Title
Constructing the Concept of Time: Roles of Perception, Language, and Culture

Author
Tillman, Katharine Allen

Publication Date
2017

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Constructing the Concept of Time:

Roles of Perception, Language, and Culture

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

Doctor of Philosophy

in

Psychology

by

Katharine A. Tillman

Committee in charge:

Professor David Barner, Chair
Professor Benjamin Bergen
Professor Lera Boroditsky
Professor Victor Ferreira
Professor Gail Heyman

2017

The Dissertation of Katharine A. Tillman is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2017

DEDICATION

For my mother and my daughter — my past, present, and future.

EPIGRAPHS

Time, unfortunately, though it makes animals and vegetables bloom and fade with amazing punctuality, has no such simple effect upon the mind of man. The mind of man, moreover, works with equal strangeness upon the body of time. An hour, once it lodges in the queer element of the human spirit, may be stretched to fifty or a hundred times its clock length; on the other hand, an hour may be accurately represented on the timepiece of the mind by one second.

— Virginia Woolf, *Orlando*

Adults, waiting for tomorrow, move in a present behind which is yesterday or the day before yesterday or at most last week: they don't want to think about the rest. Children don't know the meaning of yesterday, of the day before yesterday, or even of tomorrow, everything is this, now: the street is this, the doorway is this, the stairs are this, this is Mamma, this is Papa, this is the day, this the night.

— Elena Ferrante, *My Brilliant Friend*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Dave Barner's commitment to my development as a scientist has extended far beyond his invaluable contributions to the content of this dissertation. From our very first meeting, Dave has given me the freedom to pursue my interests, and I haven't spent a minute of graduate school working on a project I didn't truly care about. One of the most important lessons Dave has taught me over the past 6 years is that it's not how *much* you work, but *how* you work that matters. Thanks to Dave's influence, I now write (almost) every day, I (almost) never schedule meetings during my best writing times, and my email inbox is (almost) always empty. These practices were critical to finishing this work. Most importantly, Dave's compassion during my first two years of graduate school was the reason I made it to my third.

Thanks are also due to the members of my committee, whose helpful comments improved each paper contained here. The Language and Development Lab, with its many brilliant, kind, and creative members, has been my home and the highlight of my experience at UCSD. I am so grateful to my co-authors on the paper that became Chapter 2, Tyler Marghetis and Mahesh Srinivasan, who gave me the great pleasure of collaborating with dear friends. My co-authors on the manuscript that became Chapter 3, Nestor Tulagan and Eren Fukuda, who were both undergraduates during the bulk of their work on that project, gave me the joy of mentorship. Dozens of other research assistants helped with the enormous amount of data collection and subject recruitment necessary for this research, and are thanked individually in the acknowledgements of each section. In chronological order, lab coordinators Ellie Chestnut, Katie Kimura, Miriam Rubenson, Karen Bejar, and Junyi Chu each provided invaluable support. Hundreds of children, families, preschools, and elementary schools contributed to this research. Thank you especially to the Fleet Science Center, where I have frequently collected data throughout graduate school. Funding for this work was provided by a grant to David Barner from the James S. McDonnell foundation.

Many people believed in my academic career before I did, and their faith also helped to bring this document into being. For most of my time in Denis Pelli's lab at NYU, I insisted that I was not a scientist and had no interest in grad school. Through his extraordinary generosity, Denis kept the door open for me until I finally realized I was wrong. My mother also pointed out that she really didn't think I was "done" after I graduated from NYU, and she was, of course, right. I am very grateful to her and to the rest of my family of origin, who have always encouraged me on this path, even though it has taken me far from home for my entire adult life.

Thank you also to Julie Burelle, Jess Sullivan, and the many other feminist mothers in academia whose powerful examples have meant so much to me. Amber Rukaj's wise advice was essential to navigating work and life during my early years in graduate school, and Caren Walker's friendship and understanding have sustained me during the final stretch. I thank Esther and Avram Shacham, Ariel Gafni, Samantha Milliord, Christina Sandee, Karen Bejar, and the teachers of Hearts Leap North, La Petite Ecole, and the Gillispie School, for providing my daughter with loving childcare and me with the peace of mind I needed to work.

I am most indebted to Hovav Shacham, who has been a fierce advocate for my career from the moment we met. Without the many sacrifices he has made for the sake of our family, my academic work would have been exponentially more difficult, if not impossible. This thesis is dedicated to my mother, Sharon Beth Johnson, and to my daughter, Flannery Tillman Shacham, who made me who I am.

