness from facts about the world, however appealing the inference might seem. Before running these experiments, I assumed that this implicit, unconscious inference would be widespread, and because of this, I expected the vast majority of participants to report the experience of the words on the page. After all, there is a folk sense of “experience” in which we always experience the words on the page while reading. So the thought goes, how could we read if we do not see the words? Yet the fact remains, combining the three types of modal reports from these five experiments and comparing them with a one-way ANOVA, the experience of the words on the page was significantly less commonly ($M = 4.03$, $SD = 1.66$) than the experience of an inner voice ($M = 4.70$,

SD = 1.85) or visual imagery ($M = 5.08$, $SD = 1.50$), $F(93, 8093) = 444.78$, $p < .000$. This does not vindicate Julian Jaynes when he says “right at this moment... as you read, you are not conscious of the letters or even of the words or even of the syntax, or the sentences, and punctuation, but only of their meaning” (1976, p. 26). This extreme denial of visual experience is no more supported by the evidence than Siewert’s extreme embrace of it. In general reports, only 117 (7%) of the participants denied ever experiencing the words on the page, while 612 (36%) of the concrete reports were negative. Both Siewert and Jaynes fail to take into account the high level of variability in experience.

Not to be too heavy handed, Siewert believes that many people deny the visual experience of the words on the page because they are unable to precisely articulate what they are experiencing. He does not want to imply that “only a being capable of saying what it is conscious of can consciously see” (p. 251). Indeed, reports of words on the page are the most confusing to interpret. Out of the participants who reported the words on the page, only a small number explicitly mentioned the text itself in their written descriptions. When they did, it was often in contrast with other experiences, such as “I really did not experience much except for words on the page because the story did not reel me into the point of visual imagery,” “I was just reading the words, but not actually thinking about it,” or “I tried to picture the scene but I was mostly reading the words without any experience in my mind.” The converse is that, in describing modal experiences, participants often contrasted them with the experience of the words on the page.

For example, one participant said “visual imagery with inner speech narration. Words were barely noticeable,” while another claimed “Prior to the beep, I was visualizing the story in a movie-like manner of move in day on a college campus. I was not focused on the words so much so as the imagery.” Although I can do little more than speculate on the cause, here are two possibilities. First, participants could be treating the experience of words on the page as a phenomenal default, an assumed experience in the absence of other, usually modal, experiences. On this explanation, when spacing out or in doubt, participants opportunistically fill their introspective reports with the nearest relevant aspect of the world, in this case, the black words on the white background. Second, it could be that the experience of the text corresponds to a lack of engagement or poor comprehension. On this view, the contrast with modal experiences is not accidental, but is rather an important aspect of the experience itself. For example, if someone is deep into a narrative, they might transcend the words to have visual or auditory imagery, but if not, their consciousness is left with mere text.

Next, consider reports of an inner voice. When Bernard Baars said “human beings talk to themselves every moment of the waking day. Most readers of this sentence are doing it just now” (2003a, p. 7), he was not merely mistaken, the results from this study suggests that he was radically off base. Out of 2,211 total readers, only 448 (20%) say they always experience an inner voice, while 147 (7%) claim that they never do. Similarly, out of 2,237 probes for concrete reports, 1,309 (58%) were positive reports of an inner voice, while 764 (34%) were

negative. Perhaps Baars is among the 20% of constant voicers out there and is accurately describing his experience, but even still, the universal attribution of this experience is empirically dubious. Oddly, this claim comes from a psychologist — one who specializes on consciousness, no less — who should know the pitfalls of making claims about others from his armchair. Likewise, when Max Velmans (2009) describes the experience of reading the sentence “If we don’t increase the dustmen’s wages, they will refuse to take the refuse” by saying “Note that on its first occurrence in your phonemic imagery or ‘covert speech’, the word ‘refuse’ was (silently) pronounced with the stress on the second syllable (refuse), while on its second occurrence the stress was on the first syllable (refuse)” (249), his description undoubtably captures the experience of many readers. However, the application of this description to everyone does not have empirical support.

Visual imagery was the most widespread experience reported while reading in these experiments. Of the 2,211 participants who provided general reports of visual imagery, only 38 (2%) said they never experienced it, while 91 (18%) claimed to always have the experience. Out of 2,237 concrete reports, 1,600 (72%) were positive reports of visual imagery, while 526 (24%) were negative. These experiments do not support anything like a universal description of the experiences while reading, but this is the closest it comes: (almost) everyone experiences some visual imagery at times while reading. Interestingly, in researching this dissertation, I could not find a single published statement denying the existence of visual imagery, although people deny it in specific cases. So when

Wittgenstein says “[while reading] I have impressions, see pictures in my mind’s eye, etc. I make the story pass before me like pictures, like a cartoon story” (1946–1948/1975, p. 44), or Dennett claims “the act of reading, and interpreting, a text such as a novel creates some new things in my imagination: images of the characters doing the deeds” (1991, p. 366), they are describing experiences that nearly everyone is familiar with.

Finally, we have reports of mind wandering, an experience that has consistently eluded introspective discovery by philosophers. In six years of research, the only mention of mind wandering I can find among philosophers is an odd, hundred year old critique of Plato “As one reads a dialogue of Plato, one’s mind often wanders far from the printed page; indeed, this happens so frequently that there is created a feeling of difficulty in concentrating the attention on the dialogues, which inclines one to criticize Plato for inability to retain his reader’s thoughts” (Austin, 1922, p. 254). Other references to mind wandering must exist, but they are a rare topic of discussion. Looking at the extreme views in these experiments, out of 1,661 participants, 244 (15%) said that their minds never wander, while 62 (4%) said that it always does. In the 1,786 concrete reports, 437 (25%) were positive reports were of mind wandering, while 1,246 (70%) were negative. Although mind wandering is largely ignored in the philosophical literature, as I have already discussed, it has been the focus of a small literature in psychology. Notably, these results are consistent with Schooler et al. (2004) who found that participants are “caught” mind wandering while reading in 23% of

probes. Although I will have more to say on the matter soon, mind wandering is the only introspective report that has a robust connection with behavior, making its absence in philosophy all the more startling. The pessimist in me wonders if it perhaps comes from some unconscious desire to avoid conclusions with empirical consequences. I suspect the main cause of this oversight is the idea that mind wandering is a perversion of the experience of reading, not a constituent of it. I will have more to say about the fetishizing of an “ideal” experience of reading in the next section.

From these studies, we know two of the factors that contribute to the variability in experience. First, although gender, age, and the number of hours spent reading each week had no significant effect, there was a correlation between reports and the characteristics of the passage. Participants reported a higher level of inner voice while reading a dialogue, and reading abstract philosophy lead to lower reports of visual imagery. In this way, the modal characteristics of the passage manifested in a corresponding modal experience. Characters talking give rise to an inner voice, and a lack of scene or concrete nouns inhibits visual imagery. Of course, since the characteristics of the passage affect the experience of readers, all of the introspective reports provided in these experiments must be reinterpreted as representative of reading in a specific context. Although there was a broad range of passages, from poems and plays to narratives and abstract philosophy, the types of texts are infinitely varied. I recently read a book containing nothing but tattoos of text, hundred of poems and phrases that people

have chosen to have permanently etched onto their bodies. This caused me to wonder, what is the experience of reading ink on someone's skin? I do not expect it to be radically different from the experiences of reading fictional narratives, but as these experiments show, we must be careful to check our assumptions about what people will, and will not, experience in a given context.

Second, Experiment 2 found that reports of an inner voice were lower for non-native English speakers ($M = 4.79$, $SD = 1.79$) than for native speakers ($M = 5.49$, $SD = 1.61$), a 13% decrease. There was no significant difference between these two groups in Experiment 1, although the small sample size did not allow for the detection of an effect of this size.

Experiment 3 did not find that instructing participants to focus on a specific experience had a significant effect on their subsequent introspective reports. This was an unexpected result, since this is an empirical test of the refrigerator light effect. However, the statistical power of this study was low, with only 228 participants split into four groups. A post hoc test using a Bonferroni correction found that the 95% confidence interval for the mean difference in reports between conditions was 1.89–2.15 out of a seven-point Likert scale.

II. THE OTHER MOORE'S LAW

These results suggest a radical reassessment of our confidence in our self-knowledge, specifically when it comes to descriptive phenomenology. I will return

to the sources of introspective error from Chapter 2 and give some reasons for their stubborn perseverance in philosophy.

When discussing the results from these experiments, a typical conversation goes something like this: I ask someone what they experience while reading, and they describe their experience in vague general terms. I talk about the experiences that participants have reported in these experiments, and it seems as if my conversational partner is genuinely interested and open to these phenomenal possibilities. About half the time, however, something strange happens, and it seems as if a switch flips in their minds. I can not quite put my finger on it (and no one has said this explicitly), but this is my take on what they are thinking: something as obvious as the results of this study can not be important nor informative.

