Distribution Agreement

Signature:

Robert Thorstad	10/11/2019

5 Million Thoughts about the Past and Future Reveal Shared Reliance on Schemas

By

Robert Thorstad
Doctor of Philosophy
Psychology

Phillip Wolff
Advisor

Patricia Bauer
Committee Member

Robyn Fivush
Committee Member

Scott Lilienfeld
Committee Member

Stella Lourenco
Committee Member

Accepted:

Lisa Tedesco, PhD
Dean of the James T. Laney School of Graduate Studies

Date

5 Million Thoughts about the Past and Future Reveal Shared Reliance on Schemas

By

Robert Thorstad
BA, Haverford College, 2012
MA, Emory University, 2017

Advisor: Phillip Wolff, PhD

An abstract of
a dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2019

Abstract

5 Million Thoughts about the Past and Future Reveal Shared Reliance on Schemas
By Robert Thorstad

People frequently think about the future, yet little is known about the cognitive processes people use to do so. Recent research suggests that future thinking may rely on many of the same cognitive processes people use to mentally time travel into the past. I tested this proposal by automatically identifying over 5 million temporal references in web blog posts and then using linguistic markers to identify the cognitive processes that people used to generate these temporal references. I identified cognitive processes uniquely associated with mental time travel by comparing these linguistic markers in past and future references to talk about the present, which does not involve mental time travel. In Study 1, I found that talk about the past relies on more episodic language than talk about the future, but relies on equal amounts of episodic language as talk about the present. This result suggests that episodic processing is not uniquely associated with mental time travel. In Study 2, I found that talk about both the past and future relies more on schemas than talk about the present. This result suggests that the use of schemas is uniquely associated with mental time travel. In Study 3, I replicated these results using temporal thoughts evoked in the lab. Together, these results suggest that past and future thinking rely on a common cognitive process but on a different process than was previously believed: the use of schemas, not episodic processing.

5 Million Thoughts about the Past and Future Reveal Shared Reliance on Schemas

By

Robert Thorstad
BA, Haverford College, 2008
MA, Emory University, 2015

Advisor: Phillip Wolff, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2019

Acknowledgments

Thanks first to my committee: Patricia Bauer, Robyn Fivush, Scott Lilienfeld, and Stella Lourenco. I am extremely blessed to have had such an all-star dissertation committee. Thanks for taking this journey with me, and for be willing to follow as I developed a number of novel (dare I say unusual) methods in my research. I feel a tinge of guilt that I did not take even more advantage of your expertise – but know that it was incredibly valuable to me to have each and every one of you on my committee.

Thanks to my advisor, Phil Wolff. Thanks for taking a chance on me, a philosophy student with very little background in psychology. Not only were you not turned off by my background, but you openly welcomed it. Thanks also for giving me an incredible amount of creative freedom in my research. The idea that mining tweets and training intelligent machines could count as doing psychology could have seemed impossibly far-fetched, but instead you welcomed it and allowed me to unleash these tools in my research. Know that my whole view of science – what doing science is, what kind of science is possible, and what place I can take in it – has changed tremendously from working with you. From a philosopher interested in causality to a cognitive scientist interested in machine learning and data science, my entire conception of science and my place in it has changed dramatically and for the better.

Thanks to my family. Education has always been important to you, and you've consistently surrounded me with incredible teachers. That I am here today finishing my PhD only testifies to this. Thanks also for the incredible amount of trust that you have

always shown me. It can't have been easy to listen to your recent college graduate explain that, with no prior experience in psychology, I planned to pursue a PhD in psychology. You've always given me incredible freedom and trust to explore the path I want to take with my life. I'm not sure any of us would have imagined the path I am on today, but I'm very glad that I'm here.

Thanks finally to the members of the Wolff lab. Thanks to Jason Shepard. You were not just a friend but a role model from day one. I could always turn to you with even the most difficult of questions, and you always took the time to patiently work through them with me. Thanks to Jacquelyn Ellison. It's been a joy to have you in the lab and I've valued tremendously both your friendship and our many intellectual conversations in the lab.

Table of Contents

Chapter 1: Introduction 1

1.1 Evidence Supporting Common Processes for Past and Future Thinking 4

1.1.1. Neuroscientific Evidence for Common Processing of Past and Future Thoughts 4

1.1.2 Neuropsychological Evidence for Common Processing of Past and Future Thoughts 5

1.1.3 Behavioral Evidence for Common Processing of Past and Future Thoughts 6

1.2 Evidence Suggesting Differences in Processing Between Past and Future Thinking 8

1.2.1 Many Patients with Episodic Memory Deficits Have Intact Future Thinking 9

1.2.2 Much of the Evidence for Shared Processing Lacks an Appropriate Negative Control 10

1.2.3 Some Brain Regions Respond More Strongly to Future Thoughts than Past Thoughts 11

1.3 Outline of the Remainder of the Dissertation 12

1.4 Summary of Chapter 1 13

Chapter 2: Episodic Past and Future Thinking 15

2.1 Introduction 15

2.2 Approach 16

2.2.1 A Source for Naturally Occurring Temporal Thoughts: The Blog Authorship Corpus 17

2.2.2 A Method for Extracting Naturally Occurring Temporal Thoughts: The Copley & Wolff Temporal Reference Classifier 20

2.2.3 A Method for Measuring Episodic Past and Future Thinking using Language 23

2.3 Methods 27
2.4. Results 32
2.4.1 Performance of Temporal Reference Classifiers 33
2.4.2 Episodic Language in Temporal Thoughts 36
2.5. Discussion 38
Chapter 3: Semantic Past and Future Thinking 41
3.1. Introduction 41
3.1.1 Evidence for Schema Usage in Past Thinking 42
3.1.2 Evidence for Schema Usage in Future Thinking 43
3.1.3 Absence of Evidence about Similarity of Schema Usage in Past and Future Thinking 45
3.2 Approach 46
3.2.1 A Computational Model of Schemas using Topic Modeling 47
3.2.2 Learning Common Schemas: a Corpus Approach 50
3.2.3. Applying the Learned Schemas: Using Topic Models to Infer Schema Usage as a Cognitive Process 53
3.3 Methods 56
3.4 Results 60
3.4.1 Evaluation of Topic Model: Semantic Coherence 61
3.4.2 Evaluation of Topic Model: Consistency of Schemas across Time and Corpora 63
3.4.3 Evaluation of Method for Detecting Application of Schemas 65
3.4.4 Schema Usage in Thoughts about the Past, Present, and Future. 67
3.5. Discussion 69
Chapter 4: Temporal Thoughts in the Lab 74
4.1 Introduction 74

4.2 Approach ... 74

4.2.1 Eliciting Temporal Thoughts in the Lab: the Future Crovitz Task ... 74

4.2.2. Measuring Episodic and Semantic Temporal Thinking ... 76

4.3 Methods ... 77

4.4 Results ... 79

4.4.1 Episodic Processing in Lab-Elicited Temporal Thoughts ... 79

4.4.2 Semantic Processing in Lab-Elicited Temporal Thoughts ... 81

4.4.3 Lab-Based Topic Model ... 82

4.5 Discussion ... 82

Chapter 5: General Discussion ... 85

5.1 Summary of Findings ... 85

5.2 Implications of Big Data for Cognitive Science ... 86

5.2.1 Learning Common Human Schemas ... 87

5.2.2 Identifying Naturally Occurring Temporal Thoughts ... 88

5.3 Limitations and Future Directions ... 90

5.3.1 Limitations of Topic Models as a Measure of Schemas ... 91

5.3.2 Limitations of Measuring Episodic Processing Using Language ... 92

5.3.3. Alternative Interpretation of the Role of Schemas in Past and Future Thinking ... 94

5.3.4 Limitations of the Temporal Orientation Classifier ... 95

5.4 Conclusion ... 96

References ... 97

Appendix 1: Formal Description of Latent Dirichlet Allocation ... 115

Appendix 2: List of Topics Learned by Topic Model ... 117

Appendix 3: Instructions for Future Crovitz Task 130

Appendix 4: Lab-Based Topic Model.. 132

Chapter 1: Introduction

How do people think about the future? One possibility is that thoughts about the future involve largely the same mental processes that are used think about the past. In particular, thoughts about the future might involve mental simulation processes associated with the generation of episodic detail (Atance & O'Neill, 2001; Klein, Loftus, & Kihlstrom, 2002; Schacter, Addis, & Buckner, 2008). This proposal is supported by studies finding significant overlap in the content and brain areas involved in the processing of past and future thinking (Benoit & Schacter, 2015; Okuda et al., 2003; Spreng, Mar, & Kim, 2009). While there is significant empirical support for shared mental processes, several questions remain. One key issue concerns whether the processes that have been attributed to past and future thinking are, in fact, uniquely associated with these two kinds of mental time travel. The problem emerged in the interpretation of a recent study. Addis, Wong, and Schacter (2008) found that when older adults described both past and future events, the older adults relied on fewer episodic details than did younger adults. Because of age-related declines in hippocampal functioning, these results were interpreted to suggest that past and future thinking may rely on common episodic processes supported by the hippocampus. However, Schacter, Gaesser, and Addis (2013) asked participants to perform the same task but also asked older adults to perform a control task that did not involve mental time travel: describing a picture in front of them. Surprisingly, the same decline in episodic details was observed when older adults simply described a picture in front of them, despite the fact that describing a picture does not involve mental time travel. This result suggests that the decline in episodic details may not be uniquely associated with past and future thinking, but instead could extend to language production more generally. This result also demonstrates a broader logical problem with existing evidence for common processing. The logical problem is that demonstrating shared processing between past and future thinking involves identifying a process that is not just shared between past and future thinking, but is also uniquely associated with past and future thinking (for a similar argument, see Hassabis, Kumaran, & Maguire, 2007a; Hassabis, Kumaran, Vann, & Maguire, 2007b). Establishing this unique association requires the use of control conditions involving neither past nor future thinking, a condition often absent in prior literature.

This dissertation offers a new proposal on what mental processes are shared between past and future thinking, namely the use of schemas. The proposal is supported by a long history of research showing that people use schemas to think about the past. This proposal is also supported by the intuition that schemas might be especially relevant to thinking about the future because in the case of the future, people cannot rely on accessing stored information. Importantly, this later source of support remains based on intuition due to a lack of research investigating the role of schemas in future thinking.

This dissertation reports research that not only tests for the presence of mental processes in cognitions where they are expected, but also tests for their absence in cognitions where they are not expected. To investigate the mental processes uniquely associated with past and future thinking, I will test for mental processes that are not only present in past and future thinking, but also absent in thinking that is not associated with mental time travel: thoughts about the present. I predict that if schemas are uniquely associated with past and future thinking, then schemas will have a stronger impact on the way people think about the past and future than on the way people think about the present. This logic will also allow me to conduct a more stringent test than in previous accounts of the similarity between future and past thinking. In particular, if episodic thinking is a process uniquely associated with past and future thinking, then such thinking should not only be present when people engage in past and future thinking, but also should be largely absent in thinking that does not involve mental time travel, namely thinking about the present.

In effect, then, the dissertation will answer three questions about how people think about the past and future:

Question 1: Is episodic processing uniquely associated with past and future thinking compared to thinking about the present?

Question 2: Does thinking about the future involve schemas?

Question 3: Is the use of schemas uniquely associated with past and future thinking compared to thinking about the present?

In addition, the dissertation will address these questions using a unique kind of methodology. Thoughts about time are typically elicited in the lab, but people think about time many times every day. Capturing these natural thoughts about time can not only yield a large number of temporal thoughts, but can also avoid the potential limitations associated with eliciting these thoughts through explicit prompting. For these reasons, this dissertation will develop several machine learning and natural language processing techniques for automatically identifying and analyzing over five million naturally occurring thoughts about time.

1.1 Evidence Supporting Common Processes for Past and Future Thinking

The proposal that past and future thinking rely on common cognitive processes is supported by three types of evidence.

1.1.1. Neuroscientific Evidence for Common Processing of Past and Future Thoughts

The first type of evidence for common processing is that when people remember the past or imagine the future in an fMRI scanner, a common set of brain regions is engaged by both types of mental time travel. These common regions involve a broad set of broad set of frontal, parietal, and temporal lobe regions from the *default network* (Raichle et al., 2001). Such a pattern of activation has been observed in several individual studies. Szpunar, Watson, and McDermott (2007) found that regions in the posterior cingulate cortex, parahippocampal gyrus, and occipital cortex were activated both by remembering the past and imagining the future. Okuda et al. (2003) found that regions in the frontal pole and medial temporal lobe were co-activated by remembering the past and imagining the future. Similar results have been obtained in several other fMRI studies (Abraham, Schubotz, & von Cramon, 2008; Addis, Cheng, Roberts, & Schacter, 2011; Hach, Tippett, & Addis, 2014; Viard et al., 2011).

This common default network activation for past and future thinking emerges not only in individual studies, but also in two recent meta-analyses. Benoit and Schacter (2015) found in a meta-analysis that past and future thinking jointly engaged the hippocampus, parahippocampal cortex, lateral temporal lobe, inferior posterior parietal lobe, and parts of the dorsolateral prefrontal cortex. Spreng et al. (2009) found in a meta-analysis that past and future thinking jointly engaged parts of the medial temporal lobe, precuneus, posterior cingulate cortex, retrosplenial cortex, and temporal-parietal junction. Together, these meta-analyses strongly suggest that past and future thinking jointly activate many regions in the brain’s default network, suggesting the two processes share some common neural correlates.

In summary, there is strong evidence that past and future thinking both activate a shared set of neural regions, especially in the default network. This activation emerges reliably in individual studies and in meta-analyses, and could suggest that past and future thinking rely on common underlying cognitive processes.

1.1.2 Neuropsychological Evidence for Common Processing of Past and Future Thoughts

The second type of evidence for common processing of past and future thinking is neuropsychological. The key evidence is that when patients have deficits to the episodic memory system due to damage to the hippocampus, these patients sometimes to also have deficits to their ability to imagine the future. These results are taken to suggest that past and future thinking share cognitive processing. These results are further taken to suggest that the key shared process may be episodic memory.

The animating idea comes from Klein et al. (2002) who described a patient with severe retrograde amnesia (for similar results, see Andelman, Hoofien, Goldberg, Aizenstein, & Neufeld, 2010; Maguire, Vargha-Khadem, & Hassabis, 2010; Race, Keane, & Verfaellie, 2011; Tulving, 1985). The patient had trouble answering questions about the *lived past* (e.g. what did you do yesterday?) and also about the *lived future* (e.g., what will you do tomorrow?). In contrast, the patient seemed to have preserved knowledge of the *known past* and *known future* (e.g., what will be some of the most important political issues of the next ten years?). These results motivated the proposal that past and future thinking may rely on common cognitive processes, especially with respect to the lived past and future. An additional key inference comes from the fact that the patient's impairment was to the episodic memory system. Based on evidence such as this, it is broadly assumed that if past and future thinking rely on shared cognitive processes, then the shared process is likely to be episodic memory (see also Tulving, 1985).

1.1.3 Behavioral Evidence for Common Processing of Past and Future Thoughts

The third type of evidence that past and future thinking may rely on common cognitive processes is behavioral. In these studies, people are asked to generate past and future thoughts, and aspects of their language are recorded and compared. These studies either compare the language directly between the past and future thoughts, or explore the effect of manipulating some cognitive process used both to generate the past and future thoughts. Similarities are usually observed in the language used to describe the past and future.

One strategy for behaviorally studying similarities is to ask people to generate past and future thoughts, and then to use human ratings to code various aspects of people's language. The underlying logic is that similarities in the language could imply that similar cognitive processes were used to generate the language. For example, Addis, Wong, and Schacter (2008) cued people to generate past or future events, and then coded the number of episodic details in these events using the autobiographical memory interview. People who included more episodic details when remembering the past also included more episodic details when mentally imagining the future. Spreng and Levine (2006) found that when people generated past and future events and rated the distance of these events from the present, past and future thoughts both followed similar log-linear frequency distributions over temporal distance, although the distributions had different slopes. Rubin (2014) asked people to provide phenomenological ratings as they generated past and future events, such as the sense of reliving, intensity, involuntariness, and sense of perceptually seeing and hearing the event. These ratings were highly correlated, such that people who provided higher ratings for past events also provided higher ratings for future events.

A related method for studying past and future thinking behaviorally has been to manipulate some aspect of how people generate both past and future thoughts, and then to ask whether these two conditions are affected similarly by the manipulation. The underlying logic is that if past and future thinking are similarly affected by some manipulation, then they may rely on a common cognitive process being manipulated, although again a stronger non-temporal control condition is needed to make this inference more compelling. For example, D'Argembeau and Van der Linden (2004) found that when people were asked to remember and imagine positive or negative past and future events, the positive events were rated as having a higher feeling of experiencing than negative events, regardless of temporal direction. When people were asked to imagine temporally close compared to temporally distant episodic past and future events, the temporally close events were rated as having more sensory and contextual details and being associated with a greater feeling of re-experiencing. D'Argembeau and Van der Linden (2006) found that people with higher capacity for visual imagery used more sensory details both when remembering past events and when remembering future events. Similar results are obtained with an episodic specificity induction procedure, where people are induced to imagine both the past and future with more episodic detail (Jing, Madore, & Schacter, 2017; Madore, Gaesser, & Schacter, 2014; Madore, Szpunar, Addis, & Schacter, 2016).

Together, these behavioral studies suggest two similarities between past and future events. First, when people are simply asked to generate past and future events, the past and future events they generate often use similar language. Second, when a common manipulation is applied to the way that people are asked to generate past and future events, the manipulation usually has similar effects on both past and future thinking. Both of these results have been interpreted as suggesting common cognitive processing for past and future thinking.

1.2 Evidence Suggesting Differences in Processing Between Past and Future Thinking

While several types of evidence suggest commonalities between past and future thinking, there are also several types of evidence suggesting differences in processing between past and future thinking.

1.2.1 Many Patients with Episodic Memory Deficits Have Intact Future Thinking

The first piece of evidence for differences between past and future thinking is that many patients with episodic memory deficits have intact future thinking. Some of the key evidence for similar processing between past and future thinking comes from patients with episodic memory deficits. Some of these patients also have difficulty imagining the future, suggesting common processing between past and future thinking and implying that episodic memory may be the key shared process. While this evidence is illustrative, a number of patients have been observed with impaired episodic memory but intact future thinking. The presence of these patients casts doubt on shared processing for past and future thinking, and also casts doubt on the role of episodic memory as the key shared process.

There are two types of patient findings that suggest differences between past and future thinking. First, several groups have found that patients with impaired episodic memory due to medial temporal lobe damage can generate future events with the same number of internal episodic details as healthy controls (Hurley, Maguire, & Vargha-Khadem, 2011; Squire et al., 2010). These results suggest that the future simulations generated by patients with episodic memory deficits may not be impoverished, further implying that episodic memory may not be necessary for future thinking. Second, a number of preserved future thinking abilities have been demonstrated in patients with episodic memory deficits. These preserved abilities include the ability to make decisions about the future in delay discounting (Kwan et al., 2012; Kwan et al., 2015) and normal scores on survey measures of future time perspective (Kwan, Craver, Green, Myerson, & Rosenbaum, 2013; see also Craver, Kwan, Steindam, & Rosenbaum, 2014). For example, Kwan et al. (2013) report a patient with episodic memory deficits whose highest time perspective score was future orientation. Another episodic amnesiac patient in the same study made slightly more future-oriented decisions in a delay discounting task than did healthy controls. These results suggest that a number of important types of future thinking, including the ability to think regularly about the future and to make decisions about the future, do not depend on the episodic memory system.

1.2.2 Much of the Evidence for Shared Processing Lacks an Appropriate Negative Control

The second reason for caution regarding common cognitive processing is that most existing evidence has not established that commonalities are unique to mental time travel. Much of the evidence for shared processing, especially the behavioral evidence, is based on comparing some aspect of people's thoughts about the past and future. Similarities between these thoughts are usually found, and these similarities are taken to imply common cognitive processes. However, these studies suffer from a logical gap by failing to establish that the shared process is unique to past and future thinking. Establishing uniqueness requires including a control condition that is similar to past and future thinking, but does not involve mental time travel. Evidence for common cognitive processes would occur if thoughts about the past and future are more similar to each other than to this negative control condition.

There is little direct evidence involving such a negative control, but the evidence that does exist casts doubt on the amount of shared processing, and specifically whether episodic memory is the key shared process between past and future thinking. Hassabis et al. (2007a) asked patients with episodic memory deficits to verbally describe atemporal scenes such as standing in the main hall of a museum. The descriptions generated by these patients were scored as less experiential than the descriptions generated by healthy controls, suggesting that episodic memory impairment might impair an atemporal scene construction system rather than specifically the ability to mental time travel. Also, as previously described, Schacter et al. (2013) found that older adults' reduction in episodic details when describing past and future events also extended to the atemporal control task of describing a picture in front of them. Both of these studies suggest that while past and future thinking may have commonalities, the uniqueness of these commonalities to mental time travel has not been established.

1.2.3 Some Brain Regions Respond More Strongly to Future Thoughts than Past Thoughts

The third reason for caution regarding common cognitive processing is that some brain regions are observed to respond more strongly to future thoughts than to past thoughts. Some of the strongest evidence for similar processing between past and future thinking comes from fMRI studies showing similar activation for past and future thinking. While this evidence is compelling, these studies also reliably find a few regions that respond more strongly to thinking about the future than to thinking about the past. There is no compelling existing interpretation of these increased activations for future thinking compared to past thinking, even though on face such activation is not predicted by the view that past and future thinking rely on common cognitive processes. The most surprising finding is that the hippocampus, typically implicated in episodic memory, often responds more strongly to future thinking than past thinking (Addis, Pan, Vu, Laiser, & Schacter, 2009a; Addis, Wong, & Schacter, 2007; Kirwan, Ashby, & Nash, 2014). Several other regions are sometimes observed to respond more strongly to future thoughts compared to past thoughts, including the anteriomedial frontal pole (Okuda et al., 2003), premotor cortex and precuneus (Szpunar et al., 2007). There is no strong existing explanation of the increased hippocampal activity in the literature. One of the current explanations is that the increased hippocampal activation may reflect increased construction demands for future thoughts compared to past thoughts. However, it is not necessarily clear how such an explanation would be consistent with common cognitive processing for past and future thinking.

1.3 Outline of the Remainder of the Dissertation

In the remainder of the dissertation, I outline an approach to studying past and future thinking that overcomes some of the limitations of prior literature. I also describe three studies comparing the way people think about the past and future. To preview the main result, I find that past and future thinking do rely on common cognitive processes but that the use of schemas, not episodic processing, may be the key shared process.

In Chapter 2, I outline an approach to studying past and future thinking based on mining people’s large-scale, natural talk about time. I also report a preliminary study evaluating a method for automatically extracting temporal references against several other automated methods and human ratings. Finally, I report the results of a study that identifies the episodic processing evident in people’s natural language about the past, present, and future.

In Chapter 3, I develop an approach for identifying semantic past and future thinking in natural language. This approach is based on learning a large-scale model of the kinds of conceptual information people use every day from a large social media corpus. I then develop and evaluate a method to capture one commonly posited feature of schemas – that schemas can fill in missing information – using neural networks. Finally, I report the results of a study that identifies the semantic past and future thinking in people’s natural language.

In Chapter 4, I report the results of a lab-based study of past and future thinking. The approach in Chapters 2-3 is based on mining people’s large-scale natural language, but it is possible that people could think about the past and future differently when they are explicitly prompted under more controlled conditions. I report the results of a study replicating the main comparisons of Chapters 2-3 regarding episodic and semantic past and future thinking, but based on temporal thoughts in the lab.

In Chapter 5, I discuss the broader implications and future directions of this research.

1.4 Summary of Chapter 1

Much is known about how people mentally time travel into the past, but much less is known about how people mentally time travel into the future. This dissertation considers the view that people may use the same processes to mentally time travel into the future as they do to mentally time travel into the past. There is existing neuroscientific, neuropsychological, and behavioral evidence for common processes for past and future thinking. However, this evidence involves a number of limitations, including a failure to establish uniqueness to past and future thinking. This dissertation asks whether there is a process – episodic processing or the use of schemas – that is uniquely associated with thoughts about the future and past compared to thoughts about time that do not involve mental time travel: thoughts about the present. I report three studies investigating this question in Chapters 2-4.

Chapter 2: Episodic Past and Future Thinking

2.1 Introduction

Some of the earliest evidence for similarities between past and future thinking came from studies of *episodic future thinking*. For example, Tulving (1985) reported a patient with episodic memory loss who had difficulty imagining what he would do the next day, responding “I don’t know” and describing his mental picture of the next day as “blank, I guess.” Klein et al. (2002) reported a patient with episodic memory loss who had difficulty answering questions about the lived, *episodic* future (e.g., what will you do tomorrow?) but a seemingly preserved ability to answer questions about the known, *semantic* future (e.g., what will be an important issue in the next 10 years?). Following findings like these, the construct of *episodic future thinking* has been defined by analogy to the episodic memory system, and is thought to involve mental pre-experiencing or simulation of future events (Atance & O'Neill, 2001; Atance & O’Neill, 2005; Schacter, Benoit, & Szpunar, 2017). For example, mentally simulating playing tennis with a friend is an episodic future thought (Szpunar, Spreng, & Schacter, 2014).

Despite the evidence for similarities between episodic past and future thinking, this evidence has a number of limitations, as described in Chapter 1. Chief among these limitations is the tendency to use explicit prompts to elicit temporal thoughts in the lab. Additionally, as discussed in Chapter 1, there have been a number of patients with episodic memory loss but seemingly preserved episodic future thinking, again raising questions about the degree of similarity.

In this Chapter, I outline an approach for studying past and future thinking in general, and episodic past and future thinking in particular, that goes beyond the limitations of prior literature. Using this new approach, I report the results of a study of people's episodic past and future thinking using this new method. To preview the results, I find that episodic past and future thinking may not be the cognitive process shared between past and future thinking. In particular, I find that the episodic language in people's natural talk about the past is more similar to the episodic language in people's talk about the present, which does not involve mental time travel, than it is to the episodic language in people's talk about the future. I conclude by discussing the implications of these results for the view that past and future thinking rely on common cognitive processes.

2.2 Approach

A strong test of similarities between episodic past and future thinking would ask whether these kinds of thinking are similar in people's natural, unprompted thoughts about the past and future. When people naturally talk about time, these thoughts may encompass a much wider range of kinds of past and future thinking than the thoughts typically generated in lab-based tasks. For this reason, if similarities were still observed between the episodic past and future in such a natural, unprompted scenario, it would provide strong evidence for similar cognitive processing. On the other hand, it is also possible that less similarity will be observed between past and future thinking than is traditionally observed in laboratory designs due to the use of natural, unprompted temporal thoughts.