Chapter 1, in full, is a reprint of the material as it appears in Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology, 78*, 57-77. Tillman, K. A., & Barner, D. (2015). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier.

Chapter 2, in full, is a reprint of the material as it appears in Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology, 92*, 87-100. Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier.

Chapter 3, in full, is a reprint of the material currently under review at *Developmental Science.* Tillman, K. A., Tulagan, N., Fukuda, E., & Barner, D. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

2017 Doctor of Philosophy, Psychology
University of California, San Diego

2012 Master of Arts, Psychology
University of California, San Diego

2009 Bachelor of Arts, Psychology
New York University

PUBLICATIONS

Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology*, 92, 87-100.

Wagner, K., Tillman, K. A., & Barner, D. (2016). Inferring number, time, and color concepts from core knowledge and linguistic structure. *Core Knowledge and Conceptual Change*. Oxford University Press.

Tillman, K. A. & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology,* 78, 57-77.

Pelli, D. G., & Tillman, K. A. (2008) The uncrowded window of object recognition. *Nature Neuroscience,* 11(10):1129 – 1135.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007) Crowding and eccentricity determine reading rate. *Journal of Vision,* 7(2):20, 1-36.

Pelli, D. G., & Tillman, K. A. (2007) Parts, wholes, and context in reading: A triple dissociation. *PLoS ONE 2(8)*: e680.

ABSTRACT OF THE DISSERTATION

Constructing the Concept of Time: Roles of Perception, Language, and Culture

by

Katharine A. Tillman

Doctor of Philosophy in Psychology

University of California, San Diego, 2017

Professor David Barner, Chair

Understanding the nature and origin of abstract concepts, like the concept of time, is a fundamental problem in cognitive science. From infancy, humans can discriminate brief durations, represent event sequences, and associate temporal and spatial magnitudes. By adulthood, Westerners construe of time as an abstract dimension, which is described and measured using language, clocks, and calendars. Are mature concepts of time built from innate perceptual primitives? In this dissertation, I will argue that they are not, drawing on developmental evidence from 3- to 8-year-old children. In Chapter 1, I show that children do not learn duration words like “minute” by associating them with perceptual representations of duration. Instead, children's earliest meanings for duration words encode their relations to one another. For example, preschoolers know their relative ordering (e.g., *hour* > *minute* > *second*) long before they know each word’s approximate duration. Similarly, in Chapter 2, I present evidence that children do not

learn deictic time words like “yesterday” by associating them with experienced or anticipated events. I find that children’s earliest meanings for deictic time words include information about their relative order in the past and future, but not about their approximate temporal distance from the present. Both these cases suggest that children initially use linguistic cues to construct ordered semantic domains for time words, and do not map them to perception until later, *after* learning their formal definitions. Finally, in Chapter 3, I present evidence that the left-to-right “mental timeline” English-speaking adults use to organize events is not derived from innate space-time associations. I show that, unlike kindergarteners and adults, preschoolers do not spontaneously represent time linearly. Instead, conventional linear mappings between time and space develop slowly throughout early childhood, in response to increasing cultural exposure and education. Together, these studies suggest that abstract time concepts in children are not built from perceptual primitives, but from structures available in language and cultural artifacts.

INTRODUCTION

Each day, we experience the passing of time, as a succession of events arise in cognitive awareness and fade into memory. Beginning in early infancy, we have a sense of the duration of our experiences, and of their relative ordering over time. We know whether events are happening right now, whether they occurred in the recent or distant past, or how far in the future they might occur. Both our lived experiences and our inner thoughts have temporal structure. However, the experience of time is slippery and ethereal, imprecise and variable. Time seems to fly by when we are cognitively engaged, but to drag on when we are bored. Time seems to speed up as we age, to slow down when we are in danger, and even to stop when we're in "the zone" (for discussion, see Levine, 1997). Therefore, although time is intrinsic to experience, each person's experience of time is intimate, ever-changing, and unique.