It should go without saying, but there mere fact that something seems obvious in retrospect does not imply that it is uninformative. Indeed, the results of these experiments are highly informative, and you can see it by contrasting the descriptions of experience from this study with those found in the academic literature. Sure, we should have already known about the wide variations in experience and the ubiquity of mind wandering, but the fact remains that we did not.

The implications of this go to the core of the way that we study, not just the phenomenology of reading, but consciousness itself. As the results from these experiments show, experts routinely provide descriptions that dramatically

mischaracterize the experiences people have while reading, and I suspect, their own experiences as well. Although the variability in experience is obvious once it has been empirically demonstrated, before that, philosophers and psychologists possessed an equally high degree of confidence in their old descriptions of experience, descriptions that also seemed obvious. The longstanding assumption among those who perform philosophically trained phenomenology, the floor upon which the armchair sits, is that we can discover truths about consciousness by thinking really hard about it. I do not believe this general method always serves us poorly, but it does more often than we realize. When it comes to descriptive phenomenology, we are beset by all sorts of introspective biases, and the reports of a single person are often not to be trusted. This forms the other Moore's Law: we feel confident in the results of any introspective method. Reflective and sincere thinkers trained in armchair introspection will feel certain of their introspective reports. Practitioners of ESM and the DES method will be convinced that the results they obtain are accurate. Not to assume too much, but I suspect that readers of this experimental thesis likely feel confident in its results. We simply have a poor ability to evaluate introspective claims.

I am not implying that we can not have more confidence in one report over another. However, this evaluation can not be based on the level of confidence we have in any specific report. Instead, we must evaluate introspective methods. Only then can we have faith that reports are accurate, a vicarious confidence that is mediated by the method and the context. As with most empirical measurements,

the truth is probably quite nuanced, and there are likely questions and contexts that favor one introspective method over another. For example, while I believe that naive experimental participants are generally accurate in their descriptions of coarse-grained features of consciousness, I am skeptical of their ability to report finer phenomenological details.

This brings us back to the sources of introspective error discussed in Chapter 2. I already explained the ways that SOME method minimizes these sources of error, but let me say more about the ways that introspective error manifest in introspective reports from the academic literature. This is a topography of introspective error, and their combined effects have severely distorted the way philosophers and psychologists talk about consciousness.

1. Inference Errors – Introspective reports based on inference instead of introspection.

a. Theoretical Entailment: Claiming to experience the words on the page because of a prior theoretical commitment that entails this experience.

b. Theoretical Affinity: Reporting an inner voice because it intuitively fits with a prior theoretical commitment, such as the language of thought view.

c. Experimental Entailment: Claiming to experience an inner voice because of the experimental research on the phonological loop.

d. Stimulus Entailment: Reporting an experience of the words on the page based on the fact that you are looking at text.

- e. Assumption of Default Phenomenology: Reporting the experience of an inner voice because you believe that you always experience an inner voice.
- f. A Priori Phenomenal Constraints: Claiming to always experience visual imagery because of the a priori belief that an experience of reading requires visual imagery.

2. Overgeneralization – Introspective reports that are generalized beyond the introspective evidence.

- a. Salience: Unconsciously inferring a pervasive inner voice because it is a salient experience when you have one.
- b. Prototypical Context: Claiming to always experience visual imagery because it is constant during a prototypical instance of reading, such as while reading descriptive narratives like Harry Potter.
- c. Attribution to Others: Generalizing that everyone experiences an inner voice based on your own introspection.
- d. Richness of Experience: Reporting rich, cross-modal experiences because you can have each constitutive modal experience in isolation.
- e. Directed Attention: Claiming to always experience an inner voice because you have one whenever you consciously decide to look for it. Also called the refrigerator light problem.

3. Short-Term Memory Limitations

- a. Lack of Articulation: Claiming you do not experience the words on the page because you are unable to precisely articulate the words you just read.

4. Phenomenal Heuristics

a. Overextension of Metaphor: Using a metaphor as evidence for the accuracy of introspective reports (e.g. justifying reports that a round object viewed at an obtuse angle looks elliptical by reference to the way such an object is represented in a photo). Metaphors should be used as explanatory heuristics, not as a source of evidence (for a discussion, see Schwitzgebel, 2011, ch. 1-2).

All of us — philosophers, psychologists, experimental subjects, and people on the street — are susceptible to these sources of error. This explains the peculiar symmetry between descriptions found in the academic literature and those of participants in these experiments. If you treat the reports of philosophers and psychologists like those of anyone else, the picture they paint is similar to the one from this study (with the notable exception of mind wandering). The problems come when we treat the reports of experts as more reliable, because like the participants in this study, no single individual captures the wide variability. This is because experts are unknowingly affected by these sources of error, indeed, the descriptions of experience found in the academic literature display the tell-tale signs of introspective bias. They tend to focus on phenomenal slices that obscure the changes in experience over time, they make sweeping general statements about the things people experience, and they tend to describe experience in a way that fits snugly with their own previously established theories. Experimental

phenomenology can act as a check on these sources of bias, since it is probable that the biases of naive participants are different than the biases that affect experts. For example, experts tend to overlook mind wandering, while experimental participants, sitting at their computer and relatively uninvested in the outcome, are keenly aware of it. Further, experts tend to ignore individual differences, since other minds are hard to view from the armchair. The crowd intelligence of a group of experimental participants, however, is excellent at uncovering variation of this sort.

It is one thing to be susceptible to sources of error, quite another to remain in a perpetual state of bias. Why have experts on consciousness, people who have published peer-reviewed papers and books on the topic, consistently taken little action to account for introspective error? Perhaps the greatest impediment to our understanding of consciousness is not ontological, to be settled by debates between dualists and materialists, or epistemic, such as bridging the gap between objective and subjective knowledge. Rather it is technical or methodological, the widespread refusal to apply a systematic research strategy that goes beyond thinking *really hard* about consciousness. What follows are my observations on the causes of this methodological stagnation. I am entering the realm of psychological and social speculation, and although I have no way of evaluating the relative merits of these suggestions, the question is undeniably worthwhile.

One. We can not underestimate the lingering effects of Descartes. Regardless of the introspective evidence or experimental studies that call into

question our self-knowledge, Cartesian infallibilism is a hard view to shake. It not only entails that we know our experience with certainty, but that this is the one piece of knowledge that we have with absolute confidence. The feelings of the *immediacy* or *directness* of our experiences only serve to reinforce the Cartesian position. So the bias says, “how could something that is so central to our perception of the world, in a way that feels non-inferential, be drastically mistaken?”

Two. Skepticism towards introspection could lead to skepticism in other areas of philosophy. Our experiences form a bedrock upon which foundational theories of epistemology are built (e.g. BonJour, 2000) and upon which narrative theories of personal identity rest (e.g. DeGrazia, 2005; MacIntyre, 1989; Taylor, 1989). It could even lead to skepticism towards basic sources of human enjoyment and pleasure, since if we can be mistaken about our experiences, we can also be mistaken about our happiness or pleasure. This would have widespread implications for conative theories of action (e.g. Frankfurt, 1971; 1999) and consequentialist theories of ethics (e.g. Mill, 1861; Singer, 1993).

Three. Longstanding debates in philosophy of mind and epistemology appear to preempt the sort of experimental investigation of consciousness that I have been engaged in. If consciousness is only knowable from the first-person point of view, how can we perform an empirical study with a sample size larger than one? Of course, the answer is evident, as this dissertation demonstrates. I doubt many philosophers would actually defend this claim, but then again, it is not

being explicitly criticized either. Beside, the larger social effect does not require an explicit endorsement. It is enough for a sizable number of philosophers to draw methodological boundaries between two sets: our experiences and things that can be investigated experimentally. This would have a chilling effect on an already cold area of research.

Four. Philosophers often have trouble appreciating a straightforward description of experience because of its theoretical implication (or what they believe to be its theoretical implication). They can not see the trees for the forest. This is evident in the way that philosophers debate descriptions of experience. For example, let's return to a descriptive disagreement I discuss in Chapter 1: what does a round object viewed at an obtuse angle look like. Alva Noë (2004) claims that we experience it as both round and elliptical (from here), while Sean Kelly (2008) holds that we shift between the two (putative) competing descriptions. In his criticism, Kelly accuses Noë of “metaphorically stomping his foot,” saying “I deny that when we are seeing the plate to be circular it also looks elliptical, Noë insists that it does” (p. 685). Kelly goes on to push Noë on the theoretical consequences of his view, or in his words, “trying to get Noë to bite various bullets,” the largest of which is Noë’s simple insistence on his preferred description. I do not want to weigh in on this particular debate, but it is representative of the orientation philosophers have towards descriptive phenomenology. Sometimes an experience is just an experience, and should be taken as such, not explained away. This is one insidious effect of the sources of

introspective error outlined above. Philosophers do not believe that other philosophers are doing introspection well, leading to a culture of mistrust in each other's reports.