Studying past and future thinking in this way requires three elements. First, this approach requires a source of naturally occurring thoughts about the past and future.

Below, I outline such a source based on the content of people's weblog posts. Second, this approach requires a method for identifying past and future thoughts in this natural source. Below, I identify a method based on automated syntactic and lexical rules for identifying temporal references. Third, this approach requires a method to measure the degree to which each of these temporal references relies on episodic processing. Below, I outline an approach based on automatically identifying the amount of concrete, perceptual, and spatial language in people's temporal references.

2.2.1 A Source for Naturally Occurring Temporal Thoughts: The Blog Authorship Corpus

A good source for naturally occurring temporal thoughts should meet at least three criteria. First, a good source for naturally occurring temporal thoughts should involve thoughts that are relatively spontaneous and unprompted. This criterion allows the thoughts that result to be unaffected by some of the limitations of prior literature to use highly prompted temporal thoughts. Second, a good source of naturally occurring temporal thoughts should involve a large number of different individuals generating these thoughts. This criterion allows the thoughts that result to capture a wide variety of different temporal thoughts. Third, a good source of naturally occurring temporal thoughts should involve a reasonable demographic balance of the individuals generating these thoughts. This criterion allows the thoughts that result to be unbiased by being drawn from particular demographic characteristics such as gender or age.

A source of temporal thoughts that satisfies each of these criteria is the Blog Authorship Corpus (Schler, Koppel, Argamon, & Pennebaker, 2006). As shown in Table

1, the Blog Authorship corpus is a corpus of blog posts from over 19,000 different bloggers and encompassing over 140 million total words. The blog corpus satisfies the first criterion of relying on unprompted temporal thoughts because bloggers are free to write whenever they choose about a topic of their choosing. Indeed, bloggers varied heavily in the amount of text they wrote on their blogs (range 45 - 1,453,475 words), the length of time they maintained the blog (range 1 – 2,028 days), and the topics of the blog posts (some of the topics include work, aging, religion, holidays, technology, days of the week, sleep, and dating). The blog corpus satisfies the second criterion of involving temporal thoughts from a large number of different people. The blog posts include over 19,000 different authors. Finally, the blog corpus satisfies the third criterion of balanced demographics because the corpus was matched across different age groups (13-17y, *N* = 8,240; 20-27y, *N* = 8,086; 33-47*y*, *N* = 33-47) and gender (50% male, 50% female). While the authors of the corpus could in principle talk about topics of their choosing, they frequently made reference to time, including the past, present, and future. Examples of such temporal thoughts in the blog corpus are shown in Table 2.

Having identified a source for naturally occurring thoughts about time, it is necessary to identify temporal references in these thoughts. In a natural corpus, these thoughts do not come pre-identified as about the past, present, or future. Instead, linguistic markers in these sentences must be used to categorize them as referring to the past, present, or future. On a smaller scale, it may be possible to ask human raters to categorize each of these sentences as about the past, present, or future. However, because the corpus involves over 100 million words, it is impractical to identify temporal references using human raters. Therefore in the next section, I describe an automated method for extracting temporal references in the blog corpus.

Measure	Frequency
Total Words	145,245,703
Total Posts	681,237
Unique Authors	19,320
Avg. Posts / Author	35.3
Avg. Words / Post	213.2

***Table 1. Descriptive statistics of the Blog Authorship corpus** used to extract natural temporal thoughts.*

Temporal Reference	Category
They were paid by the politicians of the newly setup post-opposition government.	Past
We got shut out of the first game, and came back to finish in 2nd place in the consolation bracket.	Past
It is not 11:00 PM on a Sunday evening.	Present
I am seriously considering starting studying.	Future
If I don't get accepted at choice A, then hopefully I'll get accepted somewhere else.	Future

***Table 2. Examples of naturally occurring temporal references** in the Blog Authorship Corpus.*

2.2.2 A Method for Extracting Naturally Occurring Temporal Thoughts: The Copley & Wolff Temporal Reference Classifier

Having identified a source of natural temporal thinking, it is necessary to identify sentences in the corpus that refer to the past, present, and future. Below, I begin by describing the difficulty of identifying temporal references in English. I then describe a method for overcoming this difficult to automatically extract temporal references. In the Results section, I report a pilot study evaluating this automated method against human ratings.

Extracting temporal thoughts automatically is a difficult problem in English. It is a difficult problem because syntax alone cannot identify temporal references. Syntax alone cannot identify temporal references because while there is a dedicated past tense morphology for expressing past orientation, there is no analogous morphology for expressing future orientation. Sometimes, future references are marked with lexical items such as *will*, as in

(1) For the tennis match, I will pack a tennis racket.

(2) It will rain tomorrow.

However, these lexical items alone are also not sufficient for identifying future references. Sometimes future references can be expressed in the present tense, as in

(3) The tennis match happens tomorrow.

(4) The music festival finishes tomorrow.

A further difficulty is that the same lexical items that could indicate the future can be present in sentences which are not about the future, as in

(5) The man reviewed his will.

(6) Abraham Lincoln said “leave nothing for tomorrow which can be done today.”

What these examples suggest is that neither syntax nor lexical items alone are sufficient for identifying temporal references. Instead, a successful strategy will likely combine both syntactic and lexical items to identify temporal references.

I use a method for automatically extracting temporal references developed by Copley and Wolff (in prep). As shown schematically in Fig. 1, this method uses a series of rules to identify references to the past, present, and future. Each rule is expressed as a combination of syntactic items (such as a modal, MD) and lexical items (such as *will*), consistent with the above analysis suggesting that both syntax and lexical items are needed to identify temporal references. The rules are expressed in a tree using the parse structure of the sentence (Chen & Manning, 2014; Levy & Andrew, 2006). To classify a sentence as about the past, present, or future, the sentence is matched against each of the rules individually. The basic assumption is that most sentences will match at least one of the rules for the past or future. In these cases, a temporal reference (past or future) can be assigned by a majority vote among the rules. In some cases, a sentence will match neither the past rules nor the future rules. In these cases, a sentence is classified as referring to the present. I report a study evaluating these rules against human ratings in the Results section. Examples of sentences extracted by the classifier are also given in Table 3.

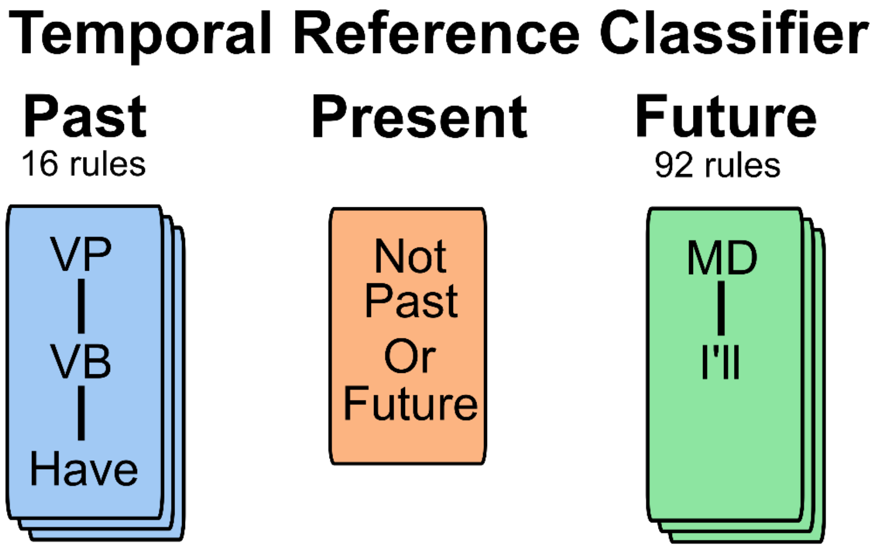

Fig. 1. Schematic of the temporal reference classifier. *The classifier uses a series of rules to identify past and future references, with other references classified as present. The rules are expressed in a hierarchical language combining syntactic constructions such as the modal, MD, with lexical items such as have and I'll*.

Temporal Reference	Category	Rule
The journey will be long and hard but I'm sure I can make it	Future	MD < can
I am very excited for these layers to dry so I can rock & roll on the little critter!	Future	VBN | JJ < excited
Comparative religions was good as usual, and science was…boring, as usual.	Past	VBD
I wish I could say I had no regrets at all.	Past	NP << regret | regrets

Table 3. Examples of temporal references identified by the Copley and Wolff classifier. *For each reference, the relevant rule is also provided, expressed in a hierarchical language known as TREGEX where < represents subordination and the*

capitalized abbreviations represent syntactic types. Note that only one of the four references was extracted using syntax alone, and that many of the temporal references are neither written in the past nor future tense, despite referring to the past or future.

2.2.3 A Method for Measuring Episodic Past and Future Thinking using Language

Once temporal references have been identified in the corpus, a method is needed for identifying the degree to which these references rely on episodic past and future thinking. In this section, I first review how episodic thinking is typically measured in existing literature. I then describe a method which extends these ideas to measure episodic thinking in the blog corpus.

Episodic thinking is typically measured in one of two ways, but neither of these methods can be extended to identify temporal references in the blog corpus. The first method for measuring episodic thinking is to instruct participants to report only episodic past and future thoughts. Highly specific instructions are used, which typically ask participants to report only events that occur in a specific place and time, are highly likely to occur to them, are personal events, and are not extended over time. While instructing participants to report only episodic thoughts is effective at measuring episodic thinking, it is not possible to use such instructions in corpus data. Instead, only the outcome of people's cognition is available (after the thoughts are retrieved), and the problem is to infer how much episodic processing was used to generate those thoughts. The second method for measuring episodic thinking is to use human raters to code for the presence of *internal* episodic details. In this method, people are interviewed based on the modified Autobiographical Memory Interview to generate past and future events (Levine, Svoboda, Hay, Wincour, & Moscovitch, 2002). The details of these events are segmented by trained human coders into *internal* details, which represent a time, place, perception, thought, emotion, or detail central to the main event, and *external* details, which represent any other kind of detail. These internal details are taken as a proxy of episodic processing (Addis et al., 2008; Irish & Piolino, 2016). While the modified Autobiographical Memory Interview is widely used in prior literature, it is neither feasible to interview participants nor to code for internal details in a large language corpus (but see Peters, Wiehler, & Bromberg, 2017 for a proof-of-concept approach to automatically scoring the Autobiographical Memory Interview). For these reasons, a new approach is needed to identify episodic processing in the blog corpus.

A strong method for identifying episodic processing should draw on key principles of prior methods while being possible to extend to a large corpus. The underlying principle of the modified Autobiographical Memory Interview is that people's language is a trace of the cognitive processing that generated that language. Following this principle, human coders can go backwards, using only the language to identify types of details, and by inference these coders can identify the cognitive processes that generated these details. While the Autobiographical Memory Interview in particular cannot be used in a corpus, a similar principle can be used to identify likely markers of episodic processing in people's language. As with the Autobiographical Memory Interview, the key idea is that people's language can provide a marker of the underlying cognitive processes that generated it.

There are several widely agreed upon features of episodic past and future thinking. My method capitalizes on these features to identify these markers in people's language. The first widely agreed feature is that episodic past and future thinking are highly concrete and perceptual. Indeed, episodic future thinking is typically described as a kind of pre-experiencing (Atance & O'Neill, 2001) or simulation (Schacter et al., 2008), and both of these descriptions have in mind a highly perceptual experience. The second widely agreed feature is that episodic past and future thoughts are thought to occur in a specific spatial location. This idea is evident in the fact that episodic future thoughts are elicited by instructions to retrieve events in a specific spatial location (Addis et al., 2011; Addis & Schacter, 2008). This idea is also evident in prior work using spatial relation words as a measure of episodic future thinking in children (Russell, Alexis, & Clayton, 2010; see also Lourenco & Frick, 2013 for evidence that spatial memory is early developing). Thus, my key assumption is that these three features: spatial, perceptual, and concrete language, provide a reasonable proxy for episodic processing in language. My approach is to measure the presence of this language in temporal thoughts as a proxy for episodic past and future thinking.

There are broadly available psychometric methods for measuring each of these linguistic indicators of episodic processing (Fig. 2). Concrete language can be measured based on a set of 40,000 common English lemmas which have been rated for their concreteness on a 1-5 scale (Brysbaert, Warriner, & Kuperman, 2014). Additionally, both perceptual language and spatial language can be measured using lists of words from the Linguistic Inquiry and Word County psychometric dictionary, a dictionary constructed and validated for psychological studies of language (Pennebaker, Boyd, Jordan, & Blackburn, 2015). Examples of words for each of these categories are provided in Table 4.

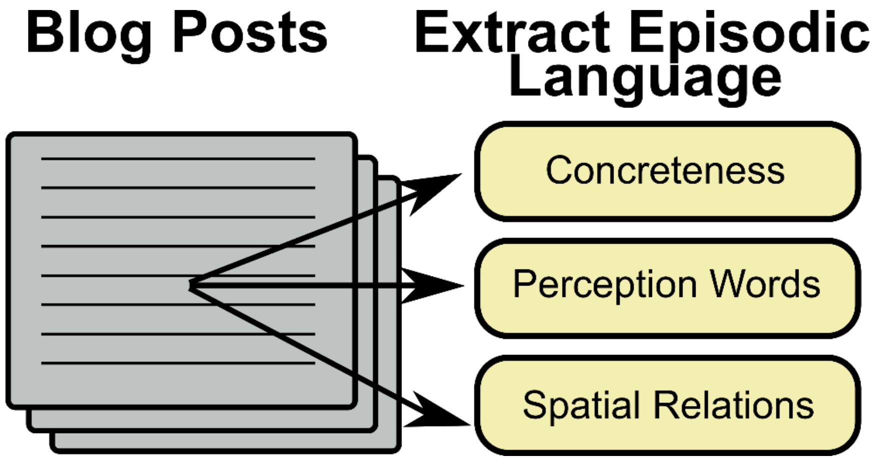

Fig. 2. Schematic of the procedure for identifying episodic language in blog posts. *I extract three indicators of episodic language in each sentence: concreteness, perceptual language, and spatial language*.

Example Words	Category
Acrid | Blue | Nasal | Pain | Taste | Whisper | Yell	Perceptual Words
Above | Across | Below | Inside | Narrowest | Uppermost	Spatial Words
Pickled | Wooded | Orange | Rusted | Minty	Highly Concrete Words (>4/5)
Infinite | Unethical | Religion | Mathematical	Low Concrete Words (<1/5)

Table 4. Examples of words used to identify episodic language. *The perceptual and spatial words come from the LIWC psychometric dictionary. The concrete words are rated 1-5 for concreteness; examples are provided of high concrete words (rated > 4/5 concreteness) and low concrete words (rated < 1/5 concreteness).*

2.3 Methods

***Materials: Blog Authorship Corpus*.** The Blog Authorship Corpus (Schler et al., 2006) was used. The corpus contains 681,288 blog posts collected from 19,320 bloggers in 2004 on www.blogger.com. Users in the corpus also self-identified their age (range 13-47 years), gender (male or female), and occupation (from 40 categories such as law or maritime). The corpus contains an equal number of male and female bloggers and a broad range of ages, although users aged 33-47 were less frequent (13-17 yrs, N = 8,248 users; 23-27 yrs, N = 8,086 users; 33-47 yrs, N = 2,944 users). Users posted on these blogs with moderate frequency (M = 35.3 posts/users), and posts had a moderate length (M = 213.2 words/post). A broad range of occupations were represented, although unknown occupation (N = 6,827 users) and student (N = 5,120 users) were most common.

***Text Preprocessing*.** Internet text corpora contain a variety of types of sentences that can be problematic for automated analyses, including sentences written in foreign languages, extremely short sentences, and sentences with unusual characters such as emoticons. The following preprocessing steps were taken to ensure that the sentences could be understood by the automated analyses that followed. First, blog posts were split into individual sentences (N = 9,453,891 sentences) using the Python Natural Language toolkit module (NLTK) version 3.2.2. Next, extremely short sentences (less than 10 tokens, including punctuation) were detected using NLTK. Sentences containing URLs, as labeled by the authors of the Blog Corpus, were also detected. These short sentences and sentences containing URLs (N = 3,402,717) were removed from the corpus. Next, sentences containing emoticons such as :) (*N* = 44,594) were removed using the Python library python-twitter version 3.3. Next, non-standard spellings were corrected (*N* = 1,309,756 words in 816,288 sentences) using a lexical normalization dictionary developed for use with online text (Han, Cook, & Baldwin, 2013). The dictionary defines 44,447 pairs of common misspellings and corrected spellings. For each misspelling, all instances of the misspelling in the corpus were detected using regular expression search and corrected to the spelling specified in the dictionary. Finally, sentences not written in English (*N* = 117,941) were removed using the Python langdetect package version 1.0.7. Langdetect is a machine learning classifier that estimates the probability that a sentence belongs to each of 55 languages. Sentences for which the largest of these probabilities was not English were removed. After processing, a total of *N* = 5,888,640 sentences from 19,315 bloggers remained.

Extraction of Temporal References**.** Each sentence in the corpus was classified as about the past, present, or future using the rules defined in Copley and Wolff (in prep). In total there were 92 future rules and 16 past rules. First, each sentence was constituency parsed using the Stanford Parser (Manning et al., 2014) implemented in Stanford CoreNLP version 3.6.0. For each rule, the number of times the construction specified by that rule occurred in the sentence was counted. *N* = 1,914 sentences could not be classified in this way and were excluded from analysis, either due to extremely long sentences that could not be parsed successfully (*N* = 1,837 sentences) or malformed output from the parser (*N* = 77 sentences). Next, each sentence was classified as about the past or future by counting the number of past and future constructions in the sentences and selecting the larger of these counts as the class label. Sentences with an equal number of past and future constructions could therefore not be clearly classified as either past or future (N = 489,537) and hence were removed. Sentences where no future or past constructions were identified were classified as present. In total, N = 2,134,357 sentences (39.5%) were classified as past, N = 1,428,626 sentences (26.5%) were classified as present, and N = 1,834,206 sentences (34.0%) were classified as future, with a total of N = 5,397,189 sentences extracted from N = 19,309 bloggers.

Evaluation of Temporal Reference Classifier. A pilot study was conducted to evaluate the temporal reference classifier. Human ratings are typically used as a gold standard for evaluating natural language processing models. In this study, human ratings were collected for whether sentences in the blogs referred to the past, present, and future. These human ratings were compared to the judgments of the temporal reference classifier, as well as those of several other automated classifiers.

Participants. N = 40 participants were recruited from Amazon Mechanical Turk and compensated $2.00 for participant. Participants were required to be located in the United States, have completed at least 100 previous tasks (HITs) on Mechanical Turk, and have at least a 95% approval rate for these HITs. Multiple submissions from the same participant were prevented programmatically.

Materials. *N* = 1,000 sentences were randomly drawn from the blog corpus with the constraint that no more than 1 sentence was written by each individual blogger. These sentences were randomly divided into 10 lists of 100 sentences to be rated. An additional *N*=5 sentences with clear answers were added to each list as unmarked attention checks. *N* = 2 sentences were incorrectly presented due to programmatic error and excluded from analysis, leaving *N* = 998 sentences remaining.

Methods. Each list of 100 sentences (+5 attention checks) was rated by 4 independent raters for whether the sentence was primarily about the past, present, future, not temporal, or unintelligible. Participants completed an online informed consent and then read the rating criteria, which provided a definition and example of each category. Participants then provided a forced-choice ratings for each sentence, 1 sentence at a time.

Classification Algorithms. The full list of sentences were presented to 5 separate classification algorithms, each of which rated every sentence as primarily about the past, present, or future. The following algorithms were used.

1. *Stanford Temporal Tagger (SUTIME)*. SUTIME is a rule-based classifier that uses a combination of regular expression, such as “yyyy-?MM-?dd,” and keywords, such as “tonight,” to identify temporal references (Chang & Manning, 2012). SUTIME was used to identify temporal references in each sentence. These references were compared to the date the blog post was created to classify it as past, present, or future using methods previously described (Thorstad & Wolff, 2018). If no temporal references were detected, the sentence was classified as atemporal. *N* = 4 sentences could not be classified due to programmatic error during classification.

2. *Schwartz et al 2015*. Schwartz et al. (2015) trained a Forest of Extremely Randomized Trees machine learning model to identify temporal references in text. The model uses several types of temporal features, as well as part-of-speech counts. The model was directly replicated as reported in Schwartz et al. (2015), with the following parameters: 1,000 estimators, gini impurity measure, and a maximum number of features equal to the square root of the total number of features. The model was evaluated with 5-fold cross-validation using the Python library scikit-learn version 0.18.2.

3. *Linguistic Inquiry and Word Count (LIWC)*. LIWC (Pennebaker et al., 2015) is a keyword-based text analysis program. LIWC includes categories for future-references (e.g. “anticipate”, “gonna”), past-references (e.g. “ate”, “met”), and present-references (e.g. “is”, “today”). The LIWC 2015 dictionary was used to identify the proportion of past, present, and future references in each text. For each sentence, the most frequently occurring temporal category was used to classify the sentence as past, present, or future. Ties were broken randomly. Sentences where no past, present, or future references were detected were classified as atemporal.

4. *Copley and Wolff (in prep)*. The Copley and Wolff model was used as described above. N = 1 sentence could not be classified due to programmatic error during parsing.

5. *Human*. Sentence ratings from the 4th set of human raters (N = 1 rating for each sentence) were used to label each sentence as about the past, present, future, atemporal, or unintelligible.

***Identification of Episodic Language*.** Three different measures of episodic language were combined to create a single score for each sentence in the corpus representing the amount of episodic language in the sentence. First, the concreteness of each sentence was scored from 1-5 using 40,000 English lemmas rated for concreteness (Brysbaert et al., 2014). This scoring was done by tokenizing each sentence into individual words and then averaging the concreteness scores for each word in the sentence. Second, the perceptual language of each sentence was scored using a list of perceptual words such as *blue* and *cold* from the LIWC dictionary (Pennebaker et al., 2015). This scoring was done by tokenizing each sentence into individual words and then counting the proportion of those words that were identified as perceptual. Third, the spatial language of each sentence was scored using a list of spatial words such as *above* and *below* from the LIWC dictionary (Pennebaker et al., 2015). The scoring procedure was identical to the procedure for perceptual language.

2.4. Results

The main result was that thoughts about the future were less episodic than thoughts about the past. This result was observed for all three measures of episodic processing. Below, I first provide the results of a pilot study evaluating the temporal reference classifier, where I found that the classifier performed better than several other algorithms and approached human-level performance. I then describe the episodic language identified in these temporal references.

2.4.1 Performance of Temporal Reference Classifiers

To evaluate whether temporal references can be reliably extracted from language, I conducted a pilot study where I compared the Copley and Wolff temporal reference classifier to several other automated classifiers and to human ratings. I conducted this comparison using 1,000 randomly drawn sentences from the Blog corpus. To preview the main result, I found that the Copley and Wolff model outperformed the other algorithms and performed almost as well as human raters.

I first verified that human raters understood the task and provided quality annotations. Human raters performed well on unmarked attention checks (M = 4.83/5 correct) representing sentences with clear temporal classes. Nevertheless, inter-rater agreement was moderate for human raters in general, 74.4% agreement. This pattern of results suggests that while human raters understood the task and broadly agreed about clear cases, even human raters often disagree about whether a sentence refers to the future.

Performance of the classification algorithms is shown in Fig. 3. The vertical axis plots accuracy as a function of *F* score (Raschka, 2015), a standard machine learning classification metric. An *F* score of 1 represents perfect performance and 0.33 represents chance performance. Two automated classification models performed worse than chance: the SUTime model (F = 0.251) and the Schwartz et al model (F = 0.296). Post-hoc exploratory analyses suggest that the SUTime model had a low ability to identify temporal references (recall), R = 0.17 where chance = 0.33, incorrectly labeling the majority of cases as atemporal. These analyses also suggest that the Schwartz et al model may have overfit to the distribution of the data, learning a strategy of rarely labeling sentences as future (*F* = 0.020, *R* = 0.011 for future where chance is 0.33 for both measures). A simple keyword-based classifier, LIWC, performed reasonably well (*F* = 0.562), although the model's ability to successfully identify temporal references, recall, was highly inconsistent (range 0.204-0.783 across classes, where chance is 0.33). The best performing automated model was the Copley and Wolff model, which had both high recall (*R* = 0.660) and high accuracy when a label was chosen (precision), *P* = 0.650. Performance of the Copley and Wolff model approached human performance (*F* = 0.672), although humans raters performed best overall.

These results suggest that while temporal references are difficult for humans and algorithms to identify, the Copley and Wolff model performs almost as well as humans and performs better than 3 automated methods from the literature. For these reasons, I chose to use the Copley and Wolff model to identify temporal references in the following studies. Descriptive statistics of the temporal references identified are shown below in Table 5. It can also be seen in Table 5 that people thought frequently about the future (34% of all thoughts), almost as frequently as they thought about the past (39.5% of all thoughts).

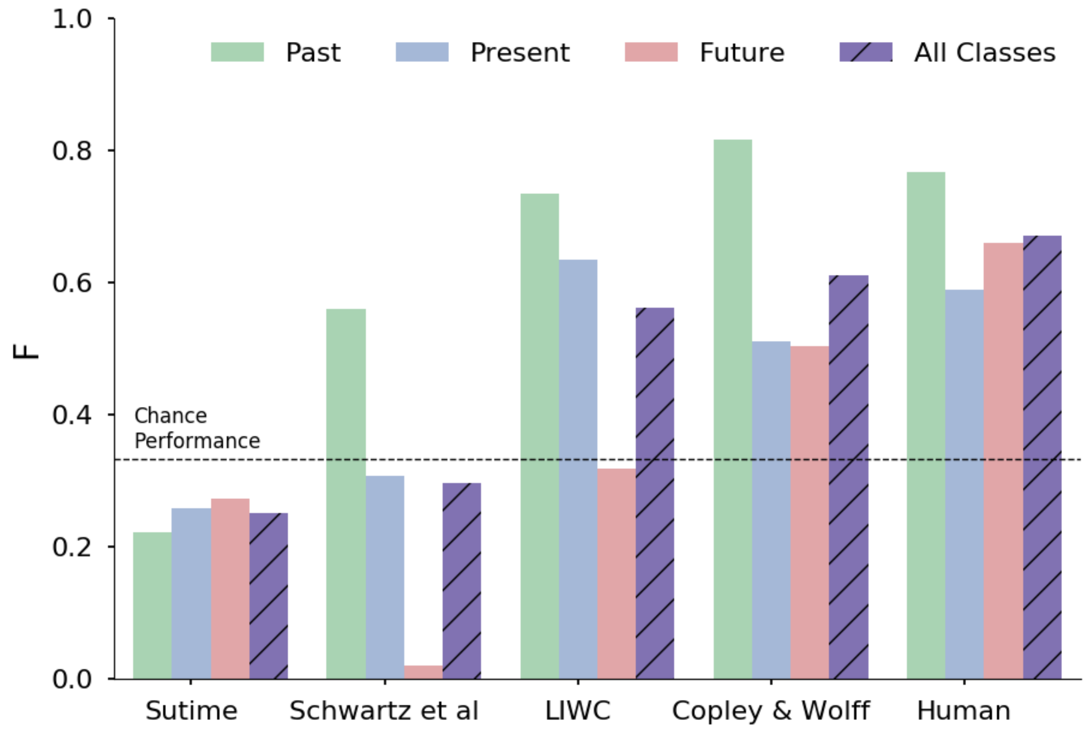

Fig. 3. Performance of automated algorithms for classifying temporal references *based on human ratings of 1,000 sentences from the blog corpus. The vertical axis plots performance (F1-score), where 0.33 is chance performance. The Copley & Wolff model performed best overall (purple bar) and approached the performance that would be expected for a single human rater.*

Category	Number of References	Percent of Sentences
Past	2,134,357	39.5%
Present	1,428,626	26.5%
Future	1,834,206	34.0%

***Table 5. Total number and proportion of temporal references** extracted from the blog corpus.*

2.4.2 Episodic Language in Temporal Thoughts

Having extracted temporal references from the blog posts, I next identified the amount of episodic language in these references separately for references to the past, present, and future. To preview the main result and as shown in Fig. 4, I found that thoughts about the past involved more episodic language than thoughts about the future for all three measures of episodic language.