Nonetheless, human communities often need to coordinate their activities in order to get things done. To deal with this problem, over the course of history, cultures have developed external symbol systems that precisely describe and measure the passing of time, including clocks, calendars, and words like "minute" and "hour" (e.g., Gell, 1992; Whitman, 1989). The gap between the subjective experience of temporal events and the formal systems we use to describe them is wide. While time is closely tied to our perception of the world, it is difficult to explain on the basis of perception alone. This paradox raises questions that have perplexed psychologists and philosophers for centuries (e.g., Augustine, 398/1992; Bender & Beller, 2014; Boroditsky, 2011; Casasanto et al., 2010; Fraisse, 1963; James, 1890; Kant, 1781/2009; McTaggart, 1908; Piaget, 1969; Whorf, 1956): What is the relationship between our experiences of events and the symbolic systems we use to describe them? How do abstract concepts like time emerge in the mind? Do innate abilities to perceive events provide the building blocks for abstract, linear time, or is this framework entirely culturally constructed? Broadly, these questions are the topic of this dissertation.

Understanding how the abstract concept of time develops in children, as they learn to describe their temporal experiences using symbolic systems, may offer a window into the nature of mature temporal concepts in adults. Although some aspects of time perception, like the capacity to discriminate brief temporal durations, are present from early infancy (e.g., Brackbrill et al., 1967; Brannon et al., 2007; Columbo & Richman, 2010; Gava et al., 2010; Provasi et al., 2010, vanMarle & Wynn, 2006), mastering formal time-keeping systems can take children a decade or more (e.g., Burny et al., 2009; Friedman & Laycock, 1989). Currently, there is no theoretical consensus on how factors like duration perception, language acquisition, and cultural transmission contribute to the development of an adult-like concept of time (e.g., McCormack, 2015; Winter, Marghetis, & Matlock, 2015). Here, as a case study in the development of abstract concepts, I explore how 3- to 8-year-old children acquire three means of expressing time: duration words, deictic time words, and linear representations of time.

Many proposed solutions to the problem of how abstract, uniquely-human thought is possible posit that we start with a set of primitive representations and combine them in such a way that we end up with ideas that are richer than the sum of their parts. Philosophers and developmental psychologists argue both about the nature of the primitives (e.g., simple and general vs. complex and specialized) and about the means by which they may be combined, but the idea that abstract concepts are composed of innate "building blocks" is both old and pervasive (e.g., Locke 1689; Fodor, 1998; see Wagner, Tillman, & Barner, 2016, for discussion). As we will discuss further below, several potential perceptual "building blocks" exist in the domain of time — including representations of duration, events, and spatial magnitudes. In principle, children might use these primitives as a foundation for the later acquisition of time words or ordinal representations of events. However, in this dissertation, I will argue that this is not what children do. Instead, I will argue that children build structured, abstract systems to represent time prior to coordinating these systems with experiences of events in the world (see also Carey, 2009), and that they rely on natural language and external spatial tools, rather than perceptual primitives, to do so (see also Barner, 2017).

In Chapter 1, I will present evidence that duration words like "minute" do not initially encode approximate durations, as mappings between words and events would predict. Instead, I show that children's very first meanings for these words encode their relations with other time words within a semantic domain — *before* the individual words are associated with durations. Similarly, in Chapter 2, I show that children do not assign meanings of deictic time words like "yesterday" by using their relations to lived events. Instead, preschoolers' first meanings for these time words also encode their interrelations with other words referring to either the past or future. Children build an abstract, yet well-ordered, semantic structure for deictic time words nearly two years before these words become individually associated with approximate temporal distances from the present. Finally, in Chapter 3, I will present evidence that the "mental timeline" adults use to represent temporal sequences is not a result of innate spatial representations of time. Instead, I show that both the ordinality and the directionality of the timeline are constructed gradually in childhood, in response to increasing cultural exposure. Together, these studies indicate that symbolic systems for representing time develop separately from the innate capacities underlying event perception, in order to provide a framework for organizing and interpreting experienced events (see also Barner, 2017).

Background Information

The adult concept of time.

A goal of this dissertation is to better elucidate how young children come to think about time as mature adults do. This is of interest in part because, in adulthood, our beliefs about time have widespread influence on our everyday lives, including the way we schedule our days, the way we plan for an uncertain future, and the way we interpret our past experiences and develop a sense of self. However, as philosophers, physicists, and psychologists have long lamented, there is still no universally agreed-upon answer to the question of what time actually *is*. Attitudes about the nature and value of time vary widely across cultures, and have changed considerably throughout human history, particularly with the rise of industrialization (Barnett, 1998; Gell, 1992; Levine, 1997; Whitman, 1989). Indeed, the great diversity of ways in which time has been understood and expressed make it particularly fertile ground for exploring how learned factors like language and culture may interact with evolutionarily-ancient ones like duration perception during conceptual development.