Five. People are generally overconfident in their knowledge or abilities when they do not receive critical feedback. Schwitzgebel (2011, Ch. 7) hints at this possibility, arguing that we become cavalier in our introspective reports because we never receive decisive evidence of error, and even if we did, there is no one to scold us when we are wrong. This is the position we are always in with respect to experience. Since there is no one to tell us we are right or wrong, all we have to go in is confidence in our introspective knowledge. As the other Moore's Law states, regardless of their accuracy, introspective reports always feel right.

Six. The intuitive view that experience is causally efficacious suggests that people with the same behavior will have (roughly) the same experiences. I have never seen this view explicitly endorsed, but I suspect many philosophers and psychologists would find it intuitive. Someone who confidently believes that they always experience an inner voice when they read might look around and infer that the other people reading are having the same experience. Assuming that this person is justified in their individual introspective report, and there is a robust one-to-one link between experience and behavior, they would have good evidence for this inference. However, as I have been arguing, both of these assumptions are highly problematic.

Finally, seven. Just below the surface of discussions of consciousness is the myth of an “ideal,” “pure,” or “true” experience. You see the effect strongest in descriptions of perceptual experiences, where the veridicality of perception provides a foothold for normativity. Hallucinations, for example, are instances of misperception, not instances of accurate perception that does not match up with the world. For some questions, this orientation is entirely acceptable, however it often gets distorted when applied to descriptions of experience. We see this at play in the idea of a prototypical experience of reading. I do not think it is an accident that the most commonly cited experiences — namely an inner voice, visual imagery, and the words on the page — are all modal interpretations of the content of a text. Academics are especially adept at extracting information while reading, and this purpose forms an implicit normative phenomenology. These experiences are seen as constituting the core of reading, so much that other experiences, like mind wandering, are seen as corruptions. As compelling as these descriptions are, they are merely a portion of the actual experiences people have while reading. They are analogous to folk theories of mind, unanalyzed beliefs from god knows where that go on to become implanted in the heart of philosophical theories.

In all of this, I speak of the need to evaluate our methods of arriving at introspective knowledge. Although I have made the case for the accuracy of SOME method at length, it is important to view it with the same critical lens that I have been applying to other views. The most developed criticism of SOME method would likely come from Hurburt and Heavey (2006), who have had a

longstanding skepticism towards the types of introspective reports provided by participants in this set of experiments. Given their explanation behind the DES method, I believe they would not be satisfied with the ways SOME method accounts for the sources of introspective error. For Hurlburt and Heavey, accurate introspective reports require an expositional interview where a trained interviewer guides a participant through their introspective biases.

To this criticism, I have three responses. First, these experiments asked participants to report on coarse-grained aspects of experience. I was not interested in the finer points of consciousness, such as difference tones or after images. I simply wanted to know, for example, if they were experiencing an inner voice or not. Reports of macro features of consciousness are less likely to be mistaken. Second, I have adopted a major part of the DES method, and participants provided both general and concrete reports. This goes a long way towards reducing the effect of faux-generalities, since concrete reports are rooted in a specific moment of time. Third, because participants provided three sets of reports — an initial and final set of general reports as well as concrete reports — we can see if participants are changing their reports throughout the experiment. Even assuming that the initial reports are largely guided by prior beliefs or implicit theories, reports at the end of the experiment likely exhibit reduced bias. Because it had the largest number of subjects, let's use Experiment 4 as an example, although results from the other experiments also match this interpretation. The correlations between initial and final general reports of an inner voice ($r = .524$), visual imagery ($r = .$

404), words on the page ($r = .340$), and mind wandering ($r = .416$) were all significant effects ($p < .000$), but still allow for a high degree of variability between the two sets of reports.

In this section, I have been criticizing philosophers and psychologists, but the fault is not all theirs. We readers have also played along, allowing ourselves to be seduced by their introspective judgments. There is something compelling about the sort of simplistic, universal claims about reading found in academic journals. I suspect that these descriptions of experience fit nicely with our own biases as readers, and together they form a feedback loop. When someone unknowingly overgeneralized their experience of an inner voice and tells you about it, the affects of directed attention on your own experiences will show that, sure enough, you have an inner voice as well. This is all well and good for casual introspection, but it is no way to engage in a rigorous scientific investigation. As culpable as readers are, we need not hold ourselves to the same standards as the writers.

III. SUBJECTIVE AND OBJECTIVE MEASURES OF EXPERIENCE

Although conscious experiences play a huge role in our lives, the results of these studies suggest that salient, coarse-grained features of experience have, at most, a minimal effect on the way that we process textual information. The most common experiences that people report while reading — an inner voice, visual imagery, and the words on the page — do not lie at the center of the cognitive system for information processing. This is to say, the modality of our experiences

while reading are causally impotent. It is as if, when God hooked up our consciousness, some of the plugs were connected to introspection and nothing else. This strongly suggests that, with regard to reading, we maintain a skeptical orientation towards a connection between experience and behavior, and the implications extend deep into the relationship between mind and world.

The results from the individual experiments in this dissertation are somewhat difficult to interpret, and viewing each experiment in isolation obscures the larger trends. For example, Experiment 2 found evidence that reports of an inner voice are predictive of accuracy on phonological questions, yet the other four experiments did not. Similarly, there is evidence that general reports of visual imagery are predictive of recollection of visual detail, or at least, this is what the results from Experiment 5 suggest. There was no evidence of this in the other four experiments. At the same time, none of the experiments found a correlation between reports of experiencing the words on the page and subsequent recollection of perceptual information about the text. According to these studies, a reader who claims to experience the words is just as likely to identify the font of the passage as a reader who denies this experience.

When taken together, these experiments suggest that pervasive, salient aspects of experience are not causally efficacious outside of our own introspective reports, or if they are, the effects are small. Of course, the lack of a robust connection between introspective reports and objective measures of experience could be the result of effects that are hard to detect. To test for this, I combined the

results from the five experiments using two separate meta-analyses (Rosenthal and Rosnow, 1991, ch. 22) (Table 5.1). The first meta-analysis combined the correlation coefficients between the subjective and objective measures of experience to provide a cumulative effect size estimate. The second meta-analysis provides an estimate of the probability of obtaining this set of p values if no relationship existed between the subjective and objective measures of experience. To give more weight to the larger studies, each correlation and p value was weighed by its degrees of freedom. In Experiment 3, I included performance on the Stroop Task in the analysis of the Visual Detail Group, and similarly, I included performance on the Baddeley Task in the analysis of the Phonological Group. I am interested in the experiences while reading, and these tasks directly involved this action. The Mental Folding Task and Memory Test were left out of the analysis because they, in and of themselves, did not involve reading directly.

Of the six hypothesized relationships between introspective reports and behavior, the only significant result comes from the meta-analysis of p values for the correlations between general reports of visual imagery and performance on visual detail questions. Even this significant result should be interpreted skeptically. The results from the two meta-analyses are in conflict, and in the analysis of effect size, the 95% confidence interval crosses zero, suggesting that this correlation is not significant after all. Further, I performed 27 analyses in the search for a connection between modal experiences and objective measures of experience. Given the multiple comparisons, one false positives due to chance is expected.

Together, the combined power of these studies allows us to detect medium to large effects. If there is a connection between our modal experiences while reading and behavior, it is a weak one.

Table 5.1

The two meta-analyses of the effect size and p values of the correlations between subjective and objective measures of experience.

Objective Measure of Experience	Subjective Measure of Experience		N	Analysis of Effect Size			Analysis of P Value
				Effect Size (r)	95% C.I. Lower	95% C.I. Upper	Sig. (p)
Phonological Group	Inner Voice Reports	General	2,211	0.055	-0.028	0.137	0.128
		Concrete	1,983	0.071	-0.006	0.147	0.273
Visual Detail Group	Visual Imagery Reports	General	2,211	0.065	-0.018	0.147	0.005
		Concrete	1,983	0.041	-0.036	0.117	0.222
Visual Perception Group	Words on the Page Reports	General	1,661	-0.018	-0.100	0.064	0.650
		Concrete	1,433	-0.023	-0.098	0.052	0.639

In addition to reports of an inner voice, visual imagery, and words on the page, Experiments 3-5 also asked participants to provide reports of mind wanderings. Unlike the modal experiences, each experiment found a significant negative correlation between reports of mind wandering and at least one objective measure of experience. This could be an artifact of multiple comparisons. If you look for enough connections, you are bound to find something eventually. For example, in Experiment 4 I performed a correlation analysis between mind

wandering and all four question groups (the Phonological Group, Visual Detail, Visual Perception, and Comprehension Groups). In contrast, the only correlation analysis I performed on reports of an inner voice was with performance on the Phonological Group. To control for this and to capture broader trends among the experiments, I performed two meta-analyses on the correlations between reports of mind wandering and objective measures of experiencing using the same procedure as above. I combined the results from Experiments 3-5, the only experiments that recorded reports of mind wandering, each weighed by the degrees of freedom (Table 5.2). As before, I included performance on the Stroop Task and the Baddeley Task, but did not include the Folding Task and Memory Tasks.

Table 5.2

The two meta-analyses of the effect size and p values of the correlations between reports of mind wandering and objective measures of experience.