First, I compared the episodic language in references to the past and future. I found significant differences for all three measures of episodic language. Thoughts about the past involved more spatial language than thoughts about the future, $t_{(18,808)} = 48.34$, $p < 0.001$, $d = 0.42$. Thoughts about the past also involved more perceptual language than thoughts about the future, $t_{(18,808)} = 23.27$, $p < 0.001$, $d = 0.21$. Finally, thoughts about the past were more concrete than thoughts about the future, $t_{(18,806)} = 46.65$, $p < 0.001$, $d = 0.35$.

Second, I compared the episodic language in references to the past and present. The general pattern was that thoughts about the past were much more similar to thoughts about the present in the level of episodic language than they were to thoughts about the future. Thoughts about the past were not significantly more spatial than thoughts about the present, $t_{(18,806)} = -0.67$, $p = 0.50$, $d = 0.001$. With respect to concreteness and perceptual language, thoughts about the past and present did differ, but the magnitude of the effects was much smaller than the differences between the past and future: for concreteness, $t_{(18,806)} = 8.39$, $p < 0.001$, $d = 0.06$ and for perceptual language, $t_{(18,806)} = 6.92$, $p < 0.001$, $d = 0.06$.

Together, these results suggest that thoughts about the future are less episodic than thoughts about the past. The results also suggest that in terms of episodic language, thoughts about the past are more similar to thoughts about the present than to thoughts about the future.

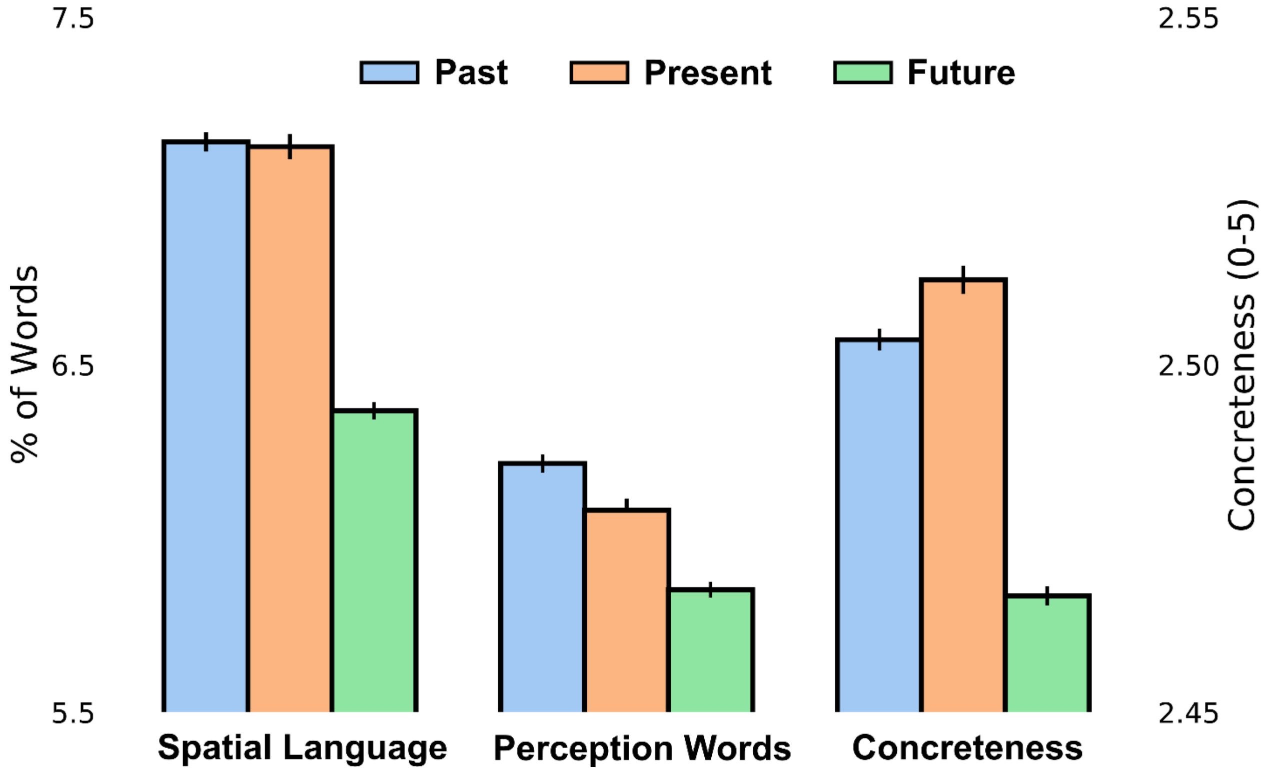

Fig. 4. Amount of episodic language in the blog posts *(+/- 95% bootstrapped CI). The amount of episodic language is plotted separately for each indicator as a percentage of words (for spatial and perceptual language) and as a 0-5 rating (for concreteness). Colors indicate type of temporal reference.*

2.5. Discussion

In this study, I extracted more than five million naturally occurring thoughts about the past, present, and future from a corpus of web blog posts. An automated method was used to identify three markers of episodic processing in these posts: the use of spatial language, the use of perceptual language, and the use of concrete language. Across all three measures, the main finding was that thoughts about the future were less episodic than both thoughts about the past and thoughts about the present.

These results contrast to existing literature, which sometimes finds similarities in the level of episodic processing between past and future thinking. As previously discussed, the difference is likely due to two features of the study design. In the current study, I investigated naturally occurring thoughts about the past and future, and I studied thoughts that were unprompted. In the existing literature, people are typically prompted to remember the past and future in a laboratory setting. My results suggest that the use of these prompts may increase the similarity between past and future thinking, perhaps by instructing people to remember the past and future in very similar ways.

An additional contribution of the study was to extract one of the largest-scale corpora of temporal thoughts to date (for a related experience-sampling approach, see Baumeister, Hofmann, Summerville, Reiss, & Vohs, 2018). One of the motivating ideas behind the study of prospection is that people think about the future quite frequently, perhaps just as often as they think about the past (Baumeister, Vohs, & Oettingen, 2016; Gilbert & Wilson, 2007). My results are consistent with this idea: people thought almost as frequently about the future (34.0% of sentences in the corpus) as they did about the past (39.5% of sentences in the corpus). Indeed, there were over 1.8 million future references in the corpus. While I used this corpus of temporal references to study episodic language, it should be possible to extend this corpus in many other ways. For example, age and gender information are available to investigate whether the frequency and content of temporal thoughts differs by age and gender. It should also be possible to ask questions about the temporal distance of past and future thoughts (Spreng & Levine, 2006; Thorstad & Wolff, 2018) or many of the other phenomenological characteristics of past and future thoughts that are typically studied such as valence (D'Argembeau & Van der Linden, 2004, 2006; Rubin, 2014; Szpunar et al., 2007).

The temporal classifier used in the current study was able to classify sentences as about the past, present, or future with nearly human-level accuracy. However, I also observed that humans often disagreed about whether a sentence refers to the past, present, or future. One reason for this variation might be that most naturally occurring thoughts are not purely about one temporal reference, but can simultaneously refer to the past and future. For example, the sentence *I was thinking about what I will do tomorrow* refers to both the past and the future. While an empirical investigation of this kind of language is beyond the scope of the current study, an interesting extension of the automated analysis method would be to identify sub-components of a sentence, such as clauses, that may individually refer to the past, present, and future.

In the current study, I found that past and future thinking differ in their level of episodic processing, which is typically one of the hallmarks of similar processing between past and future thinking. Indeed, I found that the level of episodic processing in the past was more similar to thoughts about the present, which do not involve mental time travel, than to thoughts about the future. If past and future thinking differ in their levels of episodic processing, a natural question is what processes might be shared by past and future thinking. It is well known that people draw on their conceptual knowledge of the world, often in the form of schemas, to remember the episodic past (Bauer, 1993; Bower, Black, & Turner, 1979; Brewer & Treyens, 1981; Fivush, 2002). It has also been argued that people need to draw on this schema knowledge when they imagine the future (Rubin, 2014; D'Argembeau & Mathy, 2011; Irish & Piguet, 2013). By contrast, it seems natural that people may need to draw less on this schema knowledge when experiencing the present. For these reasons, in the next chapter I asked whether the common processing underlying past and future thinking might involve schema knowledge, what is referred to as semantic future thinking.

Chapter 3: Semantic Past and Future Thinking

3.1. Introduction

In Chapter 2, I found that episodic processing may not be the shared cognitive process between past and future thinking. However, despite a historical focus on episodic processing, the view that past and future thinking rely on shared cognitive processes is not limited to episodic processing. Future thinking is typically divided into two types: episodic and semantic future thinking (Szpunar, Spreng, & Schacter, 2016; Szpunar et al., 2014). *Semantic future thinking* is thought to encompass thoughts about the future that draw on people's abstract, high-level factual knowledge about the world without the level of re-experiencing characteristic of episodic future thoughts (Irish & Piguet, 2013). For example, knowing that a tennis racket should be packed for a tennis match tomorrow is a semantic future thought. The distinction between episodic and semantic future thinking is inspired by a similar distinction in the study of human memory (Tulving, 1972; Vargha-Khadem et al., 1997), as well as by patient data showing that damage to the hippocampus can impair patients' episodic future thinking while having little effect on semantic future thinking (Klein et al., 2002; but see also Craver et al., 2014 and De Luca et al., 2018).

In this chapter, I ask whether the main similarity between past and future thinking may occur not in episodic processing, but rather in semantic processing. First, I review evidence for the use of conceptual knowledge in thoughts about the past and future. Second, I review some of the reasons that schemas have been difficult to identify in research on future thinking, and I outline an automated approach to detect schemas in past, present, and future thoughts. Third, I report the results of a study measuring the use of schemas in people's naturally occurring thoughts about the past, present, and future. To preview the main result, I find that both past and future thoughts rely more on schemas than do thoughts about the present, suggesting that semantic past and future thinking may rely on common mental processes not as commonly used to think about the present.

3.1.1 Evidence for Schema Usage in Past Thinking

Here, I describe a long tradition of evidence arguing that people use schemas often (or even always) when they remember the past. I begin by outlining some of the ways this conceptual knowledge has been operationalized in the literature: namely as either schemas or scripts. I then review some of the evidence that people use these schemas and scripts to remember the past.

The influence of conceptual knowledge on memory has variously been described as taking the form or a schema or a script. A schema is thought to be a hierarchically structured form of knowledge about the world, such as the knowledge that a chair has four legs and is used for sitting (Hintzman, 1986). Conceptual knowledge has also been described as taking the form of structured knowledge of events called scripts. An example of a script is the knowledge that at a restaurant, a waiter first brings a menu and then takes the customers' orders (Schank & Abelson, 1975).

There is broad evidence for the influence of schemas on memory. For example, when people remember items from an office, they are more likely to remember items consistent with their schema of an office such as chairs than items inconsistent with their schema such as a Frisbee (Brewer & Treyens, 1981; Pezdek, Whetstone, Reynolds, Askari, & Dougherty, 1989). These schemas are thought to affect memory from an early age, as 25 month-old boys better remember sequences showing male-stereotyped behaviors such as building a house compared to female-stereotyped behaviors such as cooking breakfast (Bauer, 1993). Children's event schemas are also thought to affect the way they remember traumatic events in early childhood (Fivush, 2002). The influence of schemas on memory has been formulated computationally (Hintzman, 1986) and neurally (van Kesteren, Ruiter, Fernández, & Henson, 2012).

Together, these studies imply strong evidence that people's conceptual knowledge about the world influences their memories of the past.

3.1.2 Evidence for Schema Usage in Future Thinking

In this section, I describe evidence that people draw on their conceptual knowledge not only when they remember the past, but also when they imagine the future. I describe two types of evidence based on patients with loss of conceptual knowledge due to semantic dementia, and based on behavioral observations of people's thoughts about the future.

The first type of evidence that people use conceptual knowledge to think about the future comes from patients with deficits to their conceptual knowledge due to semantic dementia. When these patients are asked to remember past events or imagine future events, the events they produce are scored as impoverished in the amount of *internal* (episodic) details provided based on the Autobiographical Memory Interview (Irish, Addis, Hodges, & Piguet, 2012). However, the patients' impairment is larger for thoughts about the future than the past and is distinct from the impairment in a population of patients with only episodic memory deficits, suggesting that damage to the conceptual system may more strongly impair future thinking than past thinking. A similar pattern of results was observed in patients with behavioral frontotemporal dementia (Irish & Piguet, 2013). These results have two implications. First, the results suggest that semantic knowledge is important to remembering the future because the loss of conceptual knowledge has a large impact on patients' ability to remember the future (Irish & Piguet, 2013; Irish & Piolino, 2016). Second, the results suggest that semantic knowledge may be a shared process between remembering the past and future because patients were impaired both when they remembered the past and imagined the future, although the impairment was stronger for thoughts about the future. Although suggestive, these results are limited due to the use of patient populations and small numbers of trials.

The second type of evidence that people use conceptual knowledge to imagine the future comes from behavioral studies of people generating thoughts about the future. When people are asked to verbalize their thoughts about the future as they construct these thoughts, people often begin with a schema (e.g., *I'm thinking of a birthday…*) and then use this schema to scaffold the generation of particular details of the event such as cake and presents (D'Argembeau & Mathy, 2011). Relatedly, other studies have asked people to think about the future, and then have coded the events they describe for the presence of cultural life scripts such as weddings or birthdays. These life scripts are found frequently in people's thoughts about the future, especially when they write about events in the distant future (Berntsen & Jacobsen, 2008). Finally, when people generate future thoughts in response to cues, there is often a schema-like relation between the cue and the event that people generate, such as a causal or inclusion relationship, a phenomenon known as event clustering (D'Argembeau & Demblon, 2012; Demblon & D'Argembeau, 2014).

Together, these results suggest that people frequently use schemas to imagine the future, especially when people imagine events in the distant future. The results also suggest that damage to people's conceptual knowledge via semantic dementia impairs the ability to imagine the future and may even have a stronger impact on thoughts about the future than thoughts about the past.

3.1.3 Absence of Evidence about Similarity of Schema Usage in Past and Future Thinking

While much is known about how schemas are used to remember the past and imagine the future, very little is known about whether people use schemas to the same extent (or in the same way) in past and future thoughts. In this section, I describe a likely cause of this gap: the absence of existing approaches to identify schemas at scale in thoughts about the past and future. I describe reasons that existing approaches cannot simultaneously capture schemas in past and future thinking at scale, and in the next section I outline an approach that can potentially capture these schemas.

There are two common ways to identify schemas in future thoughts, but both approaches cannot easily be extended to large corpora. The first approach is to code people's descriptions of future thoughts using the autobiographical memory interview for *internal* episodic details about the main event (Levine et al., 2002). All other details are categorized as *external*, and these external details can be used as a rough measure of semantic processing. However, such an approach is problematic because external details can also include cognitions that are not semantic such as details about another event or filler-words such as *um*. Additionally, there is no existing approach to automatically score the Autobiographical Memory Interview at scale (although see Peters et al., 2017 for a proof-of-concept approach).

A second existing approach to identifying schema usage in future thoughts is to manually code for the presence of a limited set of schemas such as cultural life scripts (Berntsen & Jacobsen, 2008). As previously discussed, these approaches have been used to reveal the influence of schemas on people's thoughts about the future, especially thoughts about the distant future. However, one of the major limitations of such an approach is the requirement that the schemas be specified in advance, for example by limiting the schemas to a set of roughly 30 important cultural life scripts. It is likely that people can draw on many more schemas than these when they remember the past and imagine the future. Additionally, there is currently no approach to automatically identify the use of cultural life scripts at scale in a large corpus.

This discussion suggests that what is needed is a method for automatically measuring schema usage in a large corpus of thoughts about the past, present, and future. In the next section, I describe such an approach.

3.2 Approach

Here, I outline an approach to automatically identifying schemas that can extend to a large corpus and can simultaneously measure schema usage in thoughts about the past, present, and future. Such an approach should meet three criteria. First, the representations learned by the model should at least approximate the types of representations understood as a schema. Second, a large corpus of everyday thinking is needed for the model to learn the kinds of everyday schemas that people use repeatedly and are broadly shared. Third, a method is needed for measuring not just whether a thought incidentally conforms to a schema, but whether the schema was likely used as part of the cognitive processes that generated the thought. Below, I describe an approach that satisfies all three of these criteria.

3.2.1 A Computational Model of Schemas using Topic Modeling

Identifying schemas at scale requires a class of models that can learn representations that meet some of the characteristics of a schema. Here, I review evidence that a class of unsupervised learning models known as topic modeling learns representations that meet several of the characteristics of schemas. In the next section, I discuss how such topic models can be applied to large language corpora to learn common everyday schemas.

Distributional semantic models are models that learn the meaning of language from the context that words occur in a corpus. Distributional semantic models comprise the state-of-the-art in machine learning approaches to processing language (e.g., Wang et al., 2018) and are motivated by the linguistic intuition that the meaning of a word is largely given by the contexts in which it can appear (Firth, 1950). Early versions of these distributional semantic models, such as Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990), focused on analyzing the distribution of words across documents, while more recent versions such as word2vec (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013), BERT (Devlin, Chang, Lee, & Toutanova, 2018) and ELMO (Peters et al., 2018) use neural networks to learn the distribution of words across short contexts such as sentences.

A particular type of distributional semantic model known as a topic model tackles the problem of learning high-level groups of semantically meaningful information from text. Here, I describe the most commonly used form of topic model: Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2002, 2003); for a recent extension of topic models to a deep learning framework, see Cao, Li, Liu, Li, & Ji, 2015). As shown in Fig. 5, Latent Dirichlet Allocation works by capitalizing on the fact that one of the strongest constraints on the choice of words is the topic, or topics underlying a particular document. For example, an author writing about basketball would be much more likely to write *court* or *net* than *library*. Latent Dirichlet Allocation reverses this generative process to use the observed distribution of words in documents to learn the likely latent topics that generated those documents, typically using an inference procedure called Collapsed Gibbs Sampling (Porteous et al., 2008; Griffiths & Steyvers, 2004). The result of the model is a set of topics that are typically represented as a set of several words that are highly probable in a given topic (although the topic is mathematically expressed as a probability distribution over every word in the vocabulary). A more detailed formal description of Latent Dirichlet Allocation is available in Appendix 1.

There is some history of arguing that distributional semantic models, especially Latent Dirichlet Allocation, are psychologically effective models of the human conceptual system. Griffiths, Steyvers, and Tenenbaum (2007) showed that Latent Dirichlet Allocation recovered a number of psychological phenomena, such as predicting the semantic priming resulting from a word, predicting the effect of a semantic intrusion in free recall, and predicting people's word association ratings. One of the strengths of Latent Dirichlet Allocation was that the model was able to recover asymmetric word associations, for example that *text* is more highly associated with *book* than *book* is with *text*. More recently, Latent Dirichlet allocation has been used to recover a number of psychological phenomena from text on social media, especially personality (Park et al., 2015; Schwartz et al., 2015) and mental health (Eichstaedt et al., 2018; Guntuku, Yaden, Kern, Ungar, & Eichstaedt, 2017). Finally, beyond Latent Dirichlet Allocation, there is a long tradition of arguing that models of the same broader class – distributional semantic models – can be used to model human semantic cognition. These distributional semantic models of human cognition include BEAGLE (Jones, Kintsch, & Mewhort, 2006; Jones & Mewhort, 2007) and HAL (Lund & Burgess, 1996; Rohde, Gonnerman, & Plaut, 2006). Latent Dirichlet Allocation improves on these models by not just learning representations for individual words, but also learning coherent higher-level semantic groupings among those words. Together, these studies suggest that distributional semantic models can provide reasonable accounts of the human conceptual system, although these models do fail to capture some features of schemas such as the fact that many schemas are hierarchical (but see Griffiths, Jordan, Tenenbaum, & Blei, 2004 for a possible extension to learn hierarchical topics).

Perhaps the largest gap in applying distributional semantic models to the human conceptual system has been the lack of large corpora to train these models. While the previous studies suggest that topic models can recover a number of phenomena in the human conceptual system, the concepts learned by the model are only as general as the text the model is trained on. In the next section, I describe an approach to this problem that capitalizes on the recent availability of very large, diverse corpora of text on social media.

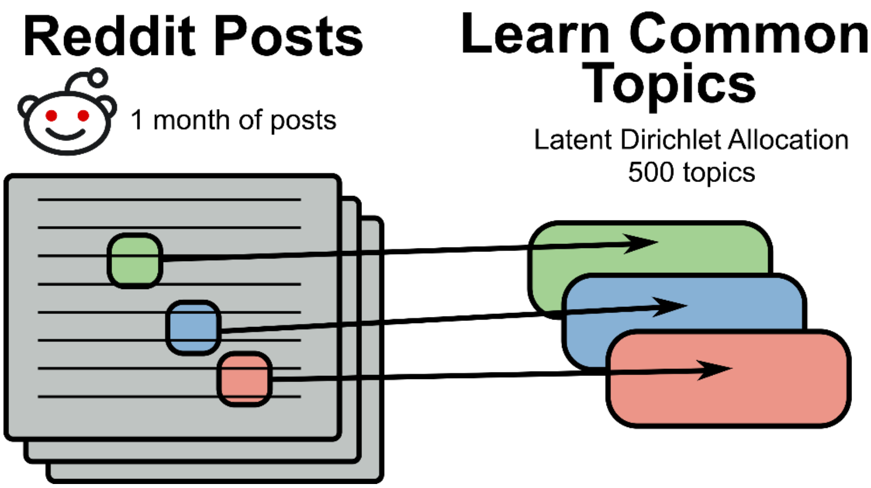

***Fig. 5. Latent Dirichlet Allocation Method for Identifying Schemas**. Latent Dirichlet Allocation works by inferring the latent topics (represented by colored boxes) that organize people's choice of words in particular documents (here, represented by Reddit posts on the left-hand side). The key assumption is that these latent topics inform the choice of which words to write in a particular document, an assumption that allows reversing the inference to use the distribution of words across documents to infer the likely latent topics used to generate those words. For a more formal definition of the model in Bayesian plate notation, see Appendix 1.*

3.2.2 Learning Common Schemas: a Corpus Approach

The first step to identifying human schemas is to learn these schemas from data. The schemas learned should capture a large proportion of the human conceptual system. In this section, I describe how the availability of large social media corpora makes possible the training of topic models in a way that can reasonably capture a large portion of the human conceptual system.

While there has been a large amount of technical development in topic modeling, there is a relative gap in training topic models on corpora large and general enough to learn about the human conceptual space. Topic models are often trained on restricted corpora such as articles in scientific journals (Blei et al., 2003; Blei & Lafferty, 2007;Griffiths & Steyvers, 2004), relatively small corpora of published books (Griffiths et al., 2007), or corpora of news articles (Cao et al., 2015). One exception is a topic model recently trained on a large corpus of Facebook status updates (Schwartz et al., 2013); however, this model learned a large enough set of topics (2,000 topics) that many of the topics are not semantically interpretable. This analysis suggests that in order to make useful inferences about the human conceptual system, a topic model is needed that is trained on a large and general corpus of human language.

In domains other than topic modeling, large corpora of human behavior have been successfully used to make a variety of inferences about human psychology. The success of these methods suggests that, in principle, large online corpora could be used to train large-scale models of the human conceptual system. For example, Vinson, Dale, and Jones (2019) recovered sequential dependencies in decision-making based on the dependency of online restaurant reviews on the previous reviews (see also Schulz et al., 2019, for a related analysis). Johns and Dye (2019) recovered gender bias from a billion-word corpus of published books. Thorstad and Wolff (2018) predicted people’s intertemporal decisions from the future thinking implicit in their tweets. Thorstad and Wolff (2019) extended these methods to predict whether individuals may have a mental illness based on the text of their posts on social media. The success of large-scale online approaches to understanding human behavior has been the subject of several recent reviews (Paxton & Griffiths, 2017; Goldstone & Lupyan, 2016; Lupyan & Goldstone, 2019; Jones, 2016). These results suggest that large-scale online corpora can be used to make inferences about human psychology.

What is needed is a method that combines the large-scale corpora that have been used to make successful inferences about human behavior with the technical developments in topic modeling enabling training of such models on large-scale corpora. In the current study, I access a large-scale corpus of everyday cognition based on a month of posts (307 million words) on the social media website Reddit. Reddit is a forum-like social media website where users can participate in one or more discourse communities called subreddits. There are more than 100,000 active subreddits on Reddit, and these subreddits are dedicated to a wide variety of different topics including sports, cooking, science, chess, mental health, and travel. The key idea is that by training a model to infer the most common latent topics across such a wide variety of discourse communities, it should be possible to infer some of the most common schemas that people bring to their everyday thinking. Critically, because the number of topics learned by the model is much lower than the number of discourse communities, the model should learn topics that are broadly shared across many different communities rather than topics that reflect individual discourse communities. I describe the topics learned by the model in more detail in the Results section. Descriptive statistics for the corpus are shown below in Table 6.