Here, when I speak of the adult's concept of time, I am referring to Western, Newtonian time. Some key elements of this multidimensional construct include that time is **absolute**, **measurable**, and **universal**. Further, time consists of a distinct **past, present,** and **future**. Time is **linear**, flowing continuously and unidirectionally from the past to the future. On this view, time "itself" is a fundamental dimension of the universe that can be described and measured separately from particular events that occupy locations in time (for more discussion, see Bardon, 2013; Newton, 1687; etc). Interestingly, while there is increasing consensus among physicists and philosophers that the Newtonian picture of time does not reflect ground truth about the world (e.g., Bardon, 2013; Einstein, 1819; Sakurai, 1995), both the general public and psychologists have been slower to question its assumptions (e.g., Slife, 1993). If Newton was wrong about the nature of time, how is it that a Newtonian view of time has become unquestionably "common sense" in almost all Western adults?

Importantly, the Western adult's view of time encompasses extensive knowledge of a formal clock and calendar system that exemplifies many of these Newtonian principles. Engagement with this system includes fluency with its conventional units, many of which "chop up" time into precise, but fairly arbitrary chucks. These units have inter-defined meanings: minutes are 60 seconds, hours are 60 minutes, days are 24 hours, weeks are 7 days, and so on[1]. I refer to such meanings as the "adult definitions" of time words. These definitions are couched in a system of mathematical knowledge (i.e., multiplication by 60), which takes many years for children to acquire. Conventional time units are also closely linked to spatial representations of time such as clocks and calendars, throughout children's instruction on time-telling, from Grade 1 to Grade 4 in the U.S. (Common Core State Standards Initiative, 2012). Here, we will focus on what

---

[1] Interesting, the fundamental unit of time, the second, which were once defined as a fraction of the mean solar day, is now defined as "the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom" (BIPM, 2008), a fact few adults ever learn.

children think words like “minute” mean, and how they think about time in general, in the years *before* they become able to use this system fluently.

Primitives in temporal cognition.

Might the well-structured, abstract concept of time found in adults be built, at least in part, from perceptual representations of time available in infancy?

In Chapter 1, we consider the possibility that children use representations of duration to learn duration words. Duration and event perception involve evolutionarily-ancient cognitive capacities. Non-human animals, such as birds and mice, demonstrate a keen sensitivity to elapsed time, when they engage in well-timed food scavenging and caching behaviors in the wild (e.g., Biebach et al, 1989; Clayton & Dickinson, 1998), and adapt to variable temporal intervals between an response and contingent rewards in conditioning paradigms (e.g., Ferster & Skinner, 1957). Human infants as young as 1.5 months of age are also sensitive to the temporal structure of events. For example, infants demonstrate anticipatory autonomic responses (e.g., a change in heart rare) timed to the onset of the next beep or flash in a conditioned temporal sequence (Brackbrill & Fitzgerald, 1972; Columbo & Richman, 2002). By 4 months of age, infants can also discriminate stimuli solely on the basis of their temporal duration (Brannon et al., 2007; Provasi et al., 2011; vanMarle & Wynn, 2006). The acuity of temporal discrimination sharpens throughout early and middle childhood (Brannon et al., 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001). Importantly, psychophysical studies on duration perception in children (and adults) rarely involve stimuli over a minute long, and many common time words label much longer periods. Nonetheless, other physiological processes in humans' evolutionary endowment – i.e., circadian rhythms – are known to operate over considerably longer timescales (e.g., Wager-Smith, & Kay, 2000).

In Chapter 2, we will explore the role of nonlinguistic representations of events in children's acquisition of deictic time words, like "yesterday." In addition to their ability to discriminate the durations of single (brief) events, infants also encode sequences of events, and have an intuitive understanding of the causal relations between them. In the first six months of life, infants can reproduce single actions (e.g., Barr et al, 1996), while 11- to 24-month-olds can imitate 2- and 3-step sequences of actions with increasingly higher accuracy over increasingly long delay periods (Bauer & Mandler, 1989; Bauer & Mandler, 1992; Mandler & McDonough, 1995). Further, infants have a basic understanding of the relationship between temporal order and causality. In violation-of-expectancy paradigms, 4-month-olds are surprised by impossible causal chains of events, including those made impossible by breaks in temporal continuity (Cohen et al., 1998; Leslie, 1982, 1984; Oakes & Cohen, 1990; Spelke et al., 1995). And finally, preschoolers can both verbally judge and spatially represent the relative recency of past events, albeit with lower accuracy than adults (e.g., Friedman, 1991; Friedman, 2003; Friedman & Kemp, 1998; McCormack & Russell, 1997).