Objective Measure of Experience	Subjective Measure of Experience	N	Analysis of Effect Size			Analysis of P Value	
			Effect Size (r)	95% C.I. Lower	95% C.I. Upper	Sig. (p)	
Phonological Group	Mind Wandering	General	2,014	-0.098	-0.168	-0.026	0.002
		Concrete	1,786	-0.029	-0.094	0.035	0.158
Visual Detail Group	Mind Wandering	General	2,014	-0.103	-0.173	-0.031	0.001
		Concrete	1,786	-0.030	-0.095	0.034	0.011
Visual Perception Group	Mind Wandering	General	1,463	-0.067	-0.134	-0.001	0.020
		Concrete	1,236	0.021	-0.035	0.077	0.682
Comprehension Group	Mind Wandering	General	1,786	-0.231	-0.291	-0.169	0.000
		Concrete	1,786	-0.136	-0.198	-0.072	0.000

There was a robust connections between reports of mind wandering and objective measures of experience. Both meta-analyses found that general reports of mind wandering correlated with performance on all of the objective measures. In addition, the analysis of p values found a significant relationship between concrete reports of mind wandering and performance on the Visual Detail and Comprehension Groups. As it turns out, experience is connected with behavior, it is just that the relationship is entirely different than I initially suspected. The modality of the experience is inconsequential. What matters is the content of experience, that is, whether or not someone is thinking about, or engaging with, the text.

This is consistent with the long history of conflicting research on the relationship between visual imagery and behavior. Marks' (1973) Vividness of Visual Imagery Questionnaire (VVIQ) is the most widely studied measure of visual imagery, and the results of decades of research are equally negative. As I discuss in Chapter 3, McKelvie's (1995) meta-analysis of 250 VVIQ studies failed to find a robust relationship between this test and visual creativity, mental rotation, or other spatial transformation tasks such as mental folding. There was some evidence that VVIQ correlates with visual memory, but even here, the data is spotty.

IV. PARAPHENOMENA AND PHENOMENAL INTERFERENCE

Despite these results, I can not shake my commitment to the essential importance of the modality of the experiences while reading. They have a

profound affect on the way we read and the way we incorporate reading into our lives. Our experiences of an inner voice or visual imagery are central to the connected web of meaning that extends out from the influential texts: the novels and treatises, love letters and poems, emails and text messages. However, the results from these experiments suggest that our modal experiences have, at most, a negligible effect on the way we process textual information. They have no impact on cognition beyond the specialized channel used for language production and memory. We seem to have two parallel processes at play while reading — one phenomenal and the other functional. Our modal experiences while reading are existentially efficacious yet causally impotent. In this section, I spell out the implications for the relationship between phenomenal and functional theories of consciousness.

The commitment to a strong connection between experience and behavior crosses metaphysical lines. Materialists and dualists alike believe that our experiences match up with behavior in far reaching ways. For materialists, this is because both consciousness and behavior are reducible (or supervene on) the same thing: physical facts about the world. For dualists, at least in their modern, science-friendly form, there are law-like connections between consciousness and our brain-bodies (e.g. Chalmers, 1996), so again, a nomological connection is expected. Even epiphenomenalism assumes a link between experiences and behavior, and employs a series of conceptual arguments to explain why this link is not a causal one. The metaphor of a steam whistle on a train, indispensable for explaining this theory to

others, illustrates this clearly (Huxley, 1874). It is a central feature of the example that the whistle covaries with the movement of the train, and the uniquely epiphenomenal move is to deny that this is the product of causation. If experience and the physical world were to not covary, epiphenomenalism would be considerably more intuitive. Following Hume (1739/1978), in the absence of constant conjunction between two phenomena, the question of causation never even arises.

Still, in this series of studies, coarse-grained features of consciousness did not have a robust correlation with objective measures of experience. This suggests that widespread, salient aspects of experience are disconnected from behavior. They are what I call *paraphenomena*. To illustrate this concept via its metaphysical extreme, contrast epiphenomenalism, the view that there is no causal connection between consciousness and the physical world, with paraphenomenalism, the view that consciousness neither causes nor covaries with the physical world. Perhaps consciousness does not float on top of the world so much as adjacent to it. I must choose my words carefully here, because I do not wish to draw metaphysical conclusions from this experimental work. Libet (1985) provides a cautionary tail, famously arguing that free will does not exist based on a (admittedly) fascinating experiment, a conclusion that has been thoroughly debunked (Mele, 1997; 2008). Yes, paraphenomenalism is a straw man position, but a useful one for our purposes. I am not arguing that aspects of consciousness are full paraphenomena, but rather, approximations. In the case of reading, our modal experiences are

paraphenomena outside of specialized channels for language production and memory. That being said, the series of experiments in this dissertation open up a metaphysical possibility that I have never seen considered. This, in turn, exposes the implicit empirical assumptions that operate behind much of the experimental work on consciousness.

One such assumption is the idea that consciousness necessarily facilitates the widespread transmission of information. My initial hypothesis in these experiments was that the auditory experience of an inner voice would aid in tasks that involve acoustic information, such as the recollection of rhyming words, and likewise for visual imagery and the perception of the words on the page. This assumption was based on two intuitive pieces of evidence. First, (with arms outstretched for emphasis) it certainly *feels* as if our experiences have an effect on the way we act. It is hard to imagine our experiences, like the little voice in our heads, being causally impotent. Second, many influential theories in psychology place consciousness at the center of information flow and intentional action. The global workspace theory of consciousness is the most well-known and widely studied example. It holds that the main function of consciousness is the widespread dissemination of information to specialized networks (Baars, 1983, 2003b; Dehaene et al., 2006). There are certain actions for which consciousness is detrimental to performance, such as when a skilled athlete is experiencing flow or conscious entrainment, but these cases lie at the periphery. By and large, consciousness gives

a general boost to information processing, leading to a measurable improvement in performance on a variety of tasks.

The assumption that consciousness is primarily about the distribution of information is called into question by the examples of parapsychological phenomena in Wegner's (2002) work on the experience of conscious will. As Wegner argues, the experience of willing an action is neither necessary nor sufficient for voluntary control. Wegner describes instances of automation, situations where the experience of conscious will disappears, even for actions that are voluntary, purposive, and complex. Although there are numerous experimental cases that illustrate this effect, the common feeling of otherworldliness associated with a Ouija board demonstrates it clearly enough. Two people, moving a tablet around a board, consistently have an experience of automation that does not comport with the functional fact that they are masters of the outcome. He also provides examples of conscious will being an illusion, and to cite just one, Nielsen (1963) instructed participants to draw a line with one of their hands inside a viewing box. Mirrors in the box projected the image of another person's hand, and given the positioning and experimental design, participants were unaware of this deception. While the participant was drawing, on cue the imposter changed the direction of the line. The participants, unaware of the setup, thought the deviation to be their fault and adjusted the position of their own arm to compensate. Again, the feeling of voluntary control over the drawing of the line is a parapsychological phenomenon, unrelated to functional facts or information dissemination.

One way to think of parphenomena is through Block's (1995, 2007) distinction between phenomenal consciousness (P-consciousness), an experiential concept, and access consciousness (A-consciousness), a functional concept. On this framework, an experience is a parphenomena if it is P-conscious and not A-conscious. Although Block does not say as much explicitly, he takes these two concepts of consciousness to typically be co-extensive. This assumption is apparent in Block's (1995) struggle to find real-world examples that distinguish between the two. However, the present study illustrates the disconnect quite well. Block's example of A-consciousness without P-consciousness is blindsight, a rare form of cognitive impairment in which a person reports of lack of visual experience in an area of their visual field yet exhibits some ability to process visual information in that area. However, the experience of reading is even more compelling, a widespread instance of the disconnect between experience and behavior. As the studies in this dissertation show, most people have functional access to the words on the page, yet few report experiencing them. This is not a situation, like blindsight, in which performance in the area of phenomenal blindness is slightly above chance. Rather, about half of people do not report the words on the page while engaged in reading while reading comprehension and memory of textual detail remains intact.

Block's example of P-consciousness without A-consciousness is the retrospective realization that there has been a pneumatic drill outside of your window. This example is problematic, since I have been arguing at length against

such a method of armchair introspection. We should be skeptical of anecdotal reports of experience, regardless of how correct they feel. It is also not clear that the sound of the drill was not A-conscious the entire time. Failing to act on a stimulus does not imply that the stimulus was unavailable to be acted upon. Imagine that the drill stops, and I suddenly notice its absence. Why think that this auditory information was not available to guide action all along? Again, however, the modal experiences of reading provide a clearer illustration. The experience of an inner voice has no measurable effect on behavior outside of the specialized channels for language production and memory. The information is P-conscious without being widely available in the way called for by functional theories of consciousness.

Although he does not speak in terms of parphenomena, Block (2011) provides another example. Perceptual overflow occurs when participants are briefly shown an array of letters, for example, three rows of four letters. Many participants report experiencing all of the letters at once, yet they can only report three to four of them from the entire array. If we are to take these reports at face value, and this is an open question, we have a situation in which the majority of the letters are phenomenally present yet functionally inert.