Measure	Frequency
Total Words	307,174,792
Total Sentences	13,346,004
Total Posts	2,253,975
Unique Authors	1,953,324
Unique Subreddits	115,293

***Table 6. Descriptive Statistics of the Reddit Corpus** used to infer common schemas. The corpus is based on every post to Reddit in the month of January 2017.*

3.2.3. Applying the Learned Schemas: Using Topic Models to Infer Schema Usage as a Cognitive Process

Detecting schemas in text involves not only learning common schemas, but also capturing when these schemas are being applied to a situation. Here, I describe a method to detect when schemas are being applied based on the idea that schemas are thought to fill in missing information in a situation. I describe a method for using neural networks, in combination with the schemas learned by a topic model, to capture this filling-in process.

Most accounts of schemas argue that people use schemas to fill in missing information in a situation. For example, Shank and Abelson's (1975) Script Applier Mechanism (SAM) was developed to take a description of a scene as input and use hand-specified schemas to help answer questions about information not explicitly provided. For example, given a story about a man ordering food from a restaurant, SAM could infer that a waiter brought the food. The notion that schemas are used to fill in missing information was also central to Bartlett's (1920) account of schemas, where people retelling stories progressively retold the story in a way consistent with their existing schemas (see also Kintsch & Greene, 1978). Together, these studies suggest that a strong way to measure schema usage is to capture the way schemas are used to fill in missing information in a situation.

This type of filling-in task is frequently used in machine learning, although it is not currently understood as a measure of schema usage. For example, state-of-the-art language models such as BERT (Devlin et al., 2018) and ELMO (Peters et al., 2018) are trained in large part by providing the model with part of a sentence and asking the model to fill in a plausible missing word in the sentence. Additionally, topic models are sometimes evaluated using a word prediction task where the model needs to predict the next word in a sentence (e.g. Wang et al., 2017). Finally, next word prediction is often used as the training objective for generative models of language (e.g. Radford, Narasimhan, Salimans, & Sutskever, 2018). Together, these training procedures suggest that many classes of machine learning models can effectively fill in missing information in a sentence.

As shown in Fig. 6, my key idea is to combine machine learning approaches to word prediction with the idea that schemas fill in missing information. I train a neural network to take posts in the Blog Authorship Corpus as input, holding out the last sentence in the post. Critically, the neural network represents the input only as a set of schemas, since the input is provided to the model in the form a distribution over the 500 schemas learned by the Latent Dirichlet Allocation Model, as described in the Method section 3.3 below. The task of the neural network is to predict the words written in the unseen last sentence of the post using only these schemas. To the extent to which this prediction is possible, the neural network is using the schemas evident in the rest of the post to fill in information about words that come next. Thus, the relative ability of the model to perform this prediction is a proxy for schema usage in the blog posts. In the Results section, I describe a preliminary test evaluating this procedure.

In summary, I develop a machine learning method to measure the application of schemas based on the idea that schemas should predict missing information in a sentence. This prediction is operationalized by a neural network using Latent Dirichlet Allocation topics to predict unseen words in the next sentence.

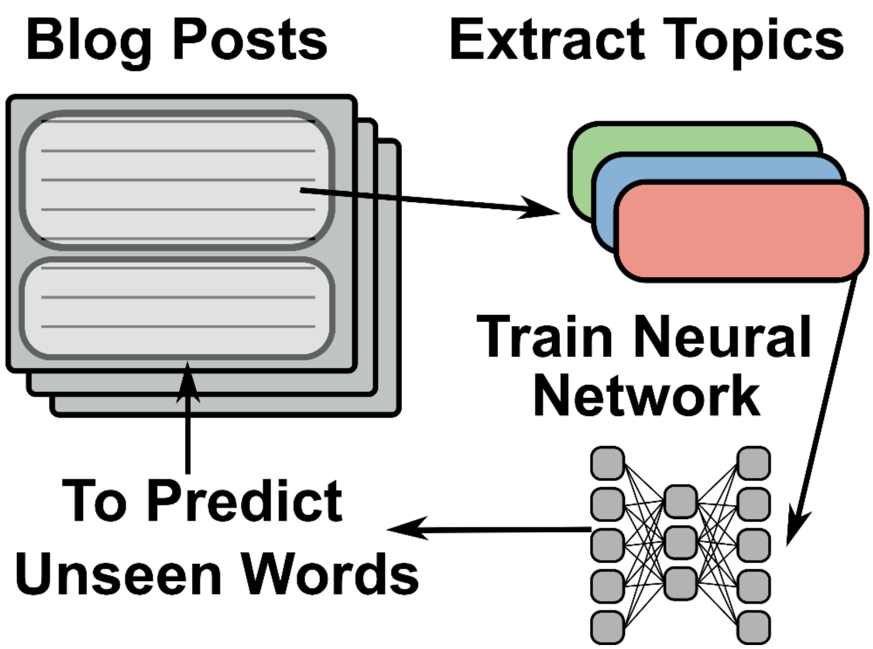

Fig. 6. Approach for Inferring Schema Usage in the Blog Corpus. *The approach is based on holding out all but the last sentence of a blog post, and using the trained topic model to identify the schemas (N=500) revealed in this text. A neural network is then*

trained to use these schemas to predict the words in the held-out last sentence of the blog post, separately for sentences about the past, present, and future.

3.3 Methods

As a brief overview and as described in Approach section 3.2, the methods involved training a topic model to identify schemas in a large corpus of social media posts. I then trained a neural network to detect the application of these schemas by using these schemas to predict missing words in the blog posts, separately for posts about the past, present, and future.

Learning Schemas: Corpus. I accessed a corpus of social media posts based on every post on the social media website Reddit in January 2017. The corpus was accessed via a publicly available repository (Baumgartner, 2019). In total, the corpus includes 307,174,729 words in 13,346,004 sentences in 2,253,975 posts. Text preprocessing involved removing URLs and special characters with regular expression-based matching and ignoring posts marked as removed or deleted by the Reddit repository. Additionally, because topic models learn to identify topics in part based on the topic of the document as a whole, I removed extremely short posts defined as posts containing only a single sentence. I retained only the most common 10,000 words in the posts, a constraint needed computationally to limit the size of the matrices learned by the model. I also removed extremely frequent words such as *the* or *and*, defined as words occurring in more than 5% of all posts. Finally, to evaluate the consistency of the topics over time (described in Section 3.3.4), I also accessed all posts on Reddit from February 2017 using the same repository (Baumgartner, 2019).

***Learning Schemas: Topic Model*.** I trained a Latent Dirichlet Allocation model (Blei, 2003) to identify the 500 most common topics in the Reddit corpus. There are three parameters in the model: *k* (the number of topics) which was set to 500, and two parameters controlling the shape of the probability distributions for the latent topics: *alpha* and *beta*. I set each of these parameters to 0.002 following the convention of 1 / number of topics. Setting these parameters equal to each other represents no strong prior assumption about the number of topics present in any given document. The model was trained for 100 passes over the corpus using the Python library *genism*, which uses a variational Bayes algorithm for model training (Hoffman, Bach, & Blei, 2010). The trained model was evaluated by visual inspection of the semantic coherence of the topics, and I found in pilot studies that 500 topics provided a large enough selection of topics while maintaining the semantic coherence of these topics.

***Applying Schemas: Next Word Prediction*.** Having trained a topic model, I next measured people's use of schemas in the Blog Authorship Corpus using a next-word prediction task. First, for each blog post, I found the set of schemas that was most closely associated with every sentence in the blog post except the last sentence. The result was a 500-dimensional vector where each element in the vector represents the probability that a post corresponds to a particular topic. Next, I trained a neural network to use the topics revealed in these posts to predict the words in the held-out last sentence of the post. As a preprocessing step, I retained only the most 5,000 common words in the held-out last sentence in order to reduce the dimensionality of the prediction task. The architecture of the network was based a single fully connected hidden layer with 200 neurons and a RELU activation function (Krizhevsky, Sutskever, & Hinton, 2012). During training, I randomly inactivated 50% of the units in the hidden layer (a procedure known as dropout) as a form of regularization (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014; for simulation studies on the effectiveness of dropout see Warde-Farley, Goodfellow, Courville, & Bengio, 2013). The output of the model was a prediction of the probability that the held-out word corresponded to each of the 5,000 most common words in the vocabulary (based on a softmax activation function, which converts activations to probabilities). The model was trained with a cross-entropy loss and ADAM optimization (Kingma & Ba, 2014). The model was trained for 25,000 training epochs with a batch size of 100 in the Python library *tensorflow* version 1. Visual inspection of the training loss revealed that this epoch size was sufficient for the model to converge. I created separate training and test sets by randomly holding out 10% of the posts in the blog corpus for model evaluation.

***Evaluation of Consistency of Topics*.** I conducted two tests to evaluate the consistency of the topics learned by the model. First, I evaluated consistency over time by training an identical topic model on posts from Reddit but in a different month: February 2017. All details of corpus acquisition and model training were identical to the procedure for the first model. Second, I evaluated consistency across corpora by training a topic model on posts from the blog corpus, again with an identical training procedure.

I also evaluated this consistency against two baselines. As a ceiling, I trained the same Reddit model twice on the same corpus of Reddit posts from January 2017. Any differences in these two models presents a noise ceiling of inconsistency due to random variations acquired over training the topic model. As a floor, I trained the Reddit model on a randomly created corpus by randomly shuffling the assignment of words to Reddit posts. Since LDA uses the distribution of words in documents to learn semantically coherent topics, it follows that the topics learned by this random model should be semantically incoherent (which I verified by visual inspection). Nevertheless, it is possible that a small amount of alignment between this random model and the model trained on real Reddit posts could be observed due to the fact that the alignment procedure, described below, can artificially create some similarity between any two models. Thus, this measurement captures a floor of similarity that would be expected between any two models due to chance.

To evaluate consistency across the models, I developed a procedure to align the topics learned by the two models and then to measure the consistency of these aligned topics. First, I aligned each topic learned by the Reddit model to the most similar topic learned by the control model. Because each topic is a probability distribution over the vocabulary, these topics can be compared with any vector-based distance metric, of which I chose Pearson correlation. I did allow a potential many-to-one mapping in the sense that more than one topic in the Reddit model could be most similar to a single topic in the control model. Next, I averaged the resulting Pearson correlations to calculate a single score ranging from 0 to 1 representing the similarity between the two models. A high correlation indicates that the models have learned similar topics in the sense that words that are highly probable in one topic are also highly probable in an analogous topic in the second model. In preliminary studies, I obtained similar results using other distance metrics, but Pearson correlation was preferred due to the fact that each topic contains relatively few highly probable words, a feature captured by the sensitivity of Pearson correlation to the squared magnitude of distances. Visual inspection also revealed that this procedure recovered reasonable alignments between the models by aligning semantically similar topics.

Human Evaluation of Topic Model. To evaluate the semantic coherence of the topic model, I asked human raters to judge how semantically coherent the topics learned by the model were. As a control, I used the random model trained with the same procedure but on randomly shuffled documents from the Reddit corpus. N = 23 raters were recruited on Amazon Mechanical Turk and each presented with 36 topics total: 18 topics from the real topic model and 18 topics from the random topic model. Each topic was presented as a list of the top five most probable words in the topic. Participants were asked to rate the semantic coherence of each topic on a 1-5 scale.

3.4 Results

To preview the main result, I found that both thoughts about the past and future involved greater use of schemas than did thoughts about the present. Below, I first describe three tests used to evaluate the procedure for learning human schemas and detecting the application of these schemas in text. I find that the model learned schemas that meet several characteristics of schemas, including being semantically coherent, consistent across time, and sufficiently general to be consistent across different corpora. I also find that the procedure for detecting schema application is reasonable, as verified by testing the prediction accuracy based on real compared to random schemas. Finally, I describe the application of these methods to detect schema usage in temporal thoughts in the blog corpus.

3.4.1 Evaluation of Topic Model: Semantic Coherence

First, I asked whether the representations learned by the topic model met the characteristics of schemas. This step is necessary to ensure that what is being measured is reasonably similar to a schema. One characteristic of schemas is that the representations should be semantically coherent. For example, the words *cake, candle*, and *presents* from the birthday schema are highly semantically coherent. I found that the topic model learned highly semantically coherent schemas, which also spanned a wide variety of different types of concepts. Examples of the schemas learned by the model are shown in Table 7, and the full list of schemas is shown in Appendix 2. To illustrate the semantic diversity of the schemas, the model learned a schema about feelings including the words *feeling*, *feels*, *felt*, *pain*, *worse*, and *bad*. In addition to schemas about mental states, the model learned schemas about other semantic categories such as common actions such as reading (*read, reading, book, books, library)*, a schema about sounds (*hear, sound, sounds, audio, hearing, noise*), and a schema about healthcare (*care, health, insurance, letter, legal, medical*).

To quantitatively evaluate the semantic coherence of the schemas, I presented the schemas to 23 human raters on Mechanical Turk and asked participants to rate the semantic coherence of the schemas. As a control, the participants also saw schemas generated by a model trained with an identical procedure but on randomly shuffled versions of the Reddit posts. Because the topic model uses the distribution of words in documents to infer semantically coherent schemas, shuffling these words across topics should reduce the semantic coherence of the schemas. As shown in Fig. 7A, I found that the topics learned by the real model were rated as more semantically coherent than the topics learned by the semantically ablated model, $t_{(21)} = 11.68$, $p < 0.001$. As shown in Fig. 7B, this difference was observed in every individual rater (23 / 23 raters).

Topic Label	Words
Feelings	Feeling **|** Feels **|** Felt **|** Pain **|** Worse
Reading	Read **|** Reading **|** Book **|** Books **|** Library
Studying	Study **|** Subject **|** Passed **|** Studying
Food	Food **|** Eat **|** Eating **|** Healthy **|** Diet

Table 7. Topics Learned by the Latent Dirichlet Allocation Model**.** *Four topics learned by the Latent Dirichlet Allocation model, represented by a list of the most probable words in the topic. The words are learned by the model; the topic label is a human-generated semantic label.*

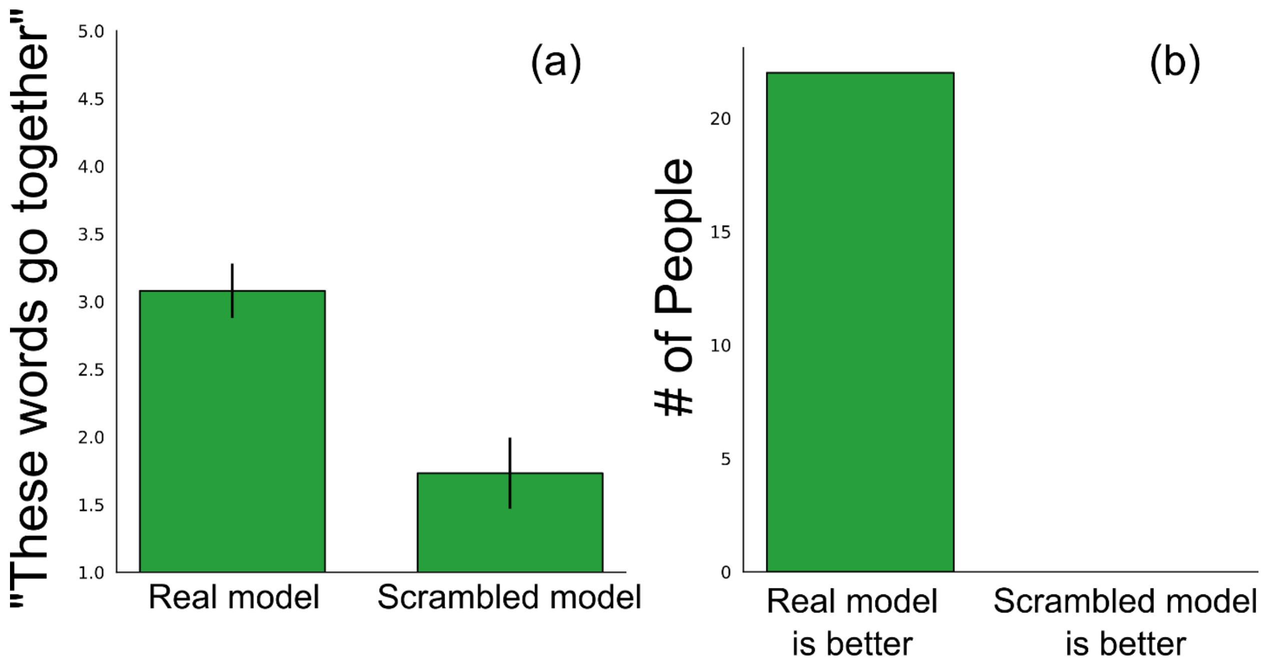

Fig. 7. Evaluation of topic model (A) Average human rating (1-5) of the semantic coherence of words in the real and scrambled topic models (+/- 95% bootstrapped CI). (B) Total number of raters who rated the topics from the real and scrambled models most highly.

3.4.2 Evaluation of Topic Model: Consistency of Schemas across Time and Corpora

I also evaluated whether the model learned schema-like representations by testing for two further characteristics of schemas. First, because schemas are broadly shared across people, schemas should change very slowly, if at all, over time. I tested for this property by training the same topic model on social media posts from two different months from the same corpus. Second, because schemas are broadly shared across people, the schemas learned should be general rather than specific to any particular corpus. I tested for this property by training the topic model on two different text corpora – Reddit posts and blog posts – and comparing the schemas learned. I also trained two baseline models. The best possible consistency of schemas is limited by the consistency that would be observed if the same model were trained twice on the same corpus, which is limited by some stochasticity inherent to the training procedure. This consistency is given by the top vertical dotted line. Additionally, the procedure I used to compare two models can induce a small amount of similarity between any two models, even if the models are not inherently similar, simply due to chance (see Methods section 3.3.4). I established this chance consistency by generating semantically incoherent topics by training the model on randomly shuffled documents of text. I then aligned these topics to the topics learned by the real topic model, plotted by the bottom vertical dotted line.

I found that the schemas learned by the model were strongly consistent across time, and were moderately consistent across corpora. I tested for correlation across time by training an identical topic model on Reddit posts from January vs. February 2017. As shown in Fig. 8, the correlation in topics observed across the two models (r = 0.44) was close to the maximum possible consistency, suggesting strong agreement of the topics over time. I also tested for correlation across corpora by training an identical topic model but on the Blog Authorship corpus rather than on Reddit posts. As shown in Fig. 8, there was still a moderate correlation in the topics learned by the two models even when these models were trained on different corpora, r = 0.32. Together, these results suggest that the topics learned are moderately consistent across corpora and strongly consistent across time. Consistency across the corpora and across time implies that the topic learning algorithm discovered packages of ideas that are likely independent of a particular training corpus. Such consistency suggests that the learning process discovered generic, widely shared packages of conceptual structure, as should be the case, assuming the process was able to learn schemas.

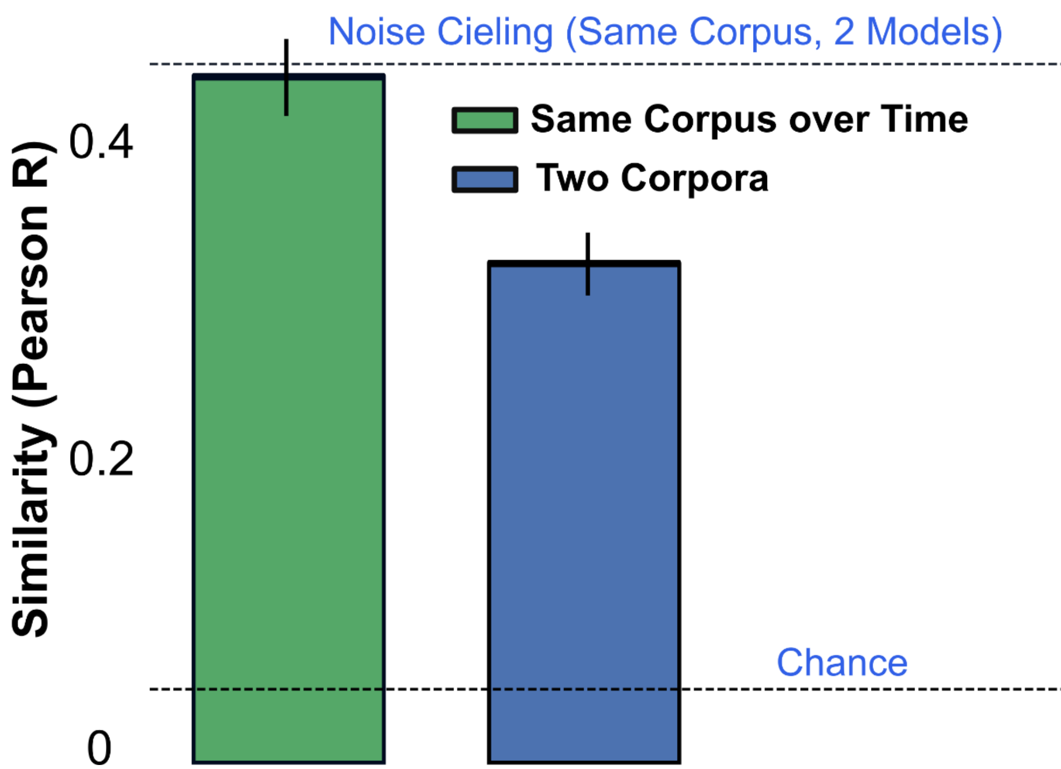

Fig. 8 Consistency of schemas across time and corpora. *The similarity of the schema model is plotted compared to the same model trained on Reddit posts from a different month (green bar) and trained on a different corpus (blue bar). The dotted lines represent the baseline of the same model trained twice on the same corpus (top dotted line) and chance alignment with a random model that is induced by the alignment procedure (bottom botted line). Vertical axis is similarity of models, represented as Pearson r. Errors +/- 95% bootstrapped CI.*

3.4.3 Evaluation of Method for Detecting Application of Schemas

My key idea is that people's use of schemas can be measured by training a neural network to use the schemas implicit in the words people just wrote to predict the words people will write next. This method replicates the property that schemas are thought to fill in missing information. To evaluate whether the model can in fact use schemas to fill in missing information, I trained two versions of this model: a neural network using schemas derived from un-randomized text and a neural network using false schemas derived from randomized text. While both models can learn to predict highly frequent words, only the model based on schemas derived from unrandomized text can also rely on schemas to fill in missing information. I therefore predicted that if the model can use schemas to fill in missing information, then the predictions of the model based on unrandomized schemas would be more accurate than the predictions of the model based on randomized schemas.

As shown in Fig. 9, prediction accuracy differed across the two models. The model trained using real schemas, derived from unrandomized text, made more accurate predictions than the model trained using random schemas as based on a lower cross-entropy loss of the model on held-out testing data, $t_{(220,174)} = 174.26$, $p < 0.001$. This result suggests that the model can in fact use schemas to fill in missing information.

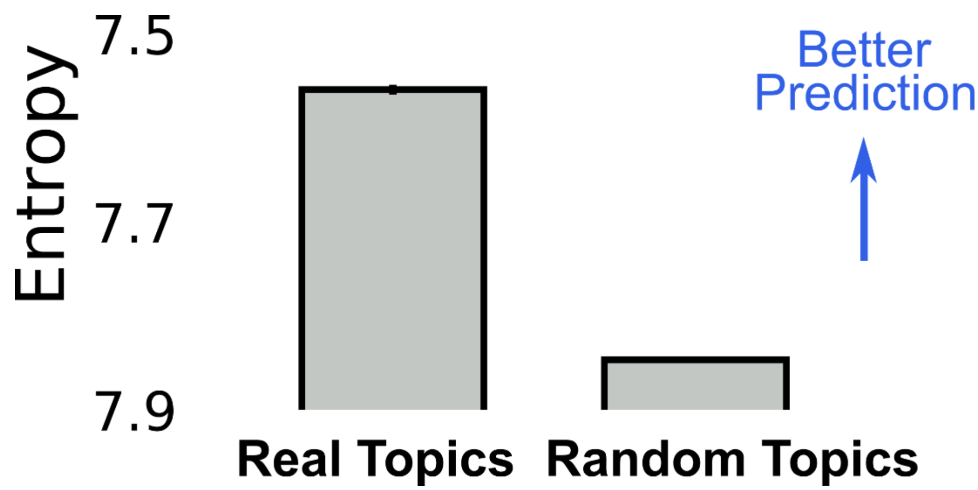

Fig. 9. Held-out word prediction accuracy. *Prediction accuracy of the neural network in predicting held-out words, when the model is provided with either real LDA topics (left bar) or randomly generated topics (right bar). The vertical axis plots error (cross-entropy loss) where a lower score represents better prediction. Error bars +/- 95% bootstrapped CI (the error for the random topics is too small to be visible).*

3.4.4 Schema Usage in Thoughts about the Past, Present, and Future.

Having established that the model can use schemas to fill in missing information, I next measured people's use of these schemas in the blog corpus separately for sentences about the past, present, and future. I did this by dividing the prediction results based on whether the sentence being predicted referred to the past, present, or future, using held-out testing data that was previously unseen by the model. As shown in Fig. 10, prediction accuracy differed by temporal class. The model was more accurate in using schemas to fill in missing information in sentences about the past compared to sentences about the present, $t_{(110,952)} = 4.60$, $p < 0.001$. The model was also more accurate in using schemas to fill in missing information in sentences about the future compared to sentences about the present, $t_{(154,144)} = 21.28$, $p < 0.001$. Thoughts about the present presumably do not involve mental time travel. Hence, these results suggest that schemas are a key ingredient in mental time travel into both the past and future. I also observed that the model was more accurate in using schemas to fill in information in sentences about the future compared to sentences about the past, $t_{(175,248)} = 18.53$, $p < 0.001$. Any interpretation of this result is speculative, but one possibility is that thoughts about the future may have relied more on schemas compared to thoughts about the past during the construction phase, but equally on schemas during the elaboration phase, due to the necessity to construct an event without access to a memory trace.