Finally, Chapter 3 considers the possibility that innate representations of *space* may form a foundation for the adult-like, linear representation of time. Infants' ability to represent approximate durations bears resemblance to other early perceptual abilities, including their ability to represent the approximate numerosity of sets (Izard et al., 2009; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). The limits of our developing capacity for discrimination of duration, quantity, and length all conform to Weber’s law: they are ratio-dependent and exhibit scalar variability. Infants also appear to spontaneously associate longer spatial magnitudes with longer temporal durations (de Hevia et al., 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010). Furthermore, there is overlap in the brain regions involved in processing time and space (e.g., Corbetta & Shulman, 2002; Pun et al, 2010). Such findings have been variously taken as evidence that space and time rely on a single system of magnitude representations (e.g., Walsh, 2003), for structural similarity between representations of time and space (e.g., Srinivasan & Carey, 2010), and for the idea that representations of time are fundamentally spatial in nature (e.g., Chatterjee, 1999; for review, see Bonato et al., 2012).

Time in language.

While some temporal primitives are in place from infancy, the ability to represent time in language develops more slowly. Despite the long history of debates about the role language plays in shaping the way we think about time (e.g., Whorf, 1956), theoretical accounts of how language acquisition relates to temporal cognitive development in children are rare (Weist, 1989; McCormack, 2015). Time is richly encoded in language, at multiple levels of representation. For instance, many languages include grammatical markers such as verb *tense*, indicating the deictic (past/present/future) status of an action, and *aspect*, indicating whether actions are complete or ongoing (Comrie, 1985). In English, the past-tense marking “*–ed*” is one of the first grammatical inflections children produce, as early as the first year of life (Brown, 1973; DeVilliers & DeVilliers, 1973). Time is also encoded in a large variety of lexical items, including words for durations, time points, temporal relations, etc. Time words are among the most frequent words in English, in both print and speech (Brysbaert & New, 2009; Kucera & Francis, 1967), and are also early to appear in the child's lexicon (e.g., Ames, 1946). Additionally, temporal information is continuously conveyed through discourse and narrative structure, such as the tendency to describe events in chronological order (Jakobson, 1966).

In Chapters 1 and 2, I explore how children first acquire words for durations (e.g., "hour") and time points relative to the present (e.g., "tomorrow"). Although children begin to produce many time words by age 2 or 3, they do not use these words in adult-like ways for several years, until at least age 6 or 7 (Brown, 1973; Busby Grant & Suddendorf, 2011; Busby and Suddendorf, 2005). However, while it is clear that children struggle with time words, past work has revealed less about precisely what children do and do not know about them, during this long delay between first production and mature comprehension. Characterizing the limits of children's understanding of time words, and their first hypotheses about their meanings, may help to illuminate the roles played by factors like event perception and linguistic structure in the initial acquisition of symbolic representations of time.

Interestingly, children's errors with time words suggest that they may have partial meanings, long before acquiring adult definitions (e.g., Harner, 1975; Shatz et al., 2010). For example, when asked questions about duration, 4-year-olds provide inaccurate answers that nonetheless include duration words, suggesting that they know these words denote durations (Shatz et al., 2010). Furthermore, 3-year-olds contrast both "tomorrow" and "yesterday" with "today", suggesting that they know these words refer to non-present times (Harner, 1975). Anecdotal reports of within-category errors, such as over-extending "yesterday" to refer to any past event (e.g., Harner, 1981), also suggest that preschoolers know particular sets of time words comprise common categories.

In the present dissertation, I build upon these previous findings, using new empirical methods to more precisely characterize the time-course of children's knowledge of time words, from age 3 to age 8. Ultimately, of course, children must learn the adult meanings of time words, as defined by their relations to one another (e.g., a *minute* is exactly 60 *seconds*). Here I consider two hypotheses about children's earlier knowledge. On the first hypothesis, children's first meanings for time words are **perceptual**. In this case, they begin by associating time words with events or durations the words are used to describe (e.g., an *hour* is approximately as long as a TV show; *last year* is approximately as long ago as a birthday party). A second hypothesis is that children's very first meanings for time words are (already) **relational** – indicating relations with other time words, rather than associations to events. In this case, children could rely on speech input to infer the relative ordering of terms (e.g., an *hour* is longer than a *minute; last week* is before *yesterday*). As we will discuss in Chapters 1 and 2, these two hypotheses make different predictions about the limits of children's early knowledge of time words, and about which facets of the meanings of time words should be learned earlier vs. later.