Assuming something like the global workspace theory, what role do our experiences play in this functional theory of consciousness? In the experiments from this dissertation, the modality of experiences was disassociated from the way the textual information was processed. Perhaps consciousness is not as global as

we thought, since widespread, salient aspects of experience are paraphenomena. Reports of mind wandering were correlated with behavior, so the results do not suggest that consciousness is always a paraphenomena. Clearly, our experiences often play a role in cognition. The question I am asking, the question that these studies bring up, is about the gritty empirical details of the connection between experience and informational processing. In these studies, there were three instances in which modal experiences were correlated with behavior, and in each, the effects were unpredictable. For example, general reports of visual imagery were negatively correlated with performance on the Folding Task ($r = -.087$, $p = .042$) and performance on the Stroop Task ($r = -.142$, $p = .040$).

Instead of improving performance, as we would expect on the global workspace theory, visual experiences may create interference with the processing of visual and spatial information. I have been emphasizing instances of paraphenomena, but it is possible that when experiences play a functional role, sometimes they aid in information dissemination while other times they inhibit it. The examples of flow or conscious entraining would seem to be examples of this type of interference. This suggests that our experiences might not always cause a boost in performance. Perhaps we should adopt a bifurcated view. To explain through metaphor, the actors in our Cartesian theater can join the informational chorus of the global workspace, but they are also apt to sing a different song entirely. Consciousness may be at the center of our existential world, but on this view, it is not at the center of our cognitive world.

There is a surprisingly small amount of experimental research that can differentiate between the effects of functional and phenomenal consciousness. This is largely due to the general distrust of introspective reports among psychologists. For an illustrative example, consider the research on subvocalization in short-term memory. Subvocalization is the term for inner speech favored by psychologists, since it does not assume that the inner sounds form whole units of meaning, such as complete words. Early evidence for the phonological loop came from research on the phonological similarity effect, the decreased capacity of short-term memory for words that are phonetically similar. Conrad and Hull (1964) hypothesized that subvocalization was the process by which visual information was converted into acoustic information. Subsequent studies found that articulatory suppression — counting out loud or repeatedly saying a word such as “colacolacola...” — lead to decreased short-term memory for phonologically similar items (Murray, 1968; Levy, 1971; Cole & Young, 1975; Baddeley, Eldridge, & Lewis, 1981). The consensus is that subvocalization is crucial for short-term memory, and because articulatory suppression prevents subvocalization, short-term memory suffers.

Despite this research, the role of subvocalization in short-term memory is entirely unclear. Because these experiments rarely recorded introspective reports, we are left with a conditional conclusion. If articulatory suppression reduces subvocalization, then the loss of subvocalization inhibits short-term memory. However, these studies assume two things: that subvocalization always improves short-term memory, and that talking out loud prevents subvocalization. This

second assumption is called into question by Daneman and Newson (1992), who found that 10% of participants did not report a decrease in subvocalization from articulatory suppression and had no corresponding decrease in short-term memory. Further, since these studies generally avoid introspective reports, we have no idea how common subvocalization is or what its effects are when present, calling into question the first assumption. This makes the experimental evidence ambivalent towards phenomenal consciousness.

For an example of interference between experience and cognitive processing, we can turn once again to the introspective psychologists. In a fascinating series of experiments, C. W. Perky (1910) asked participants to fixate on a point on a screen in front of them and visualize certain objects, such as a tomato, a book, a leaf, a banana, an orange, or a lemon. While they were engaged in this act of visual imagination, Perky surreptitiously projected a small patch of color onto the screen. None of the participants in the experiment realized that they were experiencing percepts in addition to their visualization, and in fact, many mistook these perceptions for their visualization. Some participants even expressed surprise at imagining a vertical banana instead of the horizontal banana they were trying to visualize. Perky concluded that the difference between imagination and perception is one of degree, not of kind.

The *Perky effect*, as it was later called, has since been found to cause some bizarre results. In one study, participants were asked to imagine the New York skyline at sunset while a tomato was projected on the screen. Some of them did not

notice the tomato, but they reported imaging New York at sunset (Segal, 1972). Recent studies conclude that it is not a matter of confusing perception for imagination, but rather that the cognitive mechanism behind visual imagery interferes with the mechanisms behind visual perception (Segal & Fusella, 1971; Craver-Lemley & Reeves, 1992; Craver-Lemley & Arterberry, 2001). This stands as more evidence that, just as phenomenal consciousness can facilitate visual information processing, it can inhibit it as well.

There is another explanation for the interference between phenomenal consciousness and information processing. It could be an issue of individual differences. If consciousness has a bifurcated effect on behavior, an inner voice for one person could improve reading comprehension, and for another person, it could decrease comprehension. As this study found, there are individual differences in the prevalence of an inner voice, and as Daneman and Newson (1992) have shown, there is a range of effects from articulatory suppression. It is entirely plausible that an equal range exists in the effect of an inner voice on phonological interference. I am clearly reaching beyond the scope of the current studies. However, this is an alternative explanation for why no effect was found. Speaking autobiographically for a moment, this idea gets some support from my experience while reading. Although it does not happen often, occasionally I have an inner voice while reading, although this almost always occurs when I am having trouble concentrating on the text. The causal connection certainly feels strong, like the

inner voice is distracting me, but we can not infer it from my own introspective evidence along. It could always be that this experience is a paraphenomena.

This attempt to square phenomenal consciousness with access consciousness, to get theories of consciousness from psychology and philosophers to play with each other, is promising but underdeveloped. The extent of paraphenomena, and interference between phenomenal and functional notions of consciousness, are both empirical questions that can only be settled through brute experimental labor.

V. FURTHER RESEARCH

What is a dissertation without a future direction? I end with a series of outstanding questions and avenues for future research. Some of them are responsible, others are psychedelic. All are interesting.

1. When is armchair methodology a reliable method for descriptive phenomenology? What are the conditions under which SOME method is reliable, for that matter? I agree with the widespread criticisms of the old introspective psychologists, but I still find myself attracted to the potential power of trained introspectors. I suspect that while participants in an online study are perfectly capable of reporting coarse-grained aspects of consciousness, I believe that training and expertise are required for finer detail. The supernatural abilities of master sommeliers to describe the gustatory experiences of wine is an example of the

power of introspective training. I believe that if they can ward against introspective bias, a group of trained introspectors could be a an abundant source of knowledge.

2. What are the factors that determine the way we experience a text? These studies found an effect from the characteristics of the passage and the native language of the reader. Here is a (non-exhaustive) list of other potentials factors: purpose of reading (e.g. reading for comprehension versus proofreading); whether or not you have heard the voice of the author of a text; text viewed electronically compared to text viewed on ancient technology, such as books; the age of the reader, specifically the phenomenal differences between children and adults.

3. Are there external validity issues with experiments that perform experience sampling? For example, some of the mind wandering studies have participants advance the text manually, one word at a time. Does this type of procedure significantly impact the experiences that people have? Further, what are the effects on experience of probing for concrete reports? A sizable number of participants in these experiments reported the anticipation of the beep in their concrete reports.

4. Through training, can we learn to control our experiences while reading? I have heard of a process by which people can increase the vividness of their dreams, even gaining some measure of agency over them. Our experiences while awake are quite different than those we have while dreaming, but I have never seen someone

explore the possibility of voluntary control over their experiences while reading. Combining this idea with the conception of reading as a performance, can we improve the aesthetic value of our experiences while reading? By honing our faculty of introspection, perhaps we could make the experience of reading more meaningful or beautiful. With enough practice, could we will a phenomenal orchestra into existence? This is all a roundabout way of asking, how can we all experience the rampant, psychedelic mental imagery of E. B. Titchener?

5. How can this research be applied to pedagogy? If mind wandering is the strongest phenomenal predictor of reading comprehension, what can be done in an education context to reduce mind wandering? There is a large literature on increasing student engagement, but this paradigm makes students out to be passive actors in the educational context. What can students do to decrease their mind wandering? There is already some evidence that students who are taught mindfulness meditations have improved performance on standardized tests (Mrazek et al., 2013). Further, what can instructors do to decrease mind wandering? I take it as obvious that most professors do not work to minimize mind wandering because, if they did, teaching styles would look radically different at universities across the country. I also provide this for your amusement, a description of William James' struggle with mind wandering, as well as his work around.

... take that mind-wandering which at a later age may trouble us whilst reading or listening to a discourse. If attention be the reproduction of the sensation from within, the habit of reading not merely with the eye, and of listening not merely with the ear, but of articulating to one's self the words seen or heard, ought to deepen one's attention to the latter. Experience shows that this is the case. I can keep my wandering mind a great deal more closely upon a conversation or a lecture if I actively re-echo to myself the words than if I simply hear them; and I find a number of my students who report benefit from voluntarily adopting a similar course.” (1890, p. 281)

6. This is not the first time that there has been a failure to find an intuitive connection between modality and learning. Although it remains ubiquitous in schools across the country, the theory of “learning styles” has been repeatedly debunked (Scott, 2010; Reynolds, 1997). The enduring appeal of research on learning styles speaks to the intuitive connection between our experiences and cognitive processing. Looking at the bright side, this literature could prove to be an abundant source of research on other parphenomena.