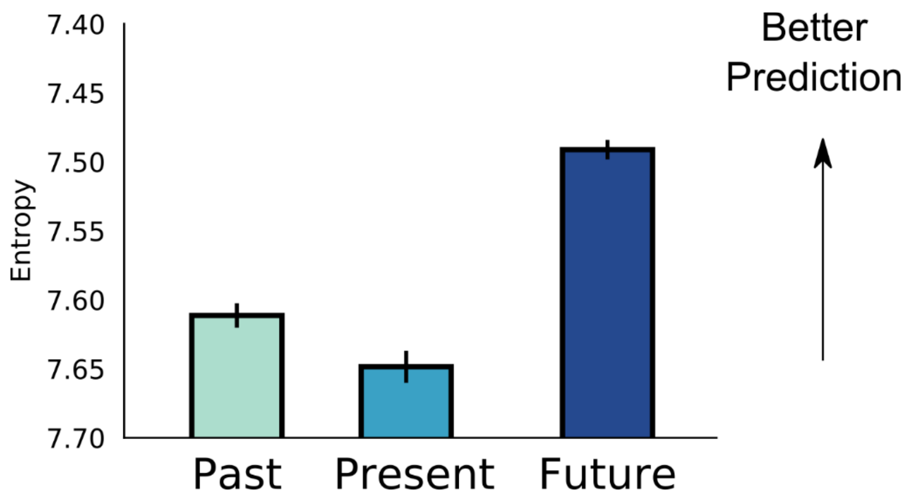

Fig. 10. Prediction accuracy by temporal class. *Prediction accuracy of the neural network in predicting held-out words, separately when the words being predicted come from sentences about the past, present, and future. Vertical axis plots error as cross-entropy loss, where smaller values represent better prediction. Errors +/- 95% bootstrapped CI.*

3.5. Discussion

In this chapter, I measured people's use of semantic past and future thinking. First, I developed a method for automatically identifying schemas based on training a topic model on a large social media corpus. Second, I developed a method for measuring schema usage as a cognitive process based on training a machine learning model to use these schemas to predict the words people subsequently wrote next. Third, I found that both thoughts about the past and future involved more use of schemas than did thoughts about the present, as evident by the increased ability of the model to use the schemas evident in people's writing to predict words in sentences about the past and present compared to words in sentences about the future.

One of the main contributions of this study is developing a large-scale model of people's everyday schemas. As described in the Approach, there is some history of arguing that distributional semantic models such as topic models can effectively model the human conceptual system (Griffiths et al., 2007; McRae & Jones, 2013). Despite the theoretical success of these models, a challenge has been to train a distributional semantic model on a large and diverse enough corpus that the representations learned can be argued to capture a large class of people's everyday schemas. In this study, I capitalized on the increasing availability of large-scale, diverse corpora on social media to train a topic model to capture a wide variety of schemas. Here, I applied these schemas to study past and future thinking, but a variety of applications should be possible. For example, while I found that schemas could predict the content of people's references, there were sentences for which the schemas predicted the words that followed worse than would be expected by chance. These cases could indicate references (or even individuals) who are less schema-driven or even deliberately violate pre-existing schemas in their writing. Additionally, I trained the topic model on a corpus of English-language writing, but because the methods used are language-agnostic, it should be possible to train several different topic models on language from different cultures to compare common schemas cross-culturally. Finally, while the words in the schemas were derived automatically, I had to manually provide a label for the topic recovered by the schema. Future work could explore automatic approaches to labeling schemas, a direction that could help understand the large-scale conceptual information being revealed by these schemas (for initial approaches to automatic topic labeling, see Magatti, Calegari, Ciucci, & Stella, 2009; Mei, Shen, & Zhai, 2007; Lau, Grieser, Newman, & Baldwin, 2011).

The other main contribution of the study was the finding that both past and future thinking rely more on schemas than do thoughts about the present. One of the driving ideas in the study of future thinking has been that thoughts about the past and future may rely on common cognitive processes (Schacter & Addis, 2007), a finding supported by a wealth of neuroscientific (Benoit & Schacter, 2015; Viard et al., 2011) and behavioral evidence (Rubin, 2014; Spreng & Levine, 2006). However, it is typically assumed that episodic processing is one of the key, if not the most important, shared process between past and future thinking. My results suggest a different interpretation. What may be most shared between past and future thinking is the need to rely on schemas. When people recall past events, a wealth of literature supports the use of schemas (Bauer, 1993; Brewer & Treyens, 1981; Fivush, 2002; Hintzman, 1986). My results suggest that schemas are at least as important when people imagine the future, perhaps because schemas provide a scaffolding for people to imagine particular types of events where they can fill in episodic details. Thus, my results suggest that the use of schemas may be the key shared cognitive process between past and future thinking.

One unexpected finding was that while schemas were involved in both thoughts about the past and future, schemas were more involved in thoughts about the future than the past. I first describe some related results in the literature, and then offer an interpretation of this finding. I am not the first to find that some component processes of mental time travel are more active for thoughts about the future than thoughts about the past. Several studies have found that the hippocampus, typically involved in episodic memory, may be even more active for thoughts about the future than the past (Addis et al., 2007; Addis & Schacter, 2012; Buckner, 2010). Indeed, one of the original fMRI studies of future thinking found a set of default network regions including the inferior frontal gyrus, middle temporal gyrus, and hippocampus that were more active during the construction of future events than recall of past events (Addis et al., 2007). Any interpretation of these results is speculative, but one possibility is that future thoughts draw more on schemas than past thoughts during the construction phase, but equally on schemas as past thoughts during the elaboration phase. Because future events have not happened, initially constructing a future event may be a largely schema-driven process. Supporting this interpretation, when people are asked to think aloud while imagining the future, people almost always begin by invoking a schema before generating specific episodic details (D'Argembeau & Mathy, 2011). Thus, while any interpretation is speculative, the difference in schema usage between past and future thinking may be due to the initial construction phase.

While I found that the model of schemas captured several key features of human schemas, the model does have some limitations. Chief among these limitations, it is widely held that people's schemas are hierarchically structured (Cooper & Shallice, 2006; Schank & Abelson, 1975). By contrast, the topics learned by the model have no such hierarchical structure. There are hierarchical extensions of the topic modeling procedure used in the study, and an interesting future extension would be to learn hierarchical versions of these schemas (Griffiths et al., 2007; Kataria, Kumar, Rastogi, Sen, & Sengamedu, 2011). Additionally, while I found a moderate degree of consistency in the schemas when the model was trained on different corpora, the model did learn some schemas that may be specific to certain contexts, such as a schema about common video games that may be particular to social media. Indeed, topic models are sometimes trained on very specific corpora with the goal of learning representations that are specific to that corpus, such as topics in American politics (Rule, Cointet, & Bearman, 2015) or news articles. Future work could explore training a topic model on multiple different large corpora, potentially with a penalty for learning topics that are over-represented in any single corpus. Such a procedure could potentially learn even more generalizable schema-like representations than the representations in the current study.

In summary, I developed a method for automatically deriving common schemas and detecting the usage of these schemas in people's writing about the past, present, and future. The main result was that past and future thoughts drew more on these schemas than thoughts about the past, as evidenced by an increased ability of a neural network to use the schemas people had just written about to predict the words they wrote next. This result may suggest that the key similarity between past and future thoughts is not episodic processing, but rather is the use of schemas.

Chapter 4: Temporal Thoughts in the Lab

4.1 Introduction

In Chapters 2-3, I studied people's naturally occurring, unprompted thoughts about time. While the results were based on a large and diverse corpus, temporal thoughts are often studied in a different way: by eliciting temporal thoughts in response to prompts in a lab. As previously discussed, in prior literature, thoughts about time are typically evoked with very specific prompts that emphasize the construction of episodic past and future thoughts. Participants are typically instructed to imagine single events that occur at a specific time instead of being extended over time, are highly likely to occur to them and occur in a specific spatial location (e.g. Addis & Schacter, 2008; Addis et al., 2011). It is currently unknown whether this method of eliciting temporal thoughts can affect how people think about time. Given that the language-based methods I developed for studying episodic and semantic processing can naturally be extended to other kinds of temporal thoughts, in this chapter I ask whether similar results would be obtained by studying temporal thoughts evoked in the lab.

4.2 Approach

4.2.1 Eliciting Temporal Thoughts in the Lab: the Future Crovitz Task

One way to study the contents of people's memories and future thoughts is to prompt participants with a variety of different cue words and then ask people to retrieve memories and future thoughts in response to these cues. The key idea is that by using a variety of different cue words, a relatively unbiased picture of people's memories and future thoughts should emerge. This idea was originally formulated as the Crovitz task used to study episodic memory (Crovitz & Schiffman, 1974). In the original Crovitz task, participants are prompted with a single cue word and asked to use the cue to retrieve an episodic memory. The paradigm was originally used to study the frequency of episodic memories as a function of temporal distance (Crovitz & Schiffman, 1974).

The basic idea of eliciting memories in response to prompts can be extended to simultaneously study past and future thinking. In the literature, this is known as the Future Crovitz task, and it is the most common way to study past and future thinking in the lab (Spreng & Levine, 2006; Bertossi, Tesini, Cappelli, & Ciaramelli, 2016; Szpunar et al., 2007; Brown et al., 2014). As shown in Fig. 11, in the modified Future Crovitz task, participants receive single word cues, one at a time. On half of the trials, participants are prompted to use the cue to remember a past event, and on the other half of the trials, participants are prompted to use the cue to imagine a future event. Because the procedure for eliciting past and future events is the same, any differences observed should indicate differences in the underlying processes of past and future thinking. I use a version of the Future Crovitz task based on cue words from Rubin (1981), which have been normed and matched for imagery and familiarity. Participants completed 30 trials in total: 15 trials where they remembered past events and 15 trials where they remembered future events.

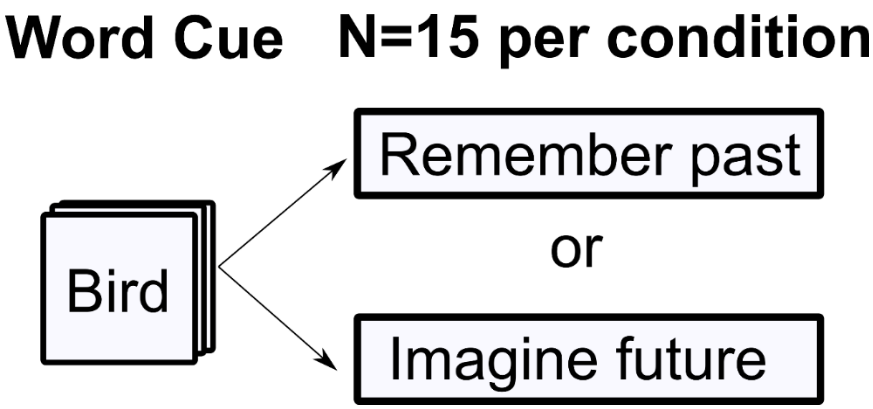

***Fig. 11. Procedure for eliciting temporal thoughts in the lab*.** In the modified future Crovitz task, participants see a series of cue words, 1 at a time (e.g., *bird*). In a block design, participants are instructed to either retrieve a past event in response to that cue (N = 15) or imagine a future event in response to that cue (N = 15).

4.2.2. Measuring Episodic and Semantic Temporal Thinking

A good method for understanding how our results might differ if temporal thoughts were elicited in the lab should vary only the method of eliciting temporal thoughts while using as similar as possible a method for measuring episodic and semantic processing in language. For this reason, I adapted the language-based measurements of episodic and semantic thinking from Chapters 2-3 to measure episodic and semantic processing in the lab. In the case of episodic thinking, I used the same keyword-based markers of spatial, perceptual, and concrete language to identify linguistic markers of episodic thinking. In the case of schema-based thinking, I created an analogue of the idea that the schemas revealed in people's thoughts could predict the words that they say next. In the Crovitz task, because people's responses are quite short, it is not feasible to use the first half of the post to identify a schema. However, since most thoughts are related in some way to the cue word, the cue word can be thought of as providing the schema. Indeed, a similar logic inspired an *event clustering* approach to identifying schemas in future thoughts, where people's language was coded for schema-like relations to the cue words (Demblon & D'Argembeau, 2014). For this reason, I used the word cues to identify a schema (based on the already trained topic model) and asked the pretrained neural network to use this schema to predict the words in people's thoughts about the past and future. I also report an unsuccessful attempt to learn these schemas directly from people's talk in the Crovitz task.

4.3 Methods

***Participants*.** 74 undergraduate participants were recruited from Emory University and participated in exchange for course credit.

***Materials*.** Two lists, each composed of 15 single-word cues, were generated from a list of word cues from Rubin (1981), with the constraint that the lists be as matched as possible for human ratings of imagery and similarity. All of the cues selected were also used in the Spreng and Levine (2006) study of future and past thinking. For each participant, I randomly assigned one of the lists to be cues for the past condition and the other list to be cues for the future condition. The first list consisted of the words *blossom, money, hide, trouble, engine, month, star, rattle, plant, power, horse, ship, window, warmth,* and *clothing*. The second list consisted of the words *girl, excuse, mother, bird, capacity, kindness, street, mountain, virtue, errand, flower, salad, door, table*, and *menace*. These lists did not differ in either imagery or familiarity (all $t < 1$).

The average imagery rating of the words was 0.50/1, and the average familiarity rating was 0.60/1.

***Crovitz Task*.** Participants were instructed that they would see simple word cues, one at a time, and that their task was to produce a “single and specific personal [past/future] event” in response to each cue. The full instructions are available in Appendix 3. Participants then completed two blocks in counterbalanced order. In the *past* block, participants generated a past event in response to each of the 15 single-word cues. In the *future* block, participants generated a future event in response to each of the 15 single-word cues.

***Analysis of Episodic Language*.** Each event was scored for three episodic language markers using identical methods to Chapter 2: spatial language, perceptual language, and concrete language. These scores were averaged across all events in a block to generate a single score for each participant for each of the past and future blocks.

***Analysis of Semantic Language*.** The method from Chapter 3 was adapted to analyze semantic language in the Crovitz task. For each trial, the word cue was inputted to the pretrained topic model described in Chapter 3 to identify the schemas present in the cue. The result was a 500-dimensional vector of the probability that the cue represented each of the 500 topics learned by the model. Next, the pretrained neural network from Chapter 3 was used to take the schema as input and predict the probability of each individual word generated in the Crovitz task. This pretrained neural network was necessary because there was insufficient data to train the neural network directly on the Crovitz task. For each participant, I averaged the error of the model, represented as categorical cross-entropy, across all trials, separately for the past and future blocks.

Topic Model. One of the major contributions of Chapter 3 was to develop a large-scale model of people's common schemas. The use of social media provided a large-scale corpus for training this model, but it is possible that less social-media specific topics could be learned using the data generated in the Future Crovitz task. On this basis, I trained a topic model using the temporal thoughts elicited in the Crovitz task. All model parameters and training procedures were identical to Chapter 3.

4.4 Results

To preview the main findings, the results elicited from the lab were highly similar to those elicited from blogs and social media in Chapters 2 and 3. As in Chapter 2, I found that thoughts about the past relied more on episodic language than did thoughts about the future. As in Chapter 3, I found that thoughts about the future relied more on schemas than did thoughts about the past. I also investigated training a topic model only on the Crovitz task, where I found that the model did not learn semantically coherent schemas.

4.4.1 Episodic Processing in Lab-Elicited Temporal Thoughts

People's responses in the Crovitz task were automatically coded for three markers of episodic language: spatial, concrete, and perceptual language. This coding was performed separately for trials where participants remembered the past and imagined the future. As shown in Fig. 12, I found that thoughts about the future were less episodic than thoughts about the past for all three measures of episodic processing, although only two of these differences reached statistical significance. Thoughts about the future were less concrete (*M* = 2.72/5 rated concreteness) than thoughts about the past (*M* = 2.77/5 rated concreteness), $t_{(72)} = 2.46$, $p = 0.016$. Thoughts about the future were also less perceptual (*M* = 0.064 perceptual words) than thoughts about the past (*M* = 0.074 perceptual words), $t_{(72)} = 2.32$, $p = 0.023$. Finally, thoughts about the future were marginally less spatial (*M* = 0.08 spatial words) than thoughts about the past (*M* = 0.09 spatial words), $t_{(72)} = 1.96$, $p = 0.054$. Together, these results suggest that people rely less on episodic processing when they imagine the future than when they remember the past, and that these effects are not specific to asking people to generate naturally occurring temporal thoughts in the blog corpus.

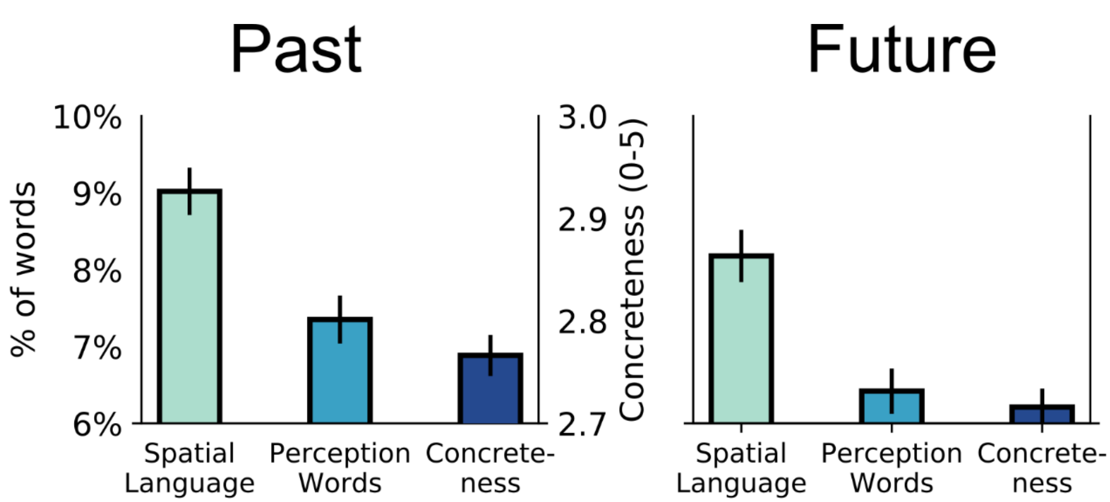

Fig. 12. Episodic language in lab-elicited temporal thoughts. *Spatial and perceptual language are measured as percentage of total words (left vertical axis) and*

concreteness is measured as an average of 1-5 ratings (right axis). Error bars bootstrapped 95% CI.

4.4.2 Semantic Processing in Lab-Elicited Temporal Thoughts

An analogue of the filling-in procedure from Study 3 was developed to measure schematic processing in the Crovitz task. For each trial, the schema evident in the cue word was used to predict the words in people's imagined events, separately for events about the past and future. As shown in Fig. 13, the main result was that schemas were significantly more effective at predicting words in references to the future (*M* entropy = 7.61, where low entropy is better prediction) than the past *(M* entropy = 7.92), $t_{(7069)}$ = 3.91, $p < 0.001$. This result suggests that people more on schematic processing when they imagine the future than when they remember the past.

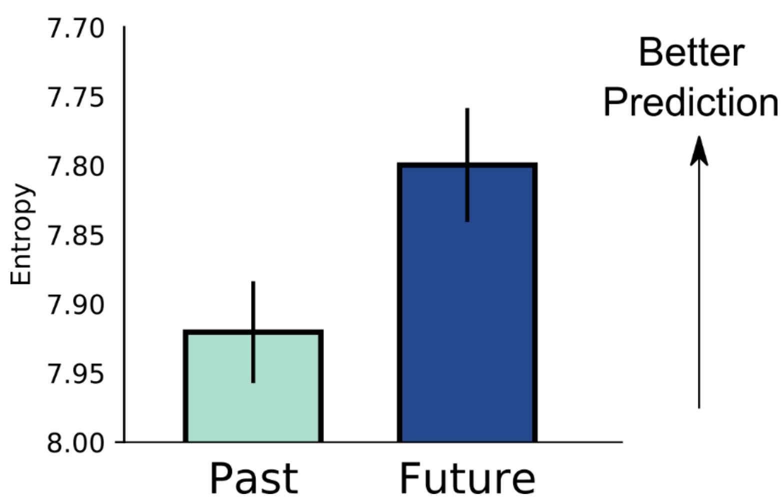

Fig. 13. Semantic processing in lab-elicited temporal thoughts. *The vertical axis plots prediction error as categorical cross-entropy, where a lower number represents better prediction. Error bars 95% bootstrapped CI.*

4.4.3 Lab-Based Topic Model

In Chapter 3, I trained a topic model on a large corpus of social media posts to learn some of people's most common schemas. While I found that these schemas were moderately consistent across time and across corpora, it is possible that better schemas could be learned using the temporal talk elicited in a more controlled laboratory environment. For this reason, I trained a separate topic model using the events people generated in the Crovitz task. However, I found that the model did not learn semantically coherent schemas. Some examples of schemas learned by the model include *physics*, *calculate*, *joy*, *photos*, *hang*, *roommates* and *hiding*, *one*, *better*, *margarine*, *research*, *figure*. The full list of schemas learned by the lab-based topic model is available in Appendix 4. The inability to learn semantically coherent schemas was most likely due to the relatively small dataset generated in the Crovitz task. Investigating the impact of semantic knowledge on thinking might require, then, access to extremely large bodies of text. The results point to how research using large bodies of text may open up areas of exploration that are not otherwise accessible using traditional data collection methods.

4.5 Discussion

In Chapters 2-3, I studied people's future thinking in a large-scale natural context based on a corpus of blog posts. Here, I asked whether similar results would hold if people's thoughts about the past and future were evoked in a more traditional way in the lab. Participants completed a modified Crovitz task where they generated past and future events in response to simple word cues. The episodic and semantic language evident in these events was measured using very similar techniques to Chapters 2-3. I found largely similar results to the previous chapters, namely that people's talk about the past used more episodic language and less schematic language than their talk about the future. Together, these results suggest that my results are not merely due to the use of natural, unprompted language.

While these results suggest that schemas rather than episodic processing may represent the common process for past and future thinking, these results are not without precedent in prior literature. First, when past and future events are coded for sensory details such as vividness, past thoughts are sometimes reported to be more vivid than future thoughts. D'Argembeau and Van der Linden (2006) found that thoughts about the future relied on fewer visual details and had less spatial context than thoughts about the past. Rasmussen and Berntsen (2013) found that future events had a lower sense of reliving, were less vivid, and involved less sensory imagery than past events. Second, there is some precedent for the idea that future thoughts may rely more on schemas than past thoughts. Rasmussen and Berntsen (2013) found that future thoughts relied more on a subset of schemas, cultural life scripts, than did thoughts about the past. As previously discussed, several studies of patients with semantic dementia also suggest that impairment to the conceptual system may have a stronger impact on thoughts about the future than the past, although these studies have a number of limitations (Irish et al., 2012; Irish & Piguet, 2013; Irish & Piolino, 2016).

While I did find that thoughts about the past were more episodic than thoughts about the future, the effect was statistically marginal for one episodic language measure: spatial language. One possible interpretation of this result could be that the most important differences between past and future in episodic processing are with respect to the vividness and perceptual imagery of past and future thoughts. However, in Study 2, I did find that the numerically largest difference in past and future thinking was with respect to spatial language, with an effect size greater than four-tenths of a standard deviation. Thus, caution should be used in interpreting any lack of effect for spatial language specifically in the lab, which could also be explained by a smaller sample size involving fewer trials.

In Chapter 3, I trained a large-scale model of people's everyday schemas based on the words they write on social media. In this Chapter, I also explored training a similar model based on the language people write in the lab. I found that the schemas learned by the model in Chapter 3, based on social media, were more semantically coherent than those schemas learned in the lab. Based on this finding, I used the model trained on large-scale social media data to analyze people's behavior in the lab. This finding may point to a broader strategy for combining research on large-scale everyday behavior with laboratory designs. Specifically, it may often be possible to train large-scale classifiers on large everyday datasets, and then transfer the model to people's behavior in the lab. A similar strategy could be useful for detecting mental illness based on people's language (Thorstad & Wolff, 2019; Eichstaedt et al., 2018) or using language to predict people's decision-making (Thorstad & Wolff, 2018).

Chapter 5: General Discussion

5.1 Summary of Findings

This dissertation asked three open questions about how people think about the past and future. First, is episodic processing uniquely associated with past and future thinking compared to thinking about the present? I found that episodic processing was not uniquely associated with past and future thinking. Instead, I found that the episodic processing in past thoughts was more similar to present thoughts, which do not involve mental time travel, than to future thoughts. Second, does thinking about the future involve schemas, and if so, third, is the use of schemas uniquely associated with past and future thinking compared to thinking about the present? I found that thinking about the future does involve schemas, and that the use of schemas was uniquely associated with thoughts about the past and future compared to thoughts about the present. Together, these results suggest that past and future thinking rely on common cognitive processes in the use of schemas, but not episodic processing.

I studied how people think about the past and future by extracting millions of naturally occurring temporal references. I did this by using an automated tool for extracting temporal references from language, based on a series of syntactic and lexical rules (Copley & Wolff, in prep). Using these rules, I was able to identify more than 5 million references to the past, present, and future from over 19,000 individuals in a corpus of web blog posts. A key feature of these temporal references is that unlike in previous research where temporal references were elicited in response to prompts, these temporal references were naturally occurring and unprompted.

I developed methods for extracting cognitive processing from the language in these temporal references. First, I extracted episodic processing by mining the temporal references for indicators of concrete, perceptual, and spatial language. Second, I extracted schema-based processing by learning a large-scale model of the kinds of topics that people talk about every day on social media. I then developed a procedure to identify the filling-in process characteristic of schema usage by training a neural network to identify the words people wrote in temporal references based on the schemas revealed in their prior writing. These new methods allowed me to go beyond prior literature. In particular, there are few existing methods for identifying schemas in past and future thinking, although there are some approaches based on manual coding or studies of neuropsychological patients. Using these new methods, I was able to automatically identify the schema usage in millions of naturally occurring thoughts about time.

Finally, I compared the way people naturally talk about time to the way people talk about time in a lab context. I did this by eliciting temporal thoughts in the lab using simple word cues. I then mined the language elicited for the same episodic and semantic indicators used to study millions of temporal thoughts in the blog corpus. I found largely similar results, suggesting that the results are not specific to studying naturally occurring thoughts about time.

5.2 Implications of Big Data for Cognitive Science

Some of the techniques used in this research are newly emerging in cognitive science. Here, I highlight two of these techniques that may have broader implications.

5.2.1 Learning Common Human Schemas

Schemas have a long history in psychology and artificial intelligence. However, to date there has been no large-scale model of the schemas that are broadly shared across people. Here, I describe how the approach to learning human schemas in this dissertation makes progress towards such a model by combining the strengths of approaches in artificial intelligence and psychology.

Central to symbolic approaches to artificial intelligence was an attempt to specify the full range of human schemas, largely by hand-coding common schemas using human intuition (Minsky, 1974; Schank & Abelson, 1975). These attempts continue today, in resources such as the ConceptNet database (Liu & Singh, 2004; Speer, Chin, & Havasi, 2017), BabelNet (Navigli & Ponzetto, 2012), MENTA (de Melo & Weikum, 2010) and WikiNet (Nastase, Strube, Börschinger, Zirn, & Elghafari, 2010). One of the strengths of these artificial intelligence approaches is their scale, which can include many thousands of schemas. A key weakness, however, is that the schemas must be specified based on human intuition, which limits the schemas that can be discovered to those schemas that are readily intuited (for a similar argument about the limitations of intuition in psychology, see Lilienfeld, 2010).