Characterizing children's understanding of time words during the long delay between production and adult-like comprehension provides a testing ground for current theories of abstract word learning. For example, as discussed further in Chapter 2, the *syntactic bootstrapping* account — on which children's hypotheses about the meanings of abstract words are guided by grammatical cues in sentences they inhabit (Brown, 1957; Gleitman, 1990; Gleitman et al., 2005; Landau & Gleitman, 1985; Naigles, 1996) — predicts that children will learn aspects of time-word meaning that are also grammatically encoded (e.g., in verb tense) more easily than those that are not. The case of time also provides a new avenue for exploring theories that are most closely associated with other abstract domains, such as number (see Wagner, Tillman, & Barner, 2016 for discussion). Similar to time words, number words are produced by children long before they have adult-like meanings (e.g., Le Corre & Carey, 2011; Wynn, 1990; 1992). Many theories suggest that number word meanings are given by innate, perceptual systems of approximate number representation and/or parallel individuation of objects (e.g., Cantlon, 2012; Gelman & Gallistel, 1978; Leslie, Gelman & Gallistel, 2008; Spelke & Tsivkin, 2001; Wagner & Johnson, 2011). An alternative account, advocated by Carey (2009), posits that most number words are defined by their *inferential roles* within a semantic domain (i.e., they have relational meanings), but that this system is nonetheless "anchored" in perception via a few small number words, which get (perceptual) meanings from innate object representations. A third account, recently proposed by Barner (2017), rejects this framework, instead arguing that number words and other symbolic representations of number "do not get their content *from* perception, but instead arise to explain it." We will return to these accounts and their relation to the case of time in the General Discussion.

Spatial representations of time.

In addition to representing time linguistically, societies across the world and throughout human history have developed many means of visually representing time. These include tools like sundials, clocks, and calendars, and graphical representations like charts and timelines (Gell, 1992; Whitman, 1989; Rosenberg & Grafton, 2013). Spatial metaphors for time (e.g., *I'm looking forward to putting my dissertation defense behind me*) are also widespread across the world's languages (Haspelmath, 1997; Grady, 1999), and the practice of reading and writing habitually associates progress in narratives with a particular spatial direction. A large body of evidence suggests that adult speakers of English and other languages that are written from left-to-right organize temporal sequences using a left-to-right "mental timeline" (MTL) from the past to the future (for a review, see Bonato et al., 2012). However, it remains unclear (and debated) whether cultural tools and practices linking space and time draw upon pre-existing, potentially-innate spatial representations of time in the mind, or, alternatively, if engaging in these cultural practices *causes* space and time to become mapped in the mind (e.g., Casasanto & Bottini, 2013; Lakoff & Johnson, 1980; Moore, 2006; Walsh, 2003; Whorf, 1956; Winter, Marghetis, & Matlock, 2015). I tackle this issue in Chapter 3, which explores the development of linear spatial representations of time in children.

As mentioned above, several studies in infants indicate that language and cultural experience are not required to associate temporal duration and spatial length (de Hevia et al., 2014; Brannon et al., 2007; Laurenco & Longo, 2010; Srinivasan & Carey, 2010), and some researchers argue that the MTL is innate (Chatterjee, Southwood, & Basilico, 1999; Vicaro et al., 2007). However, cross-cultural comparisons in adults and school-aged children show that the direction and orientation of space-time mappings vary widely, suggesting that they are a learned convention (Bergen & Lau, 2012; Boroditsky, 2001; Boroditsky et al., 2011; Boroditsky & Gaby, 2010; Brown, 2012; Fuhrman & Boroditsky, 2010; Lai & Boroditsky, 2011; Miles et al., 2011; Ouellet et al., 2010;

Nachson, 1983; Nunez & Sweetser, 2006; Tversky, Kugelmass, & Winter, 1991). Prior to the current work, it remained unknown whether children who have not yet entered school also create conventional linear representations of time. If the mental timeline is a result of a biological predisposition, English-speaking preschoolers, like adults and kindergarteners, should spontaneously represent time linearly, in the conventional LR fashion. Alternatively, if cultural conditioning is necessary for the deployment of the MTL, this behavior should emerge only after children begin to receive instruction on reading/writing and cultural artifacts for time.

In the chapters that follow, I will describe how children learn duration words (Chapter 1), deictic time words (Chapter 2), and build linear representations of time (Chapter 3). Finally, in the General Discussion, I will tie all three studies together and talk further about their implications and directions for future work.