7. It would be fascinating to apply SOME method to aesthetic experiences beyond reading. What do people experience while listening to music? Or watching movies? Or theater? Whatever the results, I suspect the experiences will be as nuanced and varied as the experiences while reading.

VI. AUTOBIOGRAPHICAL INTROSPECTION

Here we are, at the conclusion of six years of research on the things we experience while reading. To a large extent I grew up with philosophy, and started

studying it the first chance I had at the age of seventeen. In my first course I was introduced to Descartes, and I still desire the sense of certainty that I remember having after reading his *Meditations*. No matter what happens in the world, at least I have this: a definitive proof that I exist and that I can trust the content of my experiences. As such, I would like for this dissertation to end with some sort of certainty. Admittedly, I do not think that it does. Do you feel more confident in your experiences while reading?

In a way, the experience of reading constitutes the perfect phenomenal storm. It is variable, the external world gives us little guidance to the internal world, and because it is a minor issue in philosophy and psychology, people have used it as a battle ground for their own theoretical purposes. I hope that, in learning how to navigate the various issues, you have gained some insight into your own experiences.

As you can see, I am ending this dissertation on an autobiographical note. I am not a visual imager or an inner voice hearer. I rarely have these experiences. I used to think I was alone, but now I know that this is not the case. I am a member of small group, the 10% of readers who do not share in these experiences while reading. Now that I have had (literally) thousands of people describe to me their experiences, how would I describe my own? Let me give it a shot. When I read, it feels like I have a direct and unmediated access to the information contained within the text. I have long thought of it in a Platonic way, as if I come to know the form of the text, un-muddied by images or distracting sounds. It might be

something like your experience of hearing an update on Basque sports: Irujo beat Olaizola in the Basque Esku Pilota championship. You may not have any idea what Esku Pilota is, or what the individuals (or are they teams?) look like. But you can still have an understanding that something or other named “Ijuro” emerged victorious over something or other names “Olaizola” in some kind of contest. My experiences are rather like those of George Berkeley in this regard, who claims that he has no visual imagery when he reads “good thing” or “Aristotle” (1710/2009, Intro Section 20). Sometimes the process of reading is beautiful, sometimes it is a boor. Sometimes my mind is focused, sometimes it drifts. Regardless, it goes about processing the text without much in the way of modal experiences.

I do not know what to say about the words on the page. Here I am stumped, and I feel hundreds of years away from Descartes and in an entirely different epistemic world. The words certainly are not a salient part of my experience, but when I say that I do not experience them, the claim does not ring true. I bristle. Maybe I am committing a fallacy, reading facts about the world into my own experiences. However, I consider myself an expert on consciousness. I have now written a dissertation on the topic. I have read hundreds of articles and dozens of books about it. I have bracketed presuppositions. I have done everything right. Still, the answer eludes me.

REFERENCES

- Abrams, R. A., & Jonides, J. (1988). Programming saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 428-443.
- Abrams, R. A., Meyer, D. E., & Kornblum, S. (1989). Speed and accuracy of saccadic eye movements: Characteristics of impulse variability in the oculomotor system. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 529-543.
- Albee, E. (1962/1983). *Who's Afraid of Virginia Woolf*. Signet / New American Library.
- Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 85, 249-277.
- Anderson, J. R. (1979). Further arguments concerning representations for mental imagery: A response to Hayes-Roth and Pylyshyn. *Psychological Review*, 86, 395-406.
- Armstrong, D. M. (1968). *A materialist theory of the mind*. London: Routledge.
- Augustine of Hippo (4th c. CE/1991). *Confessions*. (H. Chadwick, Trans.). Oxford University Press.
- Austin, J. L. (1962). *How to do things with words* (2nd ed.). M. Sbisà and J. O. Urmson (Eds.). Oxford: Oxford University Press.
- Austin, M. J. (1922). Plato as a writer of imaginary conversations. *The Classical Journal*, 17(5), 243-255.
- Ayer, A. J. (1936/1971). *Language, truth, and logic*. Penguin Books.
- Ayer, A. J. (1940). *Foundations of empirical knowledge*. London: Macmillan.
- Baars, B. J. (1983). Conscious contents provide the nervous system with coherent, global information. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness & self-regulation* (41-79). New York, NY: Plenum Press.
- Baars, B. J. (2003a). How brain reveals mind. *Journal of Consciousness Studies*, 10(9-10), 100-114.

- Baars, B. J. (2003b). The global brainweb: An update on global workspace theory. *Science and Consciousness Review*, 2.
- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *The Quarterly Journal of Experimental Psychology*, 18(4), 362-365.
- Baddeley, A. D. (1986). *Working Memory*. Oxford University Press.
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Review Neuroscience*, 4(10), 829-839.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford University Press.
- Baddeley, A. D., Eldridge, M., & Lewis, V. (1981). The role of subvocalisation in reading. *The Quarterly Journal of Experimental Psychology*, 33(4), 439-454.
- Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158-173.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation*, 8, 47-89.
- Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, 36(2), 233-252.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 575-589.
- Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of personality and Social Psychology*, 65(2), 272.
- Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, 102(1), 146-184.
- Berkeley, G. (1710/1965). A treatise concerning the principles of human knowledge. In C. Turbayne (Ed.), *Principles, dialogues, and philosophical correspondence*. Macmillan.
- Betts, G. H. (1909). *The distribution and functions of mental imagery*. Teachers College: Columbia University.

- Blackmore, S. (2002). There Is No Stream of Consciousness. What is all this? What is all this stuff around me; this stream of experiences that I seem to be having all the time?. *Journal of Consciousness Studies*, 9(5-6), 17-28.
- Block, N. (1990). Inverted earth. *Philosophical Perspectives*, 4, 53-79.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227-287.
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30(5-6), 481-499.
- Block, N. (2009). Comparing the major theories of consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences IV* (1111-1123). MIT Press.
- Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, 15(12), 567-575.
- BonJour, L. (2000). Toward a defense of empirical foundationalism. In M. DePaul (Ed.), *Resurrecting old-fashioned foundationalism* (21-38). Lanham, MA: Rowman and Littlefield.
- Borges, G. L. (1975/1999). The disk. In *Collected fictions* (477) (A. Hurley, Trans.). Penguin Books.
- Borges, G. L. (1975/1999). Dreamtigers. In *Collected fictions* (294) (A. Hurley, Trans.). Penguin Books.
- Brooks, L. R. (1967). The suppression of visualization by reading. *The Quarterly Journal of Experimental Psychology*, 19(4), 289-299.
- Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 22(5), 349.
- Burke, E. (1775/1990). *A philosophical enquiry into the origin of our ideas of the sublime and beautiful*. Oxford: Oxford University Press.
- Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (275-307). New York: Academic Press.
- Carstensen, L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., ... & Nesselroade, J. R. (2011). Emotional experience improves with age: evidence based on over 10 years of experience sampling. *Psychology and Aging*, 26(1), 21.

- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford, England: Oxford University Press.
- Chalmers, D. J. (2004). How can we construct a science of consciousness? In M. Gazzaniga (Ed.), *The Cognitive Neurosciences III* (1111-1119). MIT Press.
- Cole, R. A., & Young, M. (1975). Effect of subvocalization on memory for speech sounds. *Journal of Experimental Psychology: Human Learning and Memory*, 1(6), 772.
- Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior*, 19(6), 722-735.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100(4), 589-608.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256.
- Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, 55(1), 75-84.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, 55(4), 429-432.
- Cova, F., Garcia, A., & Liao, S. Y. (2015). Experimental Philosophy of Aesthetics. *Philosophy Compass*, 10(12), 927-939.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 97-185.
- Craver-Lemley, C., & Arterberry, M. E. (2001). Imagery-induced interference on a visual detection task. *Spatial Vision*, 14(2), 101-119.
- Craver-Lemley, C., & Reeves, A. (1992). How visual imagery interferes with vision. *Psychological Review*, 99(4), 633-649.
- Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. Harper Collins.