In Psychology, schemas have typically been studied by specifying a few dozen of the most common schemas based on human data, such as a set of cultural life scripts (Berntsen & Rubin, 2004). A strength of these approaches is going beyond any individual person's intuition by synthesizing human data to extract the most common schemas. However, a weakness of this approach is that it has historically been limited to uncovering small categories of schemas such as cultural life scripts, in part because of the difficulty of extracting enough human data to learn more general categories of schemas.

A strong approach to uncovering human schemas would combine the strengths of the approaches from artificial intelligence and psychology. Such an approach is present in the model of human schemas trained in Chapter 3. Like artificial intelligence databases, a topic model can learn a large number of human schemas. Because the model is trained on a large social media corpus with millions of different posts on over 100,000 different topical forums, the schemas learned by the model are likely also very general and broad in the types of semantic categories they cover. However, whereas one of the weaknesses of the artificial intelligence databases (and a strength of some approaches in Psychology) was the requirement that schemas be specified by hand instead of learned, the topic model used in the current studies was able to learn common human schemas directly from human language. By combining the scale of artificial intelligence approaches with the data-driven approach in some psychological approaches, the topic model was able to learn schemas that at least approximate a sizable portion of the human semantic space.

5.2.2 Identifying Naturally Occurring Temporal Thoughts

Thoughts about time are typically extracted in the lab, yet people mentally time travel many times a day in their everyday life. A key challenge in understanding this everyday mental time travel is developing a method to identify these everyday temporal thoughts from language. At the scale of millions of temporal references, an automated method is needed to identify these temporal references. The current study evaluated several of these automated methods against a common standard as well as against human ratings, finding that two automated methods are capable of extracting temporal thoughts with near human-like accuracy.

In most existing studies, temporal thoughts are elicited in the lab. In some of these studies, temporal thoughts are elicited in an interview paradigm based on the autobiographical memory interview (Kopelman, Wilson, & Baddeley, 1989; Cole, Morrison, & Conway, 2013; Addis, Sacchetti, Ally, Budson, & Schacter, 2009b; De Brigard & Giovanello, 2012). In other studies, an episodic recombination paradigm is used where participants generate several past events, and then are asked to recombine features of the different events to generate a novel future event (Addis et al., 2009a; De Brigard, Addis, Ford, Schacter, & Giovanello, 2013). While some studies have explored eliciting thoughts about time outside the lab using experience sampling (Baumeister et al., 2018; Barsics, Van der Linden, & D'Argembeau, 2016; Shepard, Nie, Copley, & Wolff, 2017; Beaty, Seli, & Schacter, 2019), these methods have been limited in their scope and often still rely on prompting participants at various intervals throughout the day to retrieve temporal thoughts.

Outside of the lab, there are several approaches for automatically identifying temporal thoughts from language. However, these methods are not used in existing literature, in large part because they have not been evaluated against a common benchmark of human ratings. For this reason, it is unknown how well these algorithms perform compared both to each other and to human ratings. Schwartz et al. (2015) proposed a method for using decision trees trained on a variety of linguistic features to classify temporal references. Chang and Manning (2012) developed a method for identifying temporal references, such as dates, that can be extended to identify temporal references. Pennebaker et al. (2015) developed a keyword-based dictionary, LIWC, with categories for identifying past, present, and future references. Finally, Copley and Wolff (in prep.) developed an approach combining syntactic rules with lexical items.

I evaluated these automated temporal reference classifiers against a common standard of human ratings. I found two main results. First, I found that only two of these models performed above chance: the Copley and Wolff model and the LIWC model. While both of these models performed above chance, I found that only the Copley and Wolff model performed well for all classes, while the LIWC model struggled to identify references to the future. Second, I found that when humans are asked to classify sentences as about the past, present, and future, these raters often disagree. This disagreement may be due to the fact that some sentences contain more than one kind of temporal reference. Supporting this interpretation, experience sampling studies have found that when people think about the past, people often report thinking about the implications of the past for the future (Baumeister et al., 2018). Based on these two main results, my findings suggest that the best current model for automatically identifying temporal references is the Copley and Wolff model.

5.3 Limitations and Future Directions

The current studies have a few limitations, primarily arising from the use of novel corpus-based and machine learning techniques for studying past and future thinking.

5.3.1 Limitations of Topic Models as a Measure of Schemas

While the schema model in Study 2 captures several features of human schemas, the model does fail to capture two key features traditionally attributed to human schemas: that human schemas are structured hierarchically or temporally, and that there are several subtypes of human schemas. Because of these limitations, it may be more appropriate to label the representations learned by the topic model as a type of conceptual structure rather than a schema.

There is a long tradition of research arguing that human schemas are structured hierarchically and sometimes include information about temporal order. In the case of event scripts, this structure is thought to be extended over time. For example, preparing coffee involves adding coffee, then adding sugar, and then adding milk (Cooper & Shallice, 2000; 2006). Schemas are also thought to be hierarchically structured (Rosch et al, 1976). For example, adding coffee can be decomposed into actions such as holding, discarding, opening, and closing (Cooper & Shallice, 2006). While human schemas are thought to be structured hierarchically and sometimes temporally, a limitation of topic models is that they do not represent either hierarchical or temporal structure. This absence of structured representation is a limitation of the topic model in Study 2, although topic models could potentially be extended to overcome these limitations. Hierarchical representation could be added using hierarchical topic models, which in principle could be extended to represent more than one layer of hierarchical structure (Griffiths, Jordan, & Tenenbaum, 2004). Additionally, event scripts extended over time could perhaps be captured by correlated topic models (Blei & Lafferty, 2007). Correlated topic models are typically used to model correlations among topics when the corpus is extended over time, but these models could potentially be extended to capture topics that evolve over time in a single document.

A further limitation of topic models is that humans are thought to have several different types of conceptual representations, but the topic model does not distinguish among these types. A key distinction is between event scripts, which represent events, and schemas, which are typically used to represent objects. Additional types of human conceptual representations include those constructed around a particular goal, such as ad hoc categories (Barsalou, 1983), and those organized by the correlational structure of the world, natural categories (Rosch, 1973; Lassaline, Wisniewsky, & Medin, 1992). Topic models do not distinguish between these types of category representation, but instead could assign a single probability distribution to each category.

For these reasons, it may be more appropriate to label the representations learned by the topic model as conceptual structures rather than schemas. The term conceptual structure is meant to denote that the topics capture semantically coherent packets of information that are consistent over time and sufficiently generic to be consistent across corpora. Nevertheless, conceptual structures may fail to capture some of the features of schemas such as hierarchical or temporal structure and do not distinguishing among different subtypes of representations.

5.3.2 Limitations of Measuring Episodic Processing Using Language

The measurement of episodic processing was based on people’s language, which is an indirect measure of episodic processing. This measure could be limited because only three linguistic correlates of episodic processing were studied, and because it was not possible to ensure that the events extracted were bounded versus unbounded in time. While these are limitations of the episodic processing measure, I elaborate several reasons below that the results are unlikely to change if episodic processing were measured in a different way.

Episodic future thinking is typically defined as a type of simulation or pre-experience of the future (Atance & O'Neill, 2001; Atance & O'Neill, 2005; Schacter, Benoit, & Szpunar, 2017). However, more specifically episodic future thinking involves not just pre-experiencing, but also an event that is bounded in space and time. It was possible to extract two of these features from language – pre-experiencing and spatial relational language – but the current study did not extract linguistic measures that could determine whether the events referred to were bounded versus extended over time. For this reason, it is possible that the measure of episodic processing could have captured events which are highly perceptual and spatial, but are extended over time.

While the measurement of episodic language is imperfect, there are three reasons that the basic result is unlikely to be explained by the way episodic processing was measured. First, similar results were obtained for all three correlates of episodic processing. If episodic processing is sensitive to the particular method of measurement, then one would have expected less consistency among the measures of episodic processing. Second, even if temporally unbounded events were captured, there is no *a priori* reason to expect that people use episodic processing similarly for bounded past and future events, but differently for unbounded past and future events. Third, even if a difference were observed between temporally bounded and unbounded episodic processing, such a difference is not predicted by existing accounts of episodic future thinking, and thus would still require explanation. For these reasons, it is likely that the basic finding – that episodic processing is not uniquely associated with mental time travel – does not depend on the particular way that episodic processing was measured.

While my findings are theoretically novel, there are other findings using different measures of episodic processing that are consistent with my main result. First, there are several reports that patients with deficits to the episodic memory system due to hippocampal damage can think about the future in many ways, including normal delay discounting and normal future temporal orientation (Kwan et al, 2012; 2013; 2015). These results also suggest that episodic processing may not be uniquely associated with mental time travel, since loss of episodic processing can affect past thinking without affecting thoughts about the future. Second, others have found that thoughts about the past are rated as more perceptually vivid than thoughts about the future, again suggesting a greater involvement of episodic processing in thoughts about the past than thoughts about the future (Rubin, 2014). Together, these findings suggest that our results for episodic processing may not be specific to the way episodic processing was measured.

5.3.3. Alternative Interpretation of the Role of Schemas in Past and Future Thinking

The main result of these studies was that the use of schemas is uniquely associated with mental time travel. However, differences were also observed between past and future thinking. In particular, future thoughts relied relatively more on schemas than did past thoughts. This difference in amount could suggest a different conclusion: that past and future thinking instead rely on different processes. I argue that this such a conclusion is too strong. The results instead suggest that past and future thinking rely on a shared process – the use of schemas – but to a differing degree,

One reason that schemas could be used more in future thinking than in past thinking is that schemas may be needed relatively more during the construction phase of future thinking compared to the retrieval phase of past thinking. Two findings in existing literature are consistent with this explanation. First, it is sometimes observed that the hippocampus is more active for future thoughts than past thoughts, a finding that is typically explained by the increased construction demands needed for future thinking compared to past thinking (Addis, Pan, Vu, Laiser, & Schacter, 2009a; Addis, Wong, & Schacter, 2007; Kirwan, Ashby, & Nash, 2014). Second, for distant future thoughts in particular, a schema may be needed due to the highly unbounded nature of distant future events. For example, D'argembeau & Demblon (2012) found that personal goals were especially used to structure events in the distant, compared to near, future.

5.3.4 Limitations of the Temporal Orientation Classifier

The key variable of interest – the temporal orientation of sentences – was extracted using an automated classification model. While the model was extensively validated and approached human-level performance, the classifier does have two main limitations. While these limitations are unlikely to affect the main results of the studies, addressing these limitations is a tractable problem for future work.

The first limitation is that there were no rules for classifying present sentences. Instead, sentences that were neither classified as past nor future were identified as present. This assumption is limited because sentences that refer to neither the past nor the future can also be atemporal, such as the sentence *birds have wings*. The second limitation is that some sentences can have more than one temporal orientation, but the classifier assigns only a single temporal orientation to each sentence. For example, the sentence *I was thinking about what I was going to do tomorrow* refers to both the past and the present. Future work could consider extending the temporal orientation classifier to handle these cases with multiple temporal orientations. Such an extension could potentially be accomplished by assigning temporal orientations to units below the level of a sentence, such as to individual clauses in the sentence.

5.4 Conclusion

In conclusion, the results suggest that past and future thinking may share common cognitive processes. The results also suggest that the nature of these common processes is different than previously believed. The key shared process may be the use of schemas, not episodic processing. More broadly, the results suggest that people's large-scale talk about the past and future can be revealing of their underlying psychology. Future work should extend these results by using a more cognitively realistic model of schemas and by improving the temporal orientation classifier.

References

Abraham, A., Schubotz, R., & von Cramon, Y. (2008). Thinking about the future versus the past in personal and non-personal contexts. *Brain research, 1233*, 106-119.

Addis, D., Cheng, T., Roberts, R., & Schacter, D. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus, 21*(10), 1045-1052.

Addis, D., Pan, L., Vu, M.-A., Laiser, N., & Schacter, D. (2009a). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia, 47*(11), 2222-2238.

Addis, D., & Schacter, D. (2008). Constructive episodic simulation: Temporal distance and detail of past and future events modulate hippocampal engagement. *Hippocampus, 18*(2), 227-237.

Addis, D., Wong, A., & Schacter, D. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia, 45*(7), 1363-1377.

Addis, D., Wong, A., & Schacter, D. (2008). Age-related changes in the episodic simulation of future events. *Psychological Science, 19*(1), 33-41.

Addis, D. R., Sacchetti, D. C., Ally, B. A., Budson, A. E., & Schacter, D. (2009b). Episodic simulation of future events is impaired in mild Alzheimer's disease. *Neuropsychologia, 47*(12), 2660-2671.

Addis, D. R., & Schacter, D. (2012). The hippocampus and imagining the future: where do we stand? *Frontiers in human neuroscience, 5*, 173.

Andelman, F., Hoofien, D., Goldberg, I., Aizenstein, O., & Neufeld, M. Y. (2010). Bilateral hippocampal lesion and a selective impairment of the ability for mental time travel. *Neurocase, 16*(5), 426-435.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in cognitive sciences, 5*(12), 533-539.

Atance, C. M., & O’Neill, D. K. (2005). The emergence of episodic future thinking in humans. *Learning and motivation, 36*(2), 126-144.

Barsalou, L. (1983). Ad hoc categories. *Memory & cognition*, *11*(3), 211-227.

Barsics, C., Van der Linden, M., & D'Argembeau, A. (2016). Frequency, characteristics, and perceived functions of emotional future thinking in daily life. *The quarterly journal of experimental psychology, 69*(2), 217-233.

Bartlett, F. (1920). Some experiments on the reproduction of folk-stories. *Folklore, 31*(1), 30-47.

Bauer, P. (1993). Memory for gender-consistent and gender-inconsistent event sequences by twenty-five-month-old children. *Child development, 64*(1), 285-297.

Baumeister, R., Hofmann, W., Summerville, A., Reiss, P., & Vohs, K. (2018). Everyday thoughts in time: experience sampling studies of mental time travel *PsyArXIV Preprint*. doi:0.31234/osf.io/3cwre

Baumeister, R., Vohs, K., & Oettingen, G. (2016). Pragmatic prospection: How and why people think about the future. *Review of General Psychology, 20*(1), 3-16.

Baumgartner, J. (2019). Pushshift.io. Retrieved from https://pushshift.io/

Beaty, R., Seli, P., & Schacter, D. (2019). Thinking about the past and future in daily life: an experience sampling study of individual differences in mental time travel. *Psychological research, 83*(4), 805-816.

Benoit, R. G., & Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia, 75*, 450-457.

Berntsen, D., & Jacobsen, A. (2008). Involuntary (spontaneous) mental time travel into the past and future. *Journal of consciousness and cognition, 17*(4), 1093-1104.

Berntsen, D., & Rubin, D. (2004). Cultural life scripts structure recall from autobiographical memory. *Memory and cognition, 32*(3), 427-442.

Bertossi, E., Tesini, C., Cappelli, A., & Ciaramelli, E. (2016). Ventromedial prefrontal damage causes a pervasive impairment of episodic memory and future thinking. *Neuropsychologia, 90*, 12-24.

Blei, D., Ng, A., & Jordan, M. (2002). *Latent dirichlet allocation.* Paper presented at the Advances in neural information processing systems.

Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of machine learning research, 3*(Jan), 993-1022.

Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The annals of applied statistics, 1*(1), 17-35.

Bower, G., Black, J., & Turner, T. (1979). Scripts in memory for text. *Cognitive Psychology, 11*(2), 177-220.

Brewer, W., & Treyens, J. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*(2), 207-230.

Brown, A. D., Addis, D. R., Romano, T. A., Marmar, C. R., Bryant, R. A., Hirst, W., & Schacter, D. L. (2014). Episodic and semantic components of autobiographical memories and imagined future events in post-traumatic stress disorder. *Memory, 22*(6), 595-604.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods, 46*(3), 904-911.

Buckner, R. L. J. A. r. o. p. (2010). The role of the hippocampus in prediction and imagination. *61*, 27-48.

Cao, Z., Li, S., Liu, Y., Li, W., & Ji, H. (2015). *A novel neural topic model and its supervised extension.* Paper presented at the Twenty-Ninth AAAI Conference on Artificial Intelligence.

Chang, A. X., & Manning, C. D. (2012). *Sutime: A library for recognizing and normalizing time expressions.* Paper presented at the Lrec.

Chen, D., & Manning, C. (2014). *A fast and accurate dependency parser using neural networks.* Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).

Cooper, R. & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive neuropsychology, 17*(4), 297-338.

Cooper, R. & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological review*, *113*(4), 887-916.

Cole, S. N., Morrison, C. M., & Conway, M. A. (2013). Episodic future thinking: Linking neuropsychological performance with episodic detail in young and old adults. *The quarterly journal of experimental psychology, 66*(9), 1687-1706.

Cooper, R. P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior.

Craver, C. F., Kwan, D., Steindam, C., & Rosenbaum, R. S. (2014). Individuals with episodic amnesia are not stuck in time. *Neuropsychologia, 57*, 191-195.

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the psychonomic society, 4*(5), 517-518.

D'Argembeau, A., & Mathy, A. (2011). Tracking the construction of episodic future thoughts. *Journal of Experimental Psychology: General, 140*(2), 258.

D’Argembeau, A., & Demblon, J. (2012). On the representational systems underlying prospection: Evidence from the event-cueing paradigm. *Cognition, 125*(2), 160-167.

D’Argembeau, A., & Van der Linden, M. (2004). Phenomenal characteristics associated with projecting oneself back into the past and forward into the future: Influence of valence and temporal distance. *Consciousness and cognition, 13*(4), 844-858.

D’Argembeau, A., & Van der Linden, M. (2006). Individual differences in the phenomenology of mental time travel: The effect of vivid visual imagery and emotion regulation strategies. *Consciousness and cognition, 15*(2), 342-350.

De Brigard, F., Addis, D. R., Ford, J. H., Schacter, D. L., & Giovanello, K. (2013). Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia, 51*(12), 2401-2414.

De Brigard, F., & Giovanello, K. (2012). Influence of outcome valence in the subjective experience of episodic past, future, and counterfactual thinking. *Journal of consciousness and cognition, 21*(3), 1085-1096.

De Luca, F., Benuzzi, F., Bertossi, E., Braghittoni, D., di Pellegrino, G., & Ciaramelli, E. (2018). Episodic future thinking and future-based decision-making in a case of retrograde amnesia. *Neuropsychologia, 110*, 92-103.

De Melo, G., & Weikum, G. (2010). *MENTA: Inducing multilingual taxonomies from Wikipedia.* Paper presented at the Proceedings of the 19th ACM international conference on Information and knowledge management.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the american society for information science, 41*(6), 391-407.

Demblon, J., & D'Argembeau, A. (2014). The organization of prospective thinking: Evidence of event clusters in freely generated future thoughts. *Consciousness and cognition, 24*, 75-83.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv Preprint*.

Eichstaedt, J., Smith, R., Merchant, R., Ungar, L., Crutchley, P., Preoţiuc-Pietro, D., . . . Schwartz, H. A. s. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences, 115*(44), 11203-11208.

Firth, J. R. (1950). Personality and language in society. *The sociological review, 42*(1), 37-52.

Fivush, R. (2002). Scripts, schemas, and memory of trauma. *Representation, Memory, and Development: Essays in Honor of Jean Mandler*, 53-74.

Gilbert, D. T., & Wilson, T. D. J. S. (2007). Prospection: Experiencing the future. *317*(5843), 1351-1354.

Goldstone, R. L., & Lupyan, G. (2016). Discovering psychological principles by mining naturally occurring data sets. *Topics in cognitive science, 8*(3), 548-568.

Griffiths, T., Steyvers, M., & Tenenbaum, J. (2007). Topics in semantic representation. *Psychological review, 114*(2), 211.

Griffiths, T. L., Jordan, M. I., Tenenbaum, J. B., & Blei, D. M. (2004). *Hierarchical topic models and the nested Chinese restaurant process.* Paper presented at the Advances in neural information processing systems.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*(suppl 1), 5228-5235.

Guntuku, S., Yaden, D., Kern, M., Ungar, L., & Eichstaedt, J. (2017). Detecting depression and mental illness on social media: an integrative review. *Current opinion in behavioral sciences, 18*, 43-49.

Hach, S., Tippett, L., & Addis, D. (2014). Neural changes associated with the generation of specific past and future events in depression. *Neuropsychologia, 65*, 41-55.

Hassabis, D., Kumaran, D., & Maguire, E. (2007a). Using imagination to understand the neural basis of episodic memory. *Journal of neuroscience, 27*(52), 14365-14374.

Hassabis, D., Kumaran, D., Vann, S., & Maguire, E. (2007b). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences, 104*(5), 1726-1731.

Hintzman, D. L. (1986). " Schema abstraction" in a multiple-trace memory model. *Psychological review, 93*(4), 411.

Hoffman, M., Bach, F. R., & Blei, D. M. (2010). *Online learning for latent dirichlet allocation.* Paper presented at the advances in neural information processing systems.

Hurley, N. C., Maguire, E. A., & Vargha-Khadem, F. (2011). Patient HC with developmental amnesia can construct future scenarios. *Neuropsychologia, 49*(13), 3620-3628.

Irish, M., Addis, D. R., Hodges, J. R., & Piguet, O. (2012). Considering the role of semantic memory in episodic future thinking: evidence from semantic dementia. *Brain, 135*(7), 2178-2191.

Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in behavioral neuroscience, 7*, 27.

Irish, M., & Piolino, P. (2016). Impaired capacity for prospection in the dementias – Theoretical and clinical implications. *British Journal of Clinical Psychology, 55*(1), 49-68.

Jing, H. G., Madore, K. P., & Schacter, D. L. (2017). Preparing for what might happen: An episodic specificity induction impacts the generation of alternative future events. *Cognition, 169*, 118-128.

Johns, B. T., & Dye, M. (2019). Gender bias at scale: Evidence from the usage of personal names. *Behavior research methods*, 1-18.

Jones, M., Kintsch, W., & Mewhort, D. (2006). High-dimensional semantic space accounts of priming. *Journal of memory and language, 55*(4), 534-552.

Jones, M., & Mewhort, D. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological review, 114*(1), 1.

Jones, M. N. (2016). *Big data in cognitive science*: Psychology Press.

Kataria, S. S., Kumar, K. S., Rastogi, R. R., Sen, P., & Sengamedu, S. H. (2011). *Entity disambiguation with hierarchical topic models.* Paper presented at the Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining.

Kingma, D. P., & Ba, J. J. a. p. a. (2014). Adam: A method for stochastic optimization.

Kintsch, W., & Greene, E. (1978). The role of culture-specific schemata in the comprehension and recall of stories. *Discourse processes, 1*(1), 1-13.

Kirwan, B., Ashby, S., & Nash, M. (2014). Remembering and imagining differentially engage the hippocampus: a multivariate fMRI investigation. *Cognitive neuroscience, 5*(3-4), 177-185.

Klein, S., Loftus, J., & Kihlstrom, J. (2002). Memory and temporal experience: The effects of episodic memory loss on an amnesic patient's ability to remember the past and imagine the future. *Social Cognition, 20*(5), 353-379.

Kopelman, M., Wilson, B., & Baddeley, A. (1989). The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. *Journal of clinical experimental neuropsychology, 11*(5), 724-744.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks.* Paper presented at the Advances in neural information processing systems.

Kwan, D., Craver, C. F., Green, L., Myerson, J., Boyer, P., & Rosenbaum, R. S. (2012). Future decision-making without episodic mental time travel. *Hippocampus, 22*(6), 1215-1219.

Kwan, D., Craver, C. F., Green, L., Myerson, J., Gao, F., Black, S. E., & Rosenbaum, R. S. (2015). Cueing the personal future to reduce discounting in intertemporal choice: Is episodic prospection necessary? *Hippocampus, 25*(4), 432-443.

Kwan, D., Craver, C. F., Green, L., Myerson, J., & Rosenbaum, R. S. (2013). Dissociations in future thinking following hippocampal damage: evidence from discounting and time perspective in episodic amnesia. *Journal of Experimental Psychology: General, 142*(4), 1355.

Lassaline, M., Wisniewski, E., & Medin, D. (1992). 9 basic levels in artificial and natural categories: are all basic levels created equal? *Advances in psychology*, *93*, 327-378.

Lau, J. H., Grieser, K., Newman, D., & Baldwin, T. (2011). *Automatic labelling of topic models.* Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1.

Levine, B., Svoboda, E., Hay, J., Wincour, G., & Moscovitch, M. (2002). Aging and autobiographical memory: dissociating episodic from semantic retrieval. *Psychology and aging, 17*(4), 677-689.

Levy, R., & Andrew, G. (2006). *Tregex and Tsurgeon: tools for querying and manipulating tree data structures.* Paper presented at the LREC.

Lilienfeld, S. O. (2010). Can psychology become a science? *Personality and individual differences, 49*(4), 281-288.

Liu, H., & Singh, P. (2004). ConceptNet—a practical commonsense reasoning tool-kit. *BT Technology Journal, 22*(4), 211-226.

Lourenco, S. F., & Frick, A. (2013). Remembering where: The origins and early development of spatial memory. *The Wiley Handbook on the Development of Children's Memory, 1*, 361-393.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior research methods, 28*(2), 203-208.

Lupyan, G., & Goldstone, R. L. (2019). Introduction to special issue. Beyond the lab: Using big data to discover principles of cognition. *Behavior research methods*.

Madore, K. P., Gaesser, B., & Schacter, D. L. (2014). Constructive episodic simulation: Dissociable effects of a specificity induction on remembering, imagining, and describing in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 609.

Madore, K. P., Szpunar, K. K., Addis, D. R., & Schacter, D. L. (2016). Episodic specificity induction impacts activity in a core brain network during construction of imagined future experiences. *Proceedings of the National Academy of Sciences, 113*(38), 10696-10701.

Magatti, D., Calegari, S., Ciucci, D., & Stella, F. (2009). *Automatic labeling of topics.* Paper presented at the 2009 Ninth International Conference on Intelligent Systems Design and Applications.

Maguire, E. A., Vargha-Khadem, F., & Hassabis, D. (2010). Imagining fictitious and future experiences: Evidence from developmental amnesia. *Neuropsychologia, 48*(11), 3187-3192.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., & McClosky, D. (2014). *The Stanford CoreNLP natural language processing toolkit.* Paper presented at the Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations.

McRae, K., & Jones, M. (2013). Semantic Memory. *The Oxford handbook of cognitive psychology, 206*.

Mei, Q., Shen, X., & Zhai, C. (2007). *Automatic labeling of multinomial topic models.* Paper presented at the Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). *Distributed representations of words and phrases and their compositionality.* Paper presented at the Advances in neural information processing systems.

Minsky, M. (1974). A framework for representing knowledge.

Nastase, V., Strube, M., Börschinger, B., Zirn, C., & Elghafari, A. (2010). *WikiNet: A Very Large Scale Multi-Lingual Concept Network.* Paper presented at the LREC.