References

Ames, L. B. (1946). The development of the sense of time in the young child. *The Pedagogical Seminary and Journal of Genetic Psychology*,*68*(1), 97-125.

St. Augustine (398/1992) *Confessions*. (H. Chadwick, Trans.) Oxford: Oxford University Press.

Bardon, Adrian (2013). *A Brief History of the Philosophy of Time.* Oxford University Press.

Barner, D. (2017). Language, procedures, and the non-perceptual origin of number word meanings. *Journal of Child Language*, 1-38.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6-to 24-month-old infants. *Infant behavior and development*, *19*(2), 159-170.

Bauer, P. J. & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. Developmental Psychology, 25, 197–206.

Bauer, P. J., & Mandler, J. M. (1992). Putting the horse before the cart: The use of temporal order in recall of events by one-year-old children. Developmental Psychology, 28(3), 441.

Bergen, B. K., & Lau, T. T. C. (2012). Writing direction affects how people map space onto time. *Frontiers in psychology*, *3*.

Biebach, H., Gordijn, M., & Krebs, J. R. (1989). Time-and-place learning by garden warblers, Sylvia borin. Animal Behaviour, 37, 353-360

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2257-2273.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*(1), 1–22.

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently?. *Cognition*, *118*(1), 123-129.

Brown, R. W. (1957). Linguistic determinism and the part of speech. The Journal of Abnormal and Social Psychology, 55(1), 1.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, *37*(4), 395–402. http://doi.org/10.1016/S0028-3932(98)00108-0

Boroditsky, L., & Gaby, A. (2010). Remembrances of times East: absolute spatial representations of time in an Australian aboriginal community. *Psychological Science*, *21*(11), 1635–1639. http://doi.org/10.1177/0956797610386621

Brackbill, Y., Fitzgerald, H. E., & Lintz, L. M. (1967). A developmental study of classical conditioning. *Monographs of the society for research in child development*, 1-63.

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. Developmental science, 10(6), 770-777.

Brown, R. (1973). A first language: The early stages. Harvard U. Press.

Brown, P. (2012). Time and space in Tzeltal: Is the future uphill? *Frontiers in Psychology*, *3*(JUL).

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior research methods, 41(4), 977-990.

Burny, E., Valcke, M., & Desoete, A. (2009). Towards an agenda for studying learning and instruction focusing on time-related competences in children. *Educational Studies*, *35*(5), 481-492.

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. Cognitive Development, 20, 362–372.

Busby Grant, J. & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*, 87-96.

Cantlon, J. F. (2012). Math, monkeys, and the developing brain. PNAS, 109, 10725-10732.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and time in the child's mind: Evidence for a cross-dimensional asymmetry. *Cognitive Science*, *34*(3), 387-405.

Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gesture in English speakers. Cognitive Linguistics, 23(4), 643–674.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. Nature, 395(6699), 272-274.

Cohen, L. B., Amsel, G., Redford, M. A., & Casasola, M. (1998). The development of infant causal perception. In A. Slater (Ed.), Perceptual development: visual, auditory, and speech perception in infancy (pp. 167–209). East Sussex, UK: Psychology Press Ltd

Colombo, J., & Richman, W. A. (2002). Infant timekeeping: Attention and temporal estimation in 4-month-olds. *Psychological Science*, 13, 475-479.

Common Core State Standards Initiative. (2011). *Common core state standards for mathematics*.

Comrie, B. (Ed.). (1985). *Tense* (Vol. 17). Cambridge University Press.

Cooperrider, K., & Núñez, R. (2009). Across time, across the body: Transversal temporal gestures. Gesture, 9(2), 181-206.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. Neuroscience, 3(3), 201–15.

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve semantic induction?. Cognition, 123(1), 162-173.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, *111*(13), 4809-4813.

De Villiers, J. G., & De Villiers, P. A. (1978). Language acquisition. Harvard University Press.

Droit-Volet, S., Clément, A., & Wearden, J. (2001). Temporal generalization in 3-to 8-year-old children. Journal of Experimental Child Psychology, 80(3), 271-288.

Droit-Volet, S., & Wearden, J. H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology*, *80*(2), 142-159.

Droit-Volet, S., Tourret, S., & Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *The Quarterly Journal of Experimental Psychology, 57A,* 797-818.

Earnest, D. (2017). Clock Work: How Tools for Time Mediate Problem Solving and Reveal Understanding. *Journal for Research in Mathematics Education*, *48*(2), 191-223.