- Csikszentmihályi, M. & Csikszentmihályi, I. S. (1988). *Optimal experience: psychological studies in flow consciousness*. Cambridge University Press.
- Csikszentmihályi, M., & Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of Happiness Studies*, 4(2), 185-199.
- Cummings, E. E. (1994). Here is Effie's head. In *Selected Poems* (113). New York: Liveright.
- Cummings, E. E. (1994). The Noster was a ship of swank. In *Selected Poems* (110). New York: Liveright.
- Cummins, R. (1998). Reflection on reflective equilibrium. In M. DePaul & W. Ramsey (Eds.), *Rethinking Intuition: The psychology of intuition and its role in philosophical inquiry* (113–128). Rowman & Littlefield.
- Dale, H. C. A. (1973). Short-term memory for visual information. *British Journal of Psychology*, 64(1), 1-8.
- Daneman, M., & Newson, M. (1992). Assessing the importance of subvocalization during normal silent reading. *Reading and Writing*, 4(1), 55-77.
- de Vega, M. (1995). Backward updating of mental models during continuous reading of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 373-385.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17-22.
- DeGrazia, D. (2005). *Human identity and bioethics*. Cambridge: Cambridge University Press.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, 10(5), 204-211.
- Delillo, D. (1985). *White noise*. Penguin Books.
- Della Sala, S., & Logie, R. H. (2002). Neuropsychological impairments of visual and spatial working memory. *Handbook of memory disorders*, 2, 271-292.
- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37(10), 1189-1199.

- Dennett, D. C. (1991). *Consciousness explained*. USA: Back Bay Books.
- Dennis, I., & Newstead, S. E. (1981). Is phonological recoding under strategic control?. *Memory & Cognition*, 9(5), 472-477.
- Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), 5-18.
- Devine, P. & Monteith M. (1999). Automaticity and control in stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (339–360). The Guilford Press.
- Dixon, P., & Bortolussi, M. (2013). Construction, integration, and mind wandering in reading. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(1), 1-10.
- Dovidio, J., & Gaertner, S. (2004). Aversive racism. *Advances in Experimental Social Psychology*, 36, 1-51.
- Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory*, 6(3), 319-332.
- Dulin, D., Hatwell, Y., Pylyshyn, Z., & Chokron, S. (2008). Effects of peripheral and central visual impairment on mental imagery capacity. *Neuroscience & Biobehavioral Reviews*, 32(8), 1396-1408.
- Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior*, 21(3), 296-306.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. (1976). *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.
- Ernest, C. H. (1977). Imagery ability and cognition: A critical review. *Journal of Mental Imagery*, 2, 181-216.
- Everatt, J., & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, and eye movement control. *Language and Speech*, 37(3), 283-297.

- Everatt, J., Bradshaw, M. E., & Hibbard, P. B. (1998). Individual differences in reading and eye movement control. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (223-242). Oxford, England: Elsevier.
- Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *The Quarterly Journal of Experimental Psychology*, 38(4), 675-688.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, 69(6), 1013-1027.
- Folk, J. R. (1999). Phonological codes are used to access the lexicon during silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 892-906.
- Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68, 5-20.
- Frankfurt, H. (1999). On caring. In *Necessity, volition, and love* (155–180). New York: Cambridge University Press.
- 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.
- Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*, 66(12), 2289-2294.
- Franklin, M. S., Mooneyham, B. W., Baird, B., & Schooler, J. W. (2014). Thinking one thing, saying another: The behavioral correlates of mind-wandering while reading aloud. *Psychonomic Bulletin & Review*, 21(1), 205-210.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: true issues and false trails. *Psychological bulletin*, 123(1), 71-99.
- Frost, R. L. (1927/1969). The egg and the machine. In *The Poetry of Robert Frost* (269). Henry Holt and Company.
- Galton, F. (1880). Statistics of mental imagery. *Mind*, (19), 301-318.

- Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology*, 16(1), 35-50.
- Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, 81(4), 439-454.
- Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 84.
- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*, 28(5), 887-898.
- Giambra, L. M., & Grodsky, A. (1989). Task-unrelated images and thoughts while reading. In J. Shorr, P. Robin, J. A. Connella, & M. Wolpin (Eds.), *Imagery: Current perspectives* (26-31). New York: Plenum Press.
- Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26(1), 69-83.
- Goldman, A. I. (2006). *Simulating minds*. Oxford: Oxford.
- Gordon, R. (1949). An investigation into some of the factors that favour the formation of stereotyped images. *British Journal of Psychology*, 39(3), 156-167.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Green, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Historical and contemporary readings* (359-372). MIT Press.
- Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in cognitive sciences*, 6(12), 517-523.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4-27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, 74(6), 1464-1480.

- Grodsky, A., & Giambra, L. M. (1990). The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts. *Imagination, Cognition and Personality*, 10(1), 39-52.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814-834.
- Haidt, J., Bjorklund, F., & Murphy, S. (manuscript). Moral dumbfounding: When intuition finds no reason.
- Haidt, J., Rozin, P., McCauley, C., & Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology & Developing Societies*, 9(1), 107-131.
- Hayes-Roth, F. (1979). Distinguishing theories of mental representation: A critique of anderson's 'Arguments concerning mental imagery'. *Psychological Review*, 86, 376-382.
- Hektner, J. M., Schmidt, J. A., & Csíkszentmihályi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Chicago: Sage.
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14(3), 399-425.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, S., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1217-1232.
- Hume, D. (1739/1978). *A treatise of human nature*. Oxford: Oxford.
- Hurlburt, R. T., & Heavey, C. L. (2001). Telling what we know: describing inner experience. *Trends in cognitive sciences*, 5(9), 400-403.
- Hurlburt, R., & Heavey, C. L. (2004). To beep or not to beep: Obtaining accurate reports about awareness. *Journal of Consciousness Studies*, 11(7-8), 113-128.
- Hurlburt, R. T., & Heavey, C. L. (2006). *Exploring inner experience: The descriptive experience sampling method*. John Benjamins Publishing Company.

- Hurlburt, R.T. & E. Schwitzgebel (2007). *Describing inner experience: proponent meets skeptic*. MIT Press.
- Huxley, T. H. (1874/1898). On the Hypothesis that Animals are Automata, and its History. Reprinted in *Method and results: Essays by Thomas H. Huxley*. New York: D. Appleton and Company,
- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, 32, 127-136.
- Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83(5), 291-295.
- Jacobson, J. Z., & Dodwell, P. C. (1979). Saccadic eye movements during reading. *Brain and Language*, 8(3), 303-314.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt.
- Jaynes, J. (1976/2000). *The origin of consciousness and the breakdown of the bicameral mind*. Boston: Houghton Mifflin.
- Jennings, C. D. (2015). Consciousness without attention. *Journal of the American Philosophical Association*, 1, 276-295.
- 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.
- Rayner, K., Schotter, E. R., Masson, M. E., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help?. *Psychological Science in the Public Interest*, 17(1), 4-34.
- Kelly, S. D. (2008). Content and constancy: phenomenology, psychology, and the content of perception. *Philosophy and Phenomenological Research*, 76(3), 682-690.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006), 932-932.
- Kivy, P. (2006). *The performance of reading*. Blackwell Publishing.
- Kopp, K., D'Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness and Cognition*, 34, 52-62.

- Korta, K. & Perry, J. (2011). *Critical pragmatics*. Cambridge University Press.
- Kosslyn, S. M. (1980). *Image and mind*. Harvard University Press.
- Kosslyn, S. M. (1994). *Image and brain*. MIT Press.
- Lawrence, B. M., Myerson, J., Oonk, H. M., & Abrams, H. M. (2001). The effects of eye and limb movements on working memory. *Memory*, 9(4-6), 433-444.
- Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 285.
- Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 10(2), 123-132.
- Lévy-Schoen, A. (1981). Flexible and/or rigid control of oculomotor scanning behavior. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (299-316). Hillsdale, NJ: Erlbaum.
- Lewis, K. J. S., Borst, G., & Kosslyn, S. M. (2011). Integrating images and percepts: New evidence for depictive representation. *Psychological Research*, 75, 259-271.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529-566.
- Locke, J. (1690/1975). *An essay concerning human understanding*. Oxford: Clarendon.
- Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory & Cognition*, 24(3), 305-321.
- Lukatela, G., & Turvey, M. T. (1991). Phonological access of the lexicon: Evidence from associative priming with pseudohomophones. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 951-966.
- Lycan, W. G. (1996). *Consciousness and experience*. Cambridge, MA: MIT Press.
- Lycan, W. G. (2004). The superiority of HOP to HOT. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness* (93-113). Philadelphia: John Benjamins Publishing Company.

- Machery, E. (manuscript). De-Freuding Implicit Attitudes.
- MacIntyre, A. (1989). The virtues, the unity of a human life and the concept of a tradition. In S. Hauerwas and L. G. Jones (Eds.), *Why Narrative?*. Grand Rapids, MI: W.B. Eerdmans.
- Mack, A., & Rock, I. (1998). *Inattentional Blindness*. MIT Press.
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64, 17-24.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.
- Matin, E. (1974). Saccadic suppression: A review. *Psychological Bulletin*, 81, 899-917.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578-586.
- McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery*, 19, 1-106.
- Mele, A. R. (1997). Strength of motivation and being in control: Learning from Libet. *American Philosophical Quarterly*, 34(3), 319-332.
- Mele, A. R. (2008). Recent work on free will and science. *American Philosophical Quarterly*, 45(2), 107-130.
- Mencius (4th c. BCE/1895). Book 4B. In Legge, J. (Ed.), *The works of Mencius*. Clarendon Press.
- Mill, J. S. (1861/1998). *Utilitarianism*. New York: Oxford University Press.
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: a working model. *Neuroreport*, 14(3), 299-305.
- Morin, A. (2009). Inner Speech and Consciousness. In W. P. Banks (Ed.), *Encyclopedia of Consciousness, vol. 1* (389-402). Oxford: Elsevier.
- Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, 28(3), 292-312.