Navigli, R., & Ponzetto, S. P. (2012). BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial intelligence, 193*, 217-250.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., . . . Yamadori, A. (2003). Thinking of the future and past: The roles of the frontal pole and the medial temporal lobes. *Neuroimage, 19*(4), 1369-1380.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., . . . Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of personality and Social Psychology, 108*(6), 934.

Paxton, A., & Griffiths, T. (2017). Finding the traces of behavioral and cognitive processes in big data and naturally occurring datasets. *Behavior research methods, 49*(5), 1630-1638.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Retrieved from

Peters, J., Wiehler, A., & Bromberg, U. (2017). Quantitative text feature analysis of autobiographical interview data: Prediction of episodic details, semantic details and temporal discounting. *Nature scientific reports, 7*(1), 14989.

Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., & Lee, K. (2018). Deep contextualized word representations. In.

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(4), 587.

Porteous, I., Newman, D., Ihler, A., Asuncion, A., Smyth, P., & Welling, M. (2008). *Fast collapsed gibbs sampling for latent dirichlet allocation.* Paper presented at the Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining.

Race, E., Keane, M. M., & Verfaellie, M. (2011). Medial temporal lobe damage causes deficits in episodic memory and episodic future thinking not attributable to deficits in narrative construction. *Journal of neuroscience, 31*(28), 10262-10269.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.

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

Raschka, S. (2015). *Python machine learning*: Packt Publishing Ltd.

Rasmussen, A. S., & Berntsen, D. (2013). The reality of the past versus the ideality of the future: Emotional valence and functional differences between past and future mental time travel. *Journal of memory and cognition, 41*(2), 187-200.

Rohde, D. L., Gonnerman, L. M., & Plaut, D. C. (2006). An improved model of semantic similarity based on lexical co-occurrence. *Communications of the ACM, 8*(627-633), 116.

Rosch, E. (1973). Natural categories. *Cognitive psychology*, *4*(3), 328-350.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, *8*(3), 382-439.

Rubin, D. C. (1981). *Norms for 34 properties of 125 words*: American Psycholog. Ass., Journal Suppl. Abstract Service.

Rubin, D. C. (2014). Schema-driven construction of future autobiographical traumatic events: the future is much more troubling than the past. *Journal of Experimental Psychology: General, 143*(2), 612.

Rule, A., Cointet, J.-P., & Bearman, P. (2015). Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014. *Proceedings of the National Academy of Sciences, 112*(35), 10837-10844.

Russell, J., Alexis, D., & Clayton, N. (2010). Episodic future thinking in 3-to 5-year-old children: The ability to think of what will be needed from a different point of view. *Cognition, 114*(1), 56-71.

Schacter, D., Addis, D. R., & Buckner, R. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences, 1124*(1), 39-60.

Schacter, D., Benoit, R., & Szpunar, K. (2017). Episodic future thinking: Mechanisms and functions. *Current opinion in behavioral sciences, 17*, 41-50.

Schacter, D., Gaesser, B., & Addis, D. (2013). Remembering the past and imagining the future in the elderly. *J Gerontology, 59*(2), 143-151.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical transactions of the royal society B, 362*(1481), 773-786.

Schank, R. C., & Abelson, R. P. (1975). *Scripts, plans, and knowledge.* Paper presented at the IJCAI.

Schler, J., Koppel, M., Argamon, S., & Pennebaker, J. W. (2006). *Effects of age and gender on blogging.* Paper presented at the AAAI spring symposium: Computational approaches to analyzing weblogs.

Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences*, 201821028.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . . Seligman, M. E. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one, 8*(9), e73791.

Schwartz, H. A., Park, G., Sap, M., Weingarten, E., Eichstaedt, J., Kern, M., . . . Seligman, M. (2015). *Extracting human temporal orientation from Facebook language.* Paper presented at the Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

Shepard, J., Nie, A., Copley, B., & Wolff, P. (2017). *Discovering kinds of future-oriented thought using automated machine-learning techniques.* Paper presented at the CogSci.

Speer, R., Chin, J., & Havasi, C. (2017). *Conceptnet 5.5: An open multilingual graph of general knowledge.* Paper presented at the Thirty-First AAAI Conference on Artificial Intelligence.

Spreng, R. N., & Levine, B. (2006). The temporal distribution of past and future autobiographical events across the lifespan. *Memory & cognition, 34*(8), 1644-1651.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *Journal of cognitive neuroscience, 21*(3), 489-510.

Squire, L. R., van der Horst, A. S., McDuff, S. G., Frascino, J. C., Hopkins, R. O., & Mauldin, K. N. (2010). Role of the hippocampus in remembering the past and imagining the future. *Proceedings of the National Academy of Sciences, 107*(44), 19044-19048.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. J. T. j. o. m. l. r. (2014). Dropout: a simple way to prevent neural networks from overfitting. *15*(1), 1929-1958.

Szpunar, K., Spreng, N., & Schacter, D. (2016). Toward a taxonomy of future thinking. *Seeing the future: theoretical perspectives on future mental time travel*, 21-35.

Szpunar, K., Spreng, R., & Schacter, D. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences, 111*(52), 18414-18421.

Szpunar, K., Watson, J., & McDermott, K. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences, 104*(2), 642-647.

Thorstad, R., & Wolff, P. (2018). A big data analysis of the relationship between future thinking and decision-making. *Proceedings of the National Academy of Sciences, 115*(8), E1740-E1748.

Thorstad, R., & Wolff, P. (2019). Predicting future mental illness from social media: A big-data approach. *Behavior research methods*, 1-15.

Tulving, E. (1972). Episodic and semantic memory. *Organization of memory*, 381-403.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne, 26*(1), 1.

van Kesteren, M., Ruiter, D., Fernández, G., & Henson, R. (2012). How schema and novelty augment memory formation. *Trends in neurosciences, 35*(4), 211-219.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science, 277*(5324), 376-380.

Viard, A., Chételat, G., Lebreton, K., Desgranges, B., Landeau, B., de La Sayette, V., . . . Piolino, P. (2011). Mental time travel into the past and the future in healthy aged adults: an fMRI study. *Brain and cognition, 75*(1), 1-9.

Vinson, D., Dale, R., & Jones, M. (2019). Decision contamination in the wild: Sequential dependencies in online review ratings. *Behavior research methods*, 1-8.

Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. R. (2018). Glue: A multi-task benchmark and analysis platform for natural language understanding. *ArXiv Preprint*.

Wang, W., Gan, Z., Wang, W., Shen, D., Huang, J., Ping, W., . . . Carin, L. (2017). Topic compositional neural language model. *ArXiv Preprint*.

Warde-Farley, D., Goodfellow, I. J., Courville, A., & Bengio, Y. (2013). An empirical analysis of dropout in piecewise linear networks. *ArXiv Preprint*.

Appendix 1: Formal Description of Latent Dirichlet Allocation

In Chapter 3, I describe a Latent Dirichlet Allocation model used to infer common schemas. The model is described informally in the main text; here, I provide a more formal description of the model.

Latent Dirichlet Allocation (Blei, Ng, & Jordan, 2003) is a Bayesian model, and thus can be represented using a Bayesian plate notation. In Fig. S1 below, circles represent variables in the model, arrows represent dependencies among those variables, and rectangles represent iteration. Variables are separately colored by whether they are observed (dark circles) or unobserved (light circles).

The animating idea for LDA is a generative model of language, that when a speaker writes a document, the speaker chooses a multinomial distribution over topics to write about, θ_m and a number of words to write N. Next, for each word in the document N_m, the speaker draws a single topic z_{mn} probablistically from the distribution θ_m. Since a topic itself defines a probability distribution over the vocabulary, the speaker then draws a word x_{mn} probablistically from the topic and writes this word in the document, and continues this process until N_m words have been written. Because the distributions of words to topics and topics across documents follow multinomial distributions, they are naturally represented by Dirichlet priors α and β. Finally, note that Fig. S1 illustrates the generative model for using topics to write a document. The model is parameterized (learned) by reversing this generative model to use the distribution of words in documents x_{mn} to infer θ_m and ϕ_k.

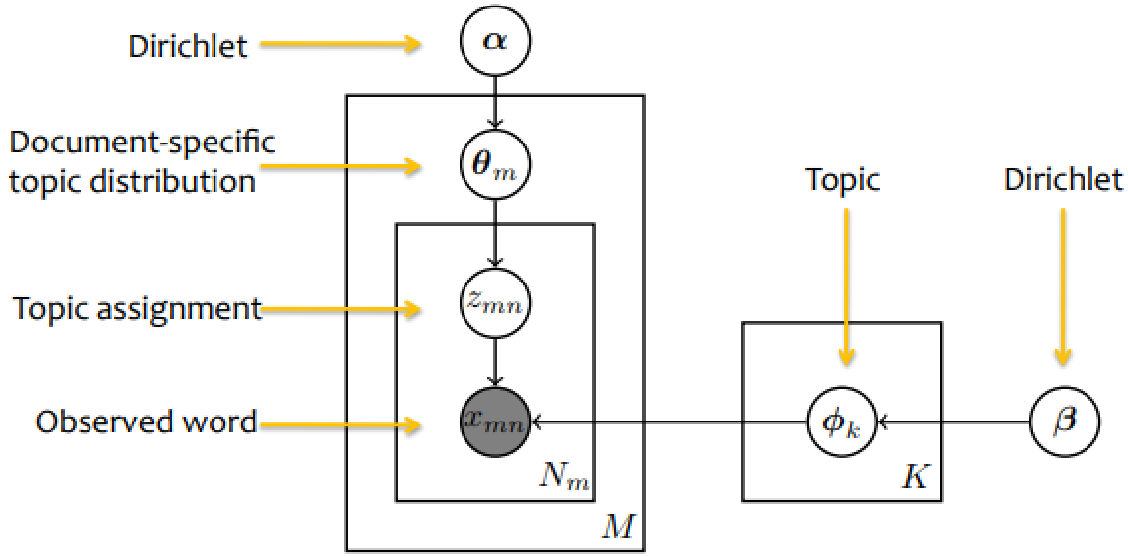

***Fig. S1. Bayesian plate notation for Latent Dirichlet Allocation*.** Figure used with permission from Matt Gormley.

Appendix 2: List of Topics Learned by Topic Model

In Chapter 3, I report a topic model trained on people's everyday language on social media. Examples of some of the topics learned by the model are listed in the main text. Below, I report the full 500 topics learned by the model, where each topic is represented by the 5 most probable words in the topic.

Topic	Term #1	Term #2	Term #3	Term #4	Term #5
1	brothers	sisters	remotely	surround	insist
2	track	art	album	band	metal
3	im	dont	cant	ive	didnt
4	gold	min	assist	victory	defender
5	internet	access	network	connect	connection
6	party	vote	votes	parties	voting
7	research	responses	survey	factors	insanely
8	minutes	allowed	information	golden	heads
9	wall	floor	walls	drain	ceiling
10	favorite	favorites	stamina	unexpected	meh
11	ios	stable	succeed	jailbreak	rift
12	added	changed	fixed	removed	remove
13	visa	beaten	musical	hats	blanket
14	vr	desk	hook	hooked	virtual
15	fit	size	wear	smaller	small
16	superior	placing	ireland	irish	ontario
17	lord	wise	lag	chaos	jon
18	deep	pool	deeper	dive	businesses
19	chase	pump	bullet	bite	candy
20	fun	talk	chat	wanna	pics
21	months	home	month	house	moving
22	sa	shell	null	si	ti
23	assets	downtown	ppl	ashamed	commute
24	speed	fast	seconds	slow	increase
25	shooting	tournaments	aiming	drone	viewing
26	boyfriend	gf	bf	uncomfortable	annoyed
27	hero	bill	deleted	trans	submitted
28	piece	pieces	piano	uploaded	centered
29	status	critical	resolve	trials	d3
30	2nd	3rd	1st	row	4th
31	hit	hold	button	miss	jump

32	volume	quiet	loud	joy	bodies
33	perform	talks	supported	officially	performing
34	called	killed	gods	nine	layers
35	pay	paying	attention	afford	debt
36	buy	price	buying	cheap	expensive
37	match	matches	bet	crowd	rumble
38	music	song	round	listen	songs
39	official	tribe	destroy	ark	dragons
40	led	sync	chip	responding	electronics
41	league	rocket	legends	wire	champions
42	pink	orange	purple	sky	lime
43	physical	target	ha	weak	wind
44	debating	walmart	hesitant	healthcare	earning
45	products	sample	mask	belt	beauty
46	dog	dogs	owner	loves	behavior
47	reddit	posting	host	tag	username
48	keyboard	owned	typing	mechanical	razer
49	centre	progressive	vancouver	slightest	camo
50	tie	wing	tied	suitable	thru
51	open	opened	opening	mouse	window
52	format	sheet	contains	excel	separated
53	believe	god	truly	spirit	faith
54	monitor	ram	fps	settings	gpu
55	cast	spell	spells	casting	lp
56	de	la	que	el	en
57	interest	rate	student	loan	rates
58	sets	strength	bench	lift	stretch
59	pc	gaming	xbox	suit	gamer
60	went	wasn	told	couldn	happened
61	force	master	jedi	luke	climb
62	top	list	bottom	tier	highest
63	pokemon	male	female	gender	nature
64	text	sent	contact	received	id
65	map	quit	porn	maps	nofap
66	option	options	select	chose	instructions
67	high	low	cost	quality	value
68	burn	burning	masters	humble	burned
69	memories	handed	verify	smith	childhood
70	grow	growing	flat	leaves	plant
71	freedom	iii	variable	dudes	mk
72	face	skin	dry	nose	acid
73	logo	missions	lang	mvp	flare
74	above	average	overall	score	compare

75	level	higher	lower	skill	levels
76	today	yesterday	tomorrow	tonight	schedule
77	drink	drinking	smoke	drug	smoking
78	zone	buddy	egg	eggs	weekends
79	star	training	wars	rising	sir
80	final	fantasy	roles	dom	bi
81	base	ebay	cam	bases	contacts
82	build	memory	case	core	atx
83	caps	checks	fishing	mx	cherry
84	ai	suggestion	disabled	enabled	applications
85	experience	experienced	experiences	effects	flair
86	race	hill	races	silent	racing
87	sub	red	bar	green	icon
88	add	items	offer	offers	accept
89	features	feature	frame	lock	wood
90	king	killer	aka	kings	prince
91	apple	earn	strike	samsung	fallen
92	paypal	pin	verified	directory	paths
93	actions	against	person	rights	identity
94	six	cases	ear	rainbow	ears
95	logic	curse	module	pointless	warranty
96	pull	pulled	grind	pulling	usd
97	law	french	laws	mexico	speech
98	practice	guitar	slot	tone	electric
99	school	college	university	high	student
100	mistake	dumb	mistakes	apologize	formatting
101	silver	boost	placed	bronze	plat
102	second	three	four	third	action
103	fate	excuses	courage	dominated	denver
104	primary	secondary	wipe	depends	med
105	birthday	duty	alex	celebrate	ops
106	stock	trigger	sight	barrel	riding
107	pick	picked	draft	picking	picks
108	clone	booster	automation	clones	consecutive
109	supplies	vault	thumb	thx	swift
110	center	island	roof	kitchen	designs
111	er	shock	til	og	det
112	society	organized	donate	revolution	ruby
113	phone	update	device	updated	android
114	dot	5x	bump	deadly	toe
115	attempt	false	sam	cave	blocking
116	market	profit	investment	rise	inc
117	between	difference	major	scale	surprised

118	google	facebook	states	searching	funny
119	downloaded	mirror	matchmakin g	sorted	clearance
120	show	series	shows	episode	friday
121	battery	charging	dual	charger	legacy
122	sales	route	continues	reaching	copies
123	equipment	equipped	traits	seattle	tolerance
124	switch	ps4	controller	launch	nintendo
125	environment	pros	logged	login	cons
126	affect	sources	ghost	guarantee	reads
127	village	practical	rival	excluding	rivals
128	wedding	german	dress	traditional	accuracy
129	stored	receives	madrid	rehab	beastball
130	shake	sprint	gifs	demons	frontier
131	advanced	beginner	embarrassi ng	entrance	rome
132	impact	fury	substantial	nz	wi
133	order	date	pre	december	ordered
134	label	inspiration	tutorials	yoga	scripts
135	side	left	hand	body	hands
136	fn	knife	factory	field	fade
137	grand	ideal	bs	crystal	contrast
138	report	charge	reports	technology	reported
139	death	dead	himself	died	die
140	cap	hat	bow	crown	guardian
141	january	bear	rounds	tap	vehicle
142	travelling	jazz	mobs	sarah	classical
143	sword	japanese	japan	titan	swing
144	address	cd	interface	exp	snapchat
145	oil	remaining	cleaned	pan	thoroughly
146	realistic	esp	alternate	wash	washing
147	plan	planning	weekend	trip	places
148	focus	paper	lens	festival	canon
149	fire	garbage	ancient	warriors	philosophy
150	nc	toilet	blown	paranoid	meditation
151	TRUE	length	spread	delay	angle
152	cpu	bios	clock	cooler	mobo
153	na	em	da	que	de
154	attack	advantage	opponent	attacks	combo
155	hopes	chemical	academy	genius	crossed
156	throw	catch	caught	rid	spawn
157	road	ride	bike	mountain	roads
158	co	op	van	counting	tables

159	pls	peter	hosts	feat	p2
160	appealing	dns	specify	stunning	fm
161	country	countries	usa	international	australia
162	ranked	skins	rude	papers	assassin
163	original	pictures	asking	shipping	brand
164	million	lvl	ultimate	mike	nick
165	usage	pen	nail	ink	marine
166	theory	fi	theories	electronic	sci
167	began	man	blood	once	body
168	paladins	dedication	pioneer	ledge	bahamut
169	trail	bubble	bath	lee	replay
170	goal	goals	kick	corner	chances
171	amount	paid	fair	transfer	orders
172	london	british	valid	highlight	messy
173	psn	abortion	women	birth	control
174	exchange	bitcoin	wallet	secure	coin
175	habits	artificial	photographer	guts	accepts
176	video	app	videos	youtube	channel
177	carry	carrying	carried	danger	dec
178	system	ship	ships	systems	crew
179	run	running	max	runs	ran
180	pro	photos	photo	crash	fish
181	interested	pm	send	message	willing
182	tree	stone	talent	santa	ep
183	indicate	coding	charity	socks	twilight
184	specific	parts	budget	included	models
185	solve	equal	insert	represent	fluid
186	card	amazon	purchase	credit	gift
187	patch	ea	patches	cameras	fifa
188	raid	guild	casual	raids	progression
189	alliance	teeth	spreadsheet	fat	filling
190	needless	certificate	addicted	trait	alice
191	cut	edge	blade	cutting	cuts
192	energy	upgrade	upgraded	upgrading	wheel
193	inches	passes	strip	wears	converted
194	room	door	outside	walk	sitting
195	ah	belong	repeated	usable	cotton
196	past	future	present	tracking	weeks
197	cover	health	insurance	interview	covered
198	1k	5k	2k	4k	flag
199	light	lights	smart	chain	craft

200	town	role	dies	suspicious	punished
201	design	teacher	animation	graphic	releases
202	police	case	legal	court	evidence
203	clear	str	matching	orb	dex
204	kicking	thankful	upwards	healed	explosion
205	read	reading	book	books	library
206	clan	monster	leader	monsters	xp
207	per	confirmation	ma	pot	psychology
208	locked	unlock	unlocked	outfit	accessories
209	shit	fucking	fuck	hate	gonna
210	board	hobby	boards	penis	promoted
211	fresh	mix	salt	cook	bowl
212	ring	cell	cells	ect	drum
213	detailed	immediate	frames	bots	individually
214	comments	comment	leave	below	helped
215	com	product	https	uk	pcpartpicker
216	yet	haven	played	style	gotten
217	evil	demon	neutral	lesser	naruto
218	team	teams	club	sports	division
219	loads	cache	cleared	clearing	reinstalling
220	account	bank	accounts	checking	100k
221	advice	appreciated	suggestions	tips	appreciate
222	wrong	guess	correct	correctly	guessing
223	folder	george	oz	extract	casey
224	version	block	net	default	client
225	story	short	strange	secret	till
226	stick	bars	selection	sticks	spectrum
227	gt	amp	comments	http	np
228	die	der	und	ich	das
229	wiki	en	org	wikipedia	https
230	pop	collection	collect	exclusive	mystery
231	relationship	together	feelings	talk	person
232	learned	iron	survivor	convert	lesson
233	cup	tight	leaning	tags	bra
234	study	subject	passed	studying	studies
235	character	characters	roll	combat	campaign
236	minor	wow	skype	swap	spam
237	package	letter	cross	packages	pause
238	boss	beat	soul	souls	root
239	sale	flash	forum	dungeon	ugly
240	section	security	act	risk	shall
241	chest	legendary	vendor	stealth	marks

242	era	recommendation	remembering	shoots	clueless
243	rotation	displayed	freak	vertical	momentum
244	quarter	membership	calendar	affordable	preparing
245	wish	turning	trash	attempts	clicking
246	task	int	void	tasks	domain
247	girls	sit	mr	th	hd
248	group	community	join	members	active
249	plastic	positions	resume	snap	translation
250	food	eat	eating	healthy	diet
251	click	auto	bot	archive	resubmit
252	intended	ac	robert	commented	apartments
253	stats	hp	dmg	tea	stat
254	feeling	feels	felt	pain	worse
255	daughter	hated	australian	furniture	pcs
256	building	built	castle	buildings	pit
257	hole	boat	screw	holes	screwed
258	form	fill	filled	requests	deposit
259	email	sign	log	reset	password
260	images	pages	spoiler	sorts	alert
261	air	cold	hot	heat	pressure
262	speaker	refund	meme	memes	propose
263	real	yes	estate	ps3	principal
264	ignoring	calculated	summons	worthwhile	composition
265	women	sex	men	woman	sexual
266	blind	paul	playthrough	stadium	conquest
267	famous	subtle	existed	keeper	freely
268	excellent	mil	deliver	vampire	inform
269	pic	recording	plate	uber	alternatives
270	eye	makeup	foundation	liquid	ds
271	hunter	hunt	hunting	gross	hunters
272	sound	tv	sounds	input	audio
273	meeting	church	raised	dh	attend
274	american	south	north	split	america
275	cool	stand	stands	instagram	robots
276	ahead	complex	csgo	tf2	ratings
277	cards	deck	card	decks	meta
278	hide	recovery	chair	hiding	seek
279	kinda	idk	tho	cuz	dunno
280	says	gets	goes	title	starts
281	non	twin	resident	bush	valentine
282	kinks	term	limits	rp	partner
283	self	yourself	confident	confidence	situations

284	file	save	error	files	program
285	realm	gaining	discipline	secretly	charlotte
286	machine	attached	python	expert	machines
287	rough	prompt	loving	daddy	abuse
288	familiar	blow	pray	indie	poop
289	tired	sick	cycle	throat	endless
290	remember	watching	movie	scene	watched
291	class	classes	wild	rogue	knight
292	military	forces	channels	operation	al
293	early	late	march	april	sometime
294	ok	oh	okay	lol	yeah
295	anxiety	depression	afraid	mental	fear
296	repair	aid	badly	damaged	uni
297	resources	farm	eu	resource	admin
298	copy	dan	animated	rotate	collective
299	big	deal	huge	fan	fans
300	screen	camera	view	display	glass
301	blood	period	heart	symptoms	cancer
302	upload	recipes	butter	subscribers	bread
303	fiber	arrangement	satellite	funeral	passage
304	replace	miles	gas	engine	replaced
305	shape	instant	lighting	sizes	shapes
306	alpha	vision	gm	glasses	companion
307	must	rules	follow	rule	apply
308	water	weather	rain	trees	warm
309	head	former	coach	coaching	cleaner
310	move	location	sun	moon	hidden
311	website	site	ad	forums	websites
312	complete	finished	finish	progress	quest
313	canada	ca	imgur	texas	rust
314	chance	cloud	pet	lightning	holy
315	bottle	taste	smooth	flavor	juice
316	thank	hope	amazing	awesome	christmas
317	space	earth	command	planet	stars
318	elf	noble	undead	elves	variations
319	city	army	hundreds	thousands	york
320	set	change	changes	step	setting
321	project	web	development	software	application
322	user	users	fellow	articles	poster
323	front	spare	rear	joint	wheels
324	train	patient	prison	underground	sequence

325	understand	sorry	english	word	explain
326	difficult	perhaps	likely	entirely	less
327	submissions	fighters	ties	athletic	nonsense
328	wife	kids	baby	age	husband
329	gt	lt	div	href	var
330	notes	sweet	smell	vanilla	hack
331	entry	atlanta	entries	georgia	relation
332	stage	kit	classic	stages	kits
333	armor	lti	sc	origin	drake
334	item	description	theme	scratch	fb
335	mark	john	jack	chris	steve
336	trump	news	media	social	president
337	katowice	7c	holo	sticker	ak
338	review	reviews	benefits	benefit	mixed
339	shot	range	gun	shoot	shots
340	noise	mic	speakers	headphones	headset
341	vs	twitch	hype	definition	popping
342	previous	search	source	table	links
343	letters	harry	mt	greek	bud
344	cage	lastly	firstly	secondly	mech
345	clothes	generation	clothing	node	shirts
346	lots	spot	spots	nearby	frequent
347	retail	lip	bn	decay	ml
348	joke	laugh	smile	angel	garden
349	call	straight	calls	calling	gay
350	standard	standards	dope	slap	12k
351	ability	magic	uses	abilities	shadow
352	rings	boom	unusual	cursed	tb
353	wherever	sweden	rt	mia	swedish
354	elite	dangerous	steal	canadian	stolen
355	aim	laid	hammer	perks	perk
356	checked	method	methods	shed	tear
357	seat	ours	oak	dining	fabric
358	money	spend	spent	spending	tickets
359	content	write	writing	written	wrote
360	tank	distance	tanks	filter	capacity
361	weight	gym	gain	lbs	exercise
362	closed	beta	prime	radio	invite
363	shortly	michael	rng	shaped	hollywood
364	server	custom	servers	script	staff
365	mod	mods	adds	skyrim	fallout
366	majority	prepare	arts	valve	martial
367	revealed	overcome	jail	spy	suspected