Ferster C.B & Skinner B.F (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Fraisse, P. (1963). The psychology of time. New York: Harper & Row.

Friedman, W. J. (1986). The development of children's knowledge of temporal structure. *Child development*, 1386-1400.

Friedman, W. J. (2003). The development of a differentiated sense of the past and the future. *Advances in child development and behavior*, *31*, 229-269.

Friedman, W. J., & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development*, 357-371.

Fuhrman, O., Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: evidence from an implicit nonlinguistic task. *Cognitive Science* 34, 1430—1451.

Gell, A. (1992). The anthropology of time: cultural constructions of temporal maps and images. Oxford: Berg.

Gelman, R., & Gallistel, C. (1978). The child's understanding of number. Cambridge, `MA: Harvard University Press.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition*, *87*(3), B87-B95.

Gleitman, L. (1990). The structural sources of verb meanings. Language Acquisition, 1(1), 3–55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. Language Learning and Development, 1, 23–64.

Grady, J. (1999). A typology of motivation for metaphor: Correlations vs. resemblances. In R. Gibbs & G. Steen (Eds.), *Metaphor in cognitive linguistics* (pp. 79—100). Amsterdam, the Netherlands: Benjamins

Grant, J. B. & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*, 87-96.

Haspelmath, M. (1997). *From space to time: Temporal adverbials in the world's languages*. Lincom Europa.

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. Developmental Psychology, 11(6), 864–865.

Harner, L. (1981). Children's understanding of time. Topics in language disorders, 2(1), 51-66.

Einstein, A. (1916). The Foundation of the General Theory of Relativity. *Annalen Der Physik*, *354*, 769–822.

Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex*, *44*(4), 454-461.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*,*106*(25), 10382-10385.

Jakobson, R. (1966). Question for the essence of language. *Diogenes*, 51, 21-37.

James, W. (1890). *The principles of psychology* (Vol. 1). *New York: Holt*.

Kamii, C., & Russell, K. A. (2012). Elapsed time: why is it so difficult to teach?. *Journal for Research in Mathematics Education*, *43*(3), 296-315.

Kant, I. (2009/1781). *Critique of pure reason*. P. Guyer (Ed.). Cambridge University Press.

Kucera, H, & Francis, W. N,.(1967). Computational Analysis of Present-day American English. Providence: Brown University Press

Labrell, F., Mikaeloff, Y., Perdry, H., & Dellatolas, G. (2016). Time knowledge acquisition in children aged 6 to 11 years and its relationship with numerical skills. *Journal of experimental child psychology*, *143*, 1-13.

Lai, V. T., & Boroditsky, L. (2013). The immediate and chronic influence of spatio temporal metaphors on the mental representations of time in English, Mandarin, and Mandarin-English speakers. *Frontiers in Psychology*, *4*(APR).

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. *Chicago: U of Chicago P*.

Landau, B., & Gleitman, L. R. (1985). Language and experience: Evidence from the blind child. Cambridge, MA: Harvard University Press.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*. 21, 873–881.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438.

Le Guen, O., Balam, P., & Ildefonsa, L. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. Frontiers in psychology, 3, 271.

Levine, R. N. (2008). A geography of time: On tempo, culture, and the pace of life. Basic Books.

Leslie, A. M. (1982). The perception of causality in infants. Perception, 11, 173–186

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. Perception, 13, 287–305

Leslie, A. M., Gelman, R., & Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends in cognitive sciences*, *12*(6), 213-218.

Locke, J (1690/1964) In AD Woozley (Ed.), An essay concerning human understanding. *Cleveland, OH: Meridian Books*.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. Child Development, 76(5), 978-988.

Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events nature or culture?. *Psychological Science*, *14*(4), 296-301.

Mandler, J. M., & McDonough, L. (1995). Long-term recall of event sequences in infancy. *Journal of Experimental Child Psychology*, *59*(3), 457-474.

McCormack, T. (2015). The development of temporal cognition. In R. Lerner, L. Liber, U. Mueller (Eds.), *Handbook of Child Psychology and Developmental Science, Volume Two, Cognitive Processes*, 7th Edition.

McCormack, T., & Russell, J. (1997). The development of recency and frequency memory: Is there a developmental shift from reliance on trace-strength to episodic recall?. Journal of Experimental Child Psychology,66(3), 376-392.

McTaggart, J. (1908) The unreality of time. *Mind* 17, 457–474

Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (2011). Can a mind have two time lines? Exploring space—time mapping in Mandarin and