- Morrow, D. G., Leirer, V., Altieri, P., & Fitzsimmons, C. (1994). Age differences in creating spatial models from narratives. *Language and Cognitive Processes*, 9(2), 203-220.
- Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, 24(5), 776-781.
- Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, 78(4p1), 679-684.
- Nakamura, J. & Csíkszentmihályi, M. (2002). The concept of flow. In C.R. Snyder & S.J. Lopez (Eds.), *Handbook of positive psychology* (89-105). New York: Oxford University Press.
- Nichols, S., & Stich, S.P. (2003). *Mindreading*. Oxford: Oxford.
- Nielsen, T. I. (1963). Volition: A new experimental approach. *Scandinavian journal of psychology*, 4(1), 225-230.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT.
- Oh, S. H., & Kim, M. S. (2004). The role of spatial working memory in visual search efficiency. *Psychonomic Bulletin & Review*, 11(2), 275-281.
- Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research*, 53(1), 13-24.
- Paivio, A. (1969). Mental imagery in associative learning. *Psychological Review*, 76(3), 241-263.
- Paivio, A. (1971). Imagery and deep structure in the recall of English nominalizations. *Journal of Verbal Learning and Verbal Behavior*, 10(1), 1-12.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, 30(3), 331-347.

- Payne, B. K. (2006). Weapon bias split-second decisions and unintended stereotyping. *Current Directions in Psychological Science*, 15(6), 287-291.
- Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, 30(4), 473-485.
- Perky, C. W. (1910). An experimental study of imagination. *The American Journal of Psychology*, 21(3), 422-452.
- Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58(3), 193.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, 16(2), 283-290.
- Piccinini, G. (2003). Data from introspective reports: Upgrading from common sense to science. *Journal of Consciousness Studies*, 10(9-10), 141-156.
- Poe, E. A. (1849/2008). A dream within a dream. In *The complete Poetry of Edgar Allan Poe* (39). New American Library.
- Price, H. H. (1932). *Perception*. London: Methuen.
- Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80, 1-24.
- Pylyshyn, Z. W. (1979). Validating computational models: A critique of anderson's indeterminacy of representation claim. *Psychological Review*, 86, 383-394.
- Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, 25, 157-182.
- Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology*, 38(4), 689-703.
- Rayner, K. (1978). Eye movement latencies for parafoveally presented words. *Bulletin of the Psychonomic Society*, 11, 13-16.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, 206, 468-469.

- 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.
- Reisberg, D. (1997). *Cognition: Exploring the science of the mind*. New York: Norton.
- Reisberg, D., Culver, L. C., Heuer, F., & Fischman, D. (1986). Visual memory: When imagery vividness makes a difference. *Journal of Mental Imagery*, 10(4), 51-74.
- De Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex*, 11(4), 341-354.
- Reynolds, M. (1997). Learning styles: a critique. *Management learning*, 28(2), 115-133.
- Richardson, J. T. E. (1980). *Mental imagery and human memory*. St. Martin's Press.
- Rinck, M., Williams, E., Bower, G. H., & Becket, E. S. (1996). Spatial situation models and narrative understanding: Some generalizations and extensions. *Discourse Processes*, 21, 23-55.
- Robinson, W. S. (2005). Thoughts without distinctive non-imagistic phenomenology. *Philosophy and Phenomenological Research*, 70(3), 534-562.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, 21(4), 803-814.
- Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, 49, 329-359.
- Rosenthal, D. M. (1993). Thinking that one thinks. In M. Davies and G. W. Humphreys (Eds.), *Consciousness: Psychological and philosophical essays* (197-223). Blackwell Publishing.
- Rosenthal, D. M. (2002). Explaining consciousness. In D. Chalmers (Ed.), *Philosophy of mind: Classical and contemporary readings* (406-421). Oxford University Press.

- Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford: Oxford University Press.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis, 2nd edition*. McGraw-Hill.
- Saenger, P. (1997). *Space between words: the origins of silent reading*. Stanford: Stanford University Press.
- Salthouse, T. A., & Ellis, C. L. (1980). Determinants of eye-fixation duration. *American Journal of Psychology*, 93, 207-234.
- Salthouse, T. A., Ellis, C. L., Diener, D. C., & Somberg, B. L. (1981). Stimulus processing during eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 611-623.
- Salway, A. F., & Logie, R. H. (1995). Visuospatial working memory, movement control and executive demands. *British Journal of Psychology*, 86(2), 253-269.
- 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.
- Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metacognition. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (203-226). Massachusetts Institute of Technology.
- Schwitzgebel, E. (2004). Introspective training apprehensively defended: Reflections on Titchener's lab manual. *Journal of Consciousness Studies*, 11(7-8), 58-76.
- Schwitzgebel, E. (2005). Difference tone training: A demonstration adapted from Titchener's Experimental Psychology. *Psyche*, 11(6).
- Schwitzgebel, E. (2011). *Perplexities of Consciousness*. MIT Press.
- Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, 4(1), 5-34.
- Scott, C. (2010). The enduring appeal of 'learning styles'. *Australian Journal of Education*, 54(1), 5-17.

- Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.
- Segal, S. J. (1972). Assimilation of a stimulus in the construction of an image: The Perky effect revisited. In P. W. Sheehan (Ed.), *The Function and Nature of Imagery* (203-230). New York & London: Academic Press.
- Segal, S. J., & Fusella, V. (1971). Effect of images in six sense modalities on detection of visual signal from noise. *Psychonomic Science*, 24(2), 55-56.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. In D. A. Balota and E. J. Marsh (Eds.), *Cognitive psychology: Key readings*. Psychology Press.
- Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, 3(2), 228-243.
- Shoemaker, S (1982). The inverted spectrum. *Journal of Philosophy*, 79, 357-81.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton: Princeton University Press.
- Siewert, C. (2006). Is the appearance of shape protean? *Psyche*, 12(3), 1-16.
- Siewert, C. (2011). Socratic introspection and the abundance of experience. *Journal of Consciousness Studies*, 18(1), 63-91.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentive blindness for dynamic events. *Perception*, 28(9), 1059-1074.
- Singer, M., Graesser, A. C., & Trabasso, T. (1994). Minimal or global inference during reading. *Journal of Memory and Language*, 33(4), 421-441.
- Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging inferences. *Cognitive Psychology*, 30, 1-38.
- Singer P. (1993). *Practical ethics, 2nd edition*. Cambridge: Cambridge University Press.
- 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., 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., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, 13(4), 657-690.
- Smyth, M. M., Pearson, N. A., & Pendleton, L. R. (1988). Movement and working memory: Patterns and positions in space. *The Quarterly Journal of Experimental Psychology*, 40(3), 497-514.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-661.
- Taylor, C. (1989). *Sources of the self: The making of modern identity*. Cambridge, MA: Harvard University Press.
- Thomas, N. J. (1999). Are theories of imagery theories of imagination?: An active perception approach to conscious mental content. *Cognitive science*, 23(2), 207-245.
- Thomas, Nigel J. T. (2014). Mental Imagery. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition). URL = <<http://plato.stanford.edu/archives/fall2014/entries/mental-imagery/>>.
- Thorndike, E. L., & Lorge, I. (1944). *Tire teacher's word book of 30,000 words*. Teachers' College, NY: Columbia University.
- Titchener E. B. (1909). *Lectures on the experimental psychology of the thought-process*. New York: Macmillan.
- Tye, M. (2000). *Consciousness, color, and content*. MIT Press.
- Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 832.
- Woodman, G. F., Vogel, E. K., & Luck, S. J. (2001). Visual search remains efficient when visual working memory is full. *Psychological Science*, 12(3), 219-224.
- Valentin, G. G. (1844). *Lehrbuch der Physiologie des Menschen, 2nd edition*. Brunswick.
- Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition*, 15, 181-198.

- Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 371.
- Vellmans, M. (2009). *Understanding consciousness, 2nd edition*. Routledge.
- Watson, J. B. (1928). *The ways of behaviorism*. New York: Harper.
- Wegner, D. M. (2002). *The illusion of conscious will*. MIT Press.
- Wegner, D. M. & Wheatley, T. P. (1999). Apparent mental causation: sources of the experience of will. *American Psychologist*, 54, 480-492.
- Wheatley, T. (2009). Everyday confabulation. In W. Hirstein (Ed.), *Confabulation: Views from neuroscience, Psychiatry, Psychology and Philosophy*. Oxford.
- Wittgenstein, L. (1953/2009). *Philosophical Investigations* (G. E. M. Anscombe, P. M. S. Hacker, & J. Shulte, Trans.). Blackwell Publishing.
- Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review*, 11(2), 269-274.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, 123(2), 162.