368	ss	wtf	fox	smash	celebrity
369	service	return	public	private	customer
370	culture	religion	religious	ur	christian
371	car	drive	driving	driver	cars
372	coffee	west	east	coast	shops
373	tax	income	taxes	returns	invest
374	trade	keys	key	sold	offers
375	battle	land	fighting	empire	fight
376	supporting	comics	counts	award	fetish
377	mood	kratom	technique	techniques	strain
378	feed	animal	animals	meat	milk
379	state	tech	ten	multiplayer	florida
380	store	box	local	pack	shop
381	ex	2x	1x	3x	ultra
382	mins	disable	basement	vent	crisis
383	black	white	asian	racist	congratulations
384	season	defense	pass	offense	defensive
385	pure	cooking	approved	hybrid	skilled
386	available	via	craigslist	ifttt	asap
387	chicken	cheese	sauce	picky	cooked
388	world	human	exist	universe	humans
389	toy	toys	scott	william	nightmares
390	updating	pts	promising	downs	egypt
391	lie	lying	lies	treasure	ranger
392	live	online	watch	stream	streaming
393	james	bond	sum	achievement	labor
394	mini	circle	coil	population	coils
395	doom	tube	kato	brutal	hive
396	unit	challenge	units	draw	challenges
397	man	captain	comic	marvel	spider
398	soldiers	soldier	column	phantom	fort
399	decision	banned	ban	choices	permanent
400	dynamic	fields	refer	referring	vendors
401	test	results	positive	result	negative
402	wrath	entirety	yourselves	cognitive	lava
403	confirm	notification	notifications	deny	fulfill
404	details	limited	lady	detail	ssr
405	guard	queen	rose	tokyo	guards
406	random	choose	choice	count	enter
407	est	le	et	de	les
408	fact	wouldn	sense	makes	opinion

409	amp	nbsp	midwife	gearhulk	liliana
410	code	sites	newer	codes	hardicus
411	push	diamond	ups	shower	index
412	during	break	fall	summer	winter
413	ps	views	experiment	france	fc
414	personality	appearance	interaction	attitude	purely
415	page	rep	coins	trading	confirmed
416	heroes	competitive	overwatch	heal	healing
417	check	note	double	profile	menu
418	plus	iphone	console	ipad	consoles
419	di	sh	youtubers	diaper	snaps
420	path	seal	bat	gallery	heavens
421	power	usb	cable	plug	plugged
422	record	records	recorded	footage	dna
423	family	mom	parents	mother	dad
424	effect	cat	mass	cats	owners
425	guide	champion	farming	guides	loop
426	meet	hotel	beach	restaurant	fancy
427	turn	turns	turned	shut	signal
428	argument	repeat	sake	conclusion	eve
429	government	national	state	states	united
430	windows	computer	laptop	installed	install
431	kill	fight	mid	enemy	kills
432	campus	sp	pilot	gray	meter
433	blue	clean	five	st	pattern
434	boots	leather	brown	boot	halo
435	problem	issue	issues	fix	problems
436	common	rare	mission	happiness	junk
437	wrap	handling	dirt	gps	wrapped
438	minimum	ap	license	requirement	chemistry
439	noob	hilarious	respectful	educated	explode
440	station	remote	wave	controls	motion
441	flight	pt	conscious	intent	commitment
442	steam	edition	dlc	ii	bundle
443	dark	dragon	ice	tower	counter
444	company	office	manager	contract	management
445	link	context	undelete	heres	vids
446	armour	homeless	tim	wa	leveled
447	trailer	summon	reveal	freezes	cord
448	points	bonus	bonuses	extra	pos
449	touch	requirements	require	gems	editing
450	night	hours	hour	morning	sleep
451	bang	catching	stance	ut	mlb

452	park	balance	guest	crafting	guests
453	print	bc	printer	turkey	printed
454	properly	article	extended	india	aspect
455	war	strategy	civil	organization	strongly
456	external	tune	slave	trusted	leak
457	business	success	production	successful	digital
458	china	chinese	europe	russian	sea
459	hair	eyes	natural	thin	naturally
460	tape	mc	protected	ceremony	transferring
461	event	events	proof	gen	legit
462	area	near	areas	snow	rewards
463	metjm	ft	ingame	nonesouvenir	mw
464	type	example	create	model	types
465	history	stories	film	modern	horror
466	ball	adamant	jolly	timid	modest
467	looked	head	eyes	face	felt
468	build	drop	gear	weapons	pvp
469	hear	heard	voice	hearing	acting
470	question	questions	ask	thread	answer
471	name	names	named	alt	undergrad
472	working	job	position	worked	jobs
473	rick	chasing	bash	cr	horribly
474	dr	tl	doctor	medical	hospital
475	share	sharing	proud	backstory	shares
476	picture	background	image	squad	colour
477	control	panel	jokes	disk	denied
478	jan	sat	dash	motor	mon
479	ground	fly	flying	paint	plane
480	cock	ass	cum	dick	mouth
481	inch	commander	curiosity	avatar	tendency
482	damage	shield	enemies	enemy	health
483	players	player	win	solo	tournament
484	cities	tall	manner	civ	cigarettes
485	bug	arena	bugs	robot	cm
486	learn	basic	learning	knowledge	reference
487	super	excited	zero	mario	loose
488	mode	rank	material	skip	materials
489	cs	hr	mount	camping	hiking
490	friends	friend	guy	girl	talking
491	dream	reality	dreams	quote	bucks
492	recommendations	request	shared	requested	complaints

493	lost	lose	losing	loss	winning
494	rush	karma	creator	gravity	complaint
495	bigger	expansion	doc	breast	directed
496	peace	lg	israel	relax	homes
497	camp	bits	modes	poorly	mmr
498	number	line	data	lines	numbers
499	coming	soon	release	released	anytime
500	truth	podcast	dare	entertaining	skull

Appendix 3: Instructions for Future Crovitz Task

Below, I describe the full instructions presented to participants in the Future Crovitz task in Chapter 4.

At the beginning of the task, participants received the following instructions:

> In this task, you will be presented with a word on the screen, for example, "WATER."
>
> Your task is to type some text in response to this word, which will be either a FUTURE event or a PAST event, depending on the trial.
>
> You will see a total of 60 words, in 2 sets of 30 words. Before each set of 30 words you will be given more specific instructions.

Before beginning the *future* block, participants received the following instructions:

> In this part of the task, your task is to type in a FUTURE event in response to each word you see. A future event is defined a single and specific personal event event that is very likely to happen in the future. The event should occur in a specific place and time, and should involve you as the main character. For example, in response to the word "WATER", you might write about your plans to go to the Water Park one day during the next Summer Break.
>
> You may use the word you see as inspiration for the event, or you may type another event of your choosing.

The instructions for the *future* block were the same as the instructions for the past block, except references to the future were replaced by references to the past.

Appendix 4: Lab-Based Topic Model

In Chapter 4, I report an attempt to train a topic model using the talk derived from the lab-based Future Crovitz task. Below, I list the full schemas learned by this model, where each schema is represented by the most probable words in the topic. Visual inspection reveals that the majority of these schemas are semantically incoherent, despite the model being trained with the same parameters as the model in Chapter 3.

Topic	Term #1	Term #2	Term #3	Term #4	Term #5
1	learned	opinions	heard	see	small
2	next	little	tuesday	ghost	thought
3	money	make	working	wind	reread
4	trouble	knew	fruit	plans	smoothies
5	sunlight	hurt	south	france	wall
6	feel	sometime	older	shorter	weekends
7	see	balcony	home	back	mom
8	hard	man	hat	white	many
9	's	thought	rate	friend	ready
10	plan	instead	squares	draw	shapes
11	task	next	alwyay	things	jordan
12	one	day	like	window	somewhere
13	came	hours	jacket	company	major
14	cubes	next	thinking	n't	really
15	star	shooting	saw	ago	blossoming
16	thinking	pleasure	virtues	therys	class
17	cox	grade	2nd	field	tell
18	plant	house	rain	wardrobe	's
19	see	eventually	sundays	strong	relationship
20	five	six	years	popcorn	tattoos
21	drive	street	car	summer	see
22	tree	playground	behind	called	school
23	warmth	fireplace	's	desert	israel
24	get	back	come	please	unlocked
25	setting	special	next	remeber	plant
26	french	else	liked	january	essay
27	said	got	beach	sunrise	's
28	revolt	news	reading	saw	drove

29	need	people	virtuous	excuse	helped
30	wearing	expensive	conference	clothing	people
31	blossom	spring	next	would	like
32	many	come	cute	go	people
33	last	riverdale	pool	episode	beat
34	gave	california	advice	read	news
35	animal	believe	played	years	butter
36	4	always	old	becuase	would
37	calender	n't	hating	wanted	visit
38	married	geometry	pay	loans	fell
39	florida	moved	house	annoying	power
40	quite	blinds	stars	yelling	growing
41	movie	beautiful	watched	caribbeans	one
42	uber	's	sometime	throw	full
43	threat	poses	remember	's	note
44	would	ambulance	taken	hopefully	stabbed
45	's	broke	friend	last	fix
46	physics	calculate	joy	photos	hang
47	house	back	go	next	park
48	boat	swim	's	deck	took
49	fire	still	flames	troubles	left
50	amount	arge	stars	camp	spent
51	said	daughter	looked	pretty	camera
52	math	personal	table	studying	study
53	sister	africa	every	week	street
54	hiding	true	one	better	margarine
55	experiment	waking	finish	houses	humanity
56	really	session	n't	3	next
57	spending	talking	involves	ethics	etc
58	walk	street	fashion	next	early
59	brother	dad	dog	fire	years
60	school	elementary	middle	fact	horses
61	debit	instead	card	polo	rolling
62	's	babysit	textbook	ten	thursday
63	hike	go	place	work	wll
64	years	mount	highly	job	levels
65	hall	running	council	imagine	across
66	pakistan	face	car	remember	reality
67	test	license	telling	would	go
68	menace	dennis	word	hated	little
69	mind	comes	classes	speaking	buddha
70	often	much	niece	nephews	love
71	swimming	go	one	's	classroms

72	square	relieved	finished	revolution	eqypt
73	younger	cousins	see	next	knew
74	club	loving	bee	remember	animations
75	chirping	cleaning	good	horror	movie
76	phone	speak	likely	upset	central
77	usually	movies	...	see	action
78	cold	crowded	part	night	coming
79	participating	forward	looking	arrived	girls
80	camping	trip	eggs	fire	gone
81	become	anything	community	see	three
82	circle	us	dance	go	lame
83	nectar	juice	drink	mango	back
84	picturing	hiking	years	ago	mexico
85	tomorrow	bank	news	something	amazon
86	ago	met	living	go	air
87	family	potter	harry	reading	old
88	butter	real	picked	pancakes	bought
89	arguing	games	use	minutes	seven
90	scared	got	really	gas	ever
91	kind	roommate	meet	's	try
92	name	books	read	time	would
93	planting	closet	room	news	got
94	hit	doodle	like	head	dad
95	storm	snow	excuse	museum	lose
96	room	bought	n't	emory	times
97	eating	salad	ducling	healthy	start
98	botanical	go	's	garden	gardens
99	childhood	strive	ball	hit	strikeing
100	mom	n't	trash	properly	see
101	asked	going	different	opinion	spanish
102	sky	fireworks	july	tank	forties
103	opinions	like	past	opinion	moment
104	finsihed	menance	google	typing	mean
105	dont	know	could	hawaii	excuses
106	''	``	trouble	man	highway
107	ocean	open	breakfast	doors	likely
108	favorite	rattle	's	toy	baby
109	groceries	grandma	every	3	feel
110	making	sandwich	see	go	got
111	colors	ago	lving	ring	fulfill
112	lunch	year	school	family	healthier
113	able	instead	lost	red	've
114	buy	han	visited	korea	south

115	feel	cherish	like	go	city
116	one	boarding	go	day	cherry
117	weekend	city	newyork	going	see
118	brownies	remake	sing	eve	section
119	school	elementary	forgot	home	front
120	take	semester	next	cross	helped
121	group	listen	fun	actually	thought
122	mental	image	grass	fishing	power
123	month	next	one	living	travelling
124	flowers	planted	make	decisions	able
125	friends	around	travel	world	lie
126	like	lazy	go	would	succulent
127	give	every	job	day	make
128	priest	side	friend	excuses	confessor
129	kids	years	last	wrong	india
130	church	priest	sunday	go	left
131	always	wanted	little	rich	miserables
132	doll	9	months	maybe	around
133	last	's	beautiful	senior	tournament
134	child	never	day	care	least
135	wo	n't	soccer	start	grew
136	sisters	housefull	movies	funny	specifically
137	make	sound	cake	play	next
138	son	's	scared	handing	september
139	probably	go	planning	since	emory
140	english	elevator	express	writing	ap
141	errands	people	ran	army	position
142	see	around	years	age	30
143	freshmen	year	next	disagrees	least
144	sick	would	day	next	school
145	offering	fire	seller	end	rent
146	taking	experience	yard	remember	every
147	birthday	present	one	really	using
148	flower	days	small	ago	flight
149	trying	staying	room	night	late
150	learn	13	relax	's	street
151	joyful	named	max	enjoying	see
152	power	friends	n't	go	house
153	homework	complete	excuse	n't	hearing
154	clothing	mall	lenox	coming	choosing
155	eat	salad	hate	burger	time
156	angry	crashing	computer	historial	revolts
157	joyfully	highlight	nickname	o'neal	paul

158	boyfriend	cherry	blossoms	go	bridge
159	school	high	got	tv	caught
160	family	n't	floating	raft	june
161	kid	's	valentine	day	choking
162	explain	excuses	try	ones	make
163	loved	bees	joy	one	fireplace
164	went	eight	hawaii	cruise	bank
165	room	source	lent	accidently	venice
166	cars	visiting	see	duc	hoped
167	emory	attend	rode	horse	took
168	drinking	started	got	room	trouble
169	tell	afraid	photo	responsibilities	ambulance
170	one	hearing	excuse	made	sounds
171	dc	see	whatever	want	day
172	food	truck	mirror	years	front
173	package	thoughts	clothings	hide	let
174	snakes	's	little	super	always
175	brownies	remake	sing	eve	section
176	beach	house	friend	visiting	cherry
177	weekend	try	glass	day	borrow
178	school	see	coming	online	south
179	ship	riding	one	cold	winter
180	airport	interview	moments	earthquake	visit
181	grandmother	's	backyard	bush	early
182	square	shapes	park	one	central
183	attending	hold	fighting	see	connected
184	entertainment	's	people	walks	past
185	writing	question	rivers	early	response
186	see	snake	next	go	hiking
187	person	babysitter	kindest	still	excuses
188	middle	school	hot	green	wrapped
189	fall	's	talking	life	atlanta
190	im	5th	taiwan	talking	respect
191	years	ago	turned	see	us
192	shower	use	started	past	events
193	picking	creating	moments	younger	would
194	12	walking	really	went	moment
195	birds	park	coming	feeding	saw
196	bought	clothes	first	last	's
197	ambulance	road	emory	online	side
198	home	go	make	tv	smoothie

199	soon	dining	months	family	next
200	rattling	cookies	bottle	fan	sound
201	set	help	bit	forced	commitment
202	look	math	anymore	done	us
203	asking	one	's	last	jump
204	desk	day	one	showed	blossoms
205	regarding	's	online	read	article
206	friday	held	last	play	missing
207	cruise	complete	equestrian	talking	nonsense
208	clubs	using	difficult	start	motorcycle
209	years	four	born	ago	elated
210	wanting	drawing	wall	admiring	afar
211	talked	spongebob	15	life	kindness
212	like	qualities	middle	math	parallelogram s
213	surprise	friends	birthday	etc	cake
214	growing	uk	pigeons	chase	used
215	call	see	difference	explained	stars
216	black	friend	eye	remember	face
217	standing	window	front	one	another
218	see	visit	blossoms	dc	cherery
219	want	go	visit	like	many
220	weeks	ago	learning	couple	months
221	turn	candle	n't	time	flame
222	played	grades	n't	good	go
223	took	administratio n	present	revolts	fights
224	brownies	remake	sing	eve	section
225	beginning	told	fight	september	year
226	capacity	full	laying	people	next
227	way	spain	phone	hit	looking
228	farm	's	window	house	family
229	talk	restroom	hand	holding	little
230	meaning	word	know	see	old
231	brownies	remake	sing	eve	section
232	money	lot	spent	last	flights
233	perform	next	philosophy	might	year
234	show	kindness	sharing	sleeping	came
235	christmas	crossing	street	till	wait
236	one	would	bedtime	happened	began
237	garden	snake	rattle	saw	one
238	may	next	see	play	visit
239	colorado	top	waving	aspen	hid

240	working	virtue	improve	character	accidentally
241	campus	natural	reserve	near	walk
242	funny	would	comedic	hope	teacher
243	new	horse	riding	york	saturday
244	nice	sit	bad	preach	remember
245	plants	trees	countury	emory	around
246	parents	something	yelled	comes	mean
247	front	welcoming	horse	making	fireplace
248	car	engine	street	got	see
249	faced	problems	trip	road	past
250	street	car	places	upset	people
251	play	game	dare	next	museum
252	series	school	one	read	free
253	reading	ferrari	would	present	mind
254	brownies	remake	sing	eve	section
255	picture	fixing	garage	along	personalities
256	important	7	years	father	often
257	pretty	midtown	really	question	7
258	test	science	object	calculate	due
259	apartment	plant	means	hide-and-seek	shopping
260	calendar	bright	shining	looking	degree
261	everything	finally	felt	got	patch
262	back	would	grade	9th	time
263	melting	reason	reached	butter	storage
264	hugging	brothers	took	see	everybody
265	going	park	national	big	sociology
266	butter	bread	breakfast	ducling	next
267	building	might	wedding	let	oven
268	studying	woman	roommates	leftovers	fridge
269	pick	running	see	money	prettiest
270	windows	scary	notes	last	pacifiers
271	opinion	asked	cookie	dad	eaten
272	telling	middle	4th	summers	2
273	nicer	parents	brother	car	``
274	father	strike	bill	hazare	lok
275	mid				
276	future	schedule	change	registration	dog
277	opened	day	door	someone	miserables
278	school	main	would	right	working
279	street	looking	singing	professor	really
280	learning	wiht	history	revolution	russian
281	last	day	yummy	ate	avocado
282	small	emory	quad	different	excluded

283	much	powerpuff	december	joy	like
284	revolting	people	textbooks	past	articles
285	belt	shower	time	nothing	comes
286	fall	'm	event	people	part
287	bird	see	's	humming	feeder
288	everyone	easter	egg	always	family
289	tonight	moment	text	good	take
290	near	honey	tea	emory	bubble
291	manchester	united	would	coming	thanksgiving
292	problems	sorry	money	friend	man
293	art	last	bee	rise	hiked
294	restaurant	clear	everywhere	sitting	many
295	book	read	reading	religion	see
296	last	's	ambulance	ago	police
297	mountain	stone	go	like	would
298	night	last	stars	watched	family
299	'm	see	powerpuff	week	bunch
300	round	access	credit	card	use
301	joy	across	feel	singing	world
302	college	family	july	freshman	tubing
303	wish	away	told	life	ends
304	onto	spread	butter	bagel	seed
305	car	accident	got	remember	find
306	sleep	holding	bank	reading	mediums
307	seeing	thanksgiving	friends	best	girl
308	popcorn	see	circle	circus	preform
309	friend	best	red	senior	science
310	seen	n't	almost	really	months
311	involvement	newspaper	athletics	writing	single
312	things	classroom	like	vegetables	math
313	debate	documentary	dad	hide	comes
314	listening	stand-up	talk	fishing	make
315	thinking	relatives	grandfather	west	live
316	truth	live	people	spread	nonreligious
317	hummingbird s	kept	turkey	window	see
318	today	later	go	's	n't
319	summer	japan	end	airplane	coming
320	happiness	reach	meet	able	day
321	many	see	last	ambulance	faught
322	someone	office	flying	plane	land
323	avoid	salads	awful	today	child
324	someday	last	game	--	lights

325	good	grade	india	think	orphanages
326	tennis	police	brother	match	felt
327	astronomy	facebook	one	freshman	year
328	cvs	shampoo	read	books	list
329	children	hospital	's	volunteering	played
330	dream	ago	celebrity	years	les
331	year	last	dad	met	much
332	left	people	'm	interact	asked
333	's	drawing	one	ripples	admiring
334	ways	atlanta	later	see	smelling
335	people	krispes	achieve	rice	's
336	sleeping	living	early	cry	classes
337	everyday	would	waiting	childern	earn
338	great	bwah	make	decisions	movie
339	life	filled	spend	joy	back
340	climbed	streets	province	huagai	lap
341	soccer	factory	away	cheesecake	running
342	hill	station	holiday	go	another
343	online	budget	counting	chekcing	overspend
344	zoo	snakes	hate	cage	're
345	buddy	gift	waterpolo	giving	godparents
346	opinion	discouraged	offered	though	even
347	works	revolution	russian	store	rich
348	nature	game	hamper	fireplace	day
349	yesterday	ambulance	go	saw	travel
350	paid	ask	tell	example	got
351	might	's	rafey	cream	final
352	cousin	baby	adopted	greet	warmly
353	library	books	returned	last	time
354	's	closed	anatomy	grey	midnight
355	window	cuddle	breeze	open	smells
356	expressing	another	one	independent	confident
357	ate	last	cooked	rice	worked
358	doctor	ambulance	hope	's	mom
359	return	seattle	stay	christmas	reminds
360	class	Topic #| sat	root	calculus	number
361	china	rivers	different	geography	little
362	december	back	bedroom	hampshire	new
363	really	english	finding	n't	11th
364	honeysuckle	''	understand	trying	old
365	running	work	currently	time	balance
366	put	shut	see	resteraunt	yellowclaw
367	party	birthday	cake	bake	using

368	freshman	studied	year	french	google
369	next	november	blue	costa	rica
370	cancer	sad	pyshcology	ridding	reading
371	used	draw	squares	pre-k	plan
372	break	winter	visit	holidays	time
373	lots	sugar	bake	butter	cookies
374	people	new	wonderful	wednesday	club
375	sure	made	top	work	excuses
376	start	died	car	battery	engine
377	grow	clean	walmart	int	much
378	watch	comedy	movie	see	go
379	shirt	find	hide	maybe	boyfriend
380	one	morning	slice	flight	beijing
381	ride	horses	go	would	grade
382	year	could	outside	senior	years
383	midwest	falling	would	flops	minnesota
384	times	titanic	watched	massive	next
385	weather	trouble	cold	called	coat
386	's	feeling	canoeing	rattlesnake	tattoo
387	calling	wish	leaving	laundry	healthier
388	run	stressed	many	without	errandsd
389	excuse	lab	make	gave	able
390	screaming	pain	last	present	ambulance
391	administratio n	sheridan	street	new	never
392	cat	started	vase	scold	broke
393	got	pink	started	service	coding
394	water	happy	rafting	white	see
395	2	downtown	writing	placing	explication
396	hide	seek	playing	go	behind
397	toast	called	likely	lived	``
398	bed	blanket	extra	pull	power
399	close	see	priest	next	back
400	norway	horse	friend	vacation	countury
401	tommorow	snakes	rattle	paper	school
402	climbing	pay	minutes	30	korea
403	enjoy	f	haiti	noise	alays
404	daycare	shows	like	's	miami
405	find	position	power	roommate	setting
406	old	steps	years	vase	use
407	dress	feel	rides	adult	life
408	opening	window	roommate	one	street
409	warm	peach	drinking	see	nectar

410	dressed	outfit	friend	gets	hoenst
411	sitting	chair	next	front	little
412	study	keep	alive	one	day
413	love	heat	iphone	often	people
414	closing	hopes	better	ability	teaching
415	walking	street	see	's	excuse
416	gym	working	come	meetings	late
417	behind	means	locked	one	like
418	know	n't	better	around	even
419	climb	large	diet	go	aspire
420	labeled	postdoc	lab	houston	dozens
421	atlanta	excited	visit	's	market
422	longer	day	comic	wife	wait
423	week	next	psychology	test	preparation
424	told	hot	upcoming	never	say
425	fly	city	next	york	new
426	green	ate	lunch	85th	called
427	camp	grade	last	built	fire
428	help	people	save	ceiling	one
429	first	name	street	senior	prom
430	lock	door	properly	's	tania
431	mountains	bulgaria	shop	next	summeer
432	graduation	afternoon	see	later	talking
433	girl	nephew	meet	's	go
434	traveling	laughing	nights	dry	youtube
435	morning	street	crossed	busy	way
436	one	teachers	date	due	push
437	3	back	moment	took	least
438	hope	days	lighten	volunteer	've
439	blooming	laptop	flowers	go	see
440	watering	breeze	square	chemistry	kid
441	brownies	remake	sing	eve	section
442	girlfriend	lunch	'm	--	evil
443	fire	went	friends	one	end
444	getting	post	go	something	run
445	see	next	temple	lot	besides
446	babysitting	chinese	grandparent s	washington	history
447	sister	go	school	dinner	pick
448	go	rent	paying	laundry	outfitters
449	word	learnt	``	see	''
450	door	open	room	leave	like
451	husband	next	reading	lying	bed

452	helping	others	china	get	activities
453	river	driving	west	american	closest
454	fire	virtues	alarm	goes	able
455	's	teacher	names	wrong	would
456	crows	crying	sun	saw	clothings
457	watching	movie	like	act	protagonist
458	little	hope	day	ambulence	one
459	closing	dressed	coming	italy	ago
460	apple	outdoor	member	family	emory
461	band	revolt	tired	got	sister
462	really	volunteer	14	procrastinate	assignments
463	flower	see	missing	assignments	professors
464	cream	bagel	butter	cheese	parlor
465	student	another	school	talk	tripping
466	mother	's	day	choi	ellie
467	path	less	patient	sucess	chance
468	mom	shopping	grocery	see	couch
469	window	outside	looking	house	apartment
470	zoo	see	think	go	south
471	every	iris	see	go	blossoms
472	's	elderly	dad	filled	got
473	clothes	'll	fall	get	see
474	recently	cousins	play	experience	use
475	playing	game	's	neighbor	video
476	change	world	lights	engine	revolt
477	time	long	go	vacation	physics
478	advertisemen t	garage	guatemala	volcano	see
479	exam	pan	professor	revolt	``
480	taught	song	months	us	remember
481	halloween	taught	visited	stress	emotional
482	count	see	discharged	military	month
483	field	one	go	staff	say
484	haunted	eatting	marriage	gay	subway
485	day	sit	questions	timelapse	bud
486	day	improv	one	kid	good
487	next	europe	driving	since	days
488	care	giving	including	dad	take
489	math	drawer	wallet	box	hid
490	tournament	badly	tennis	's	politics
491	table	dinner	window	screen	ripped
492	errand	find	definition	brought	way
493	'm	next	go	emory	hospital

494	cruise	go	jumping	family	around
495	kitchen	window	table	got	tabletalk
496	two	ago	months	long	waited
497	bar	meant	old	time	lied
498	foward	looking	people	go	excuse
499	dorm	return	room	psychology	research
500	island	sailing	vacation	chance	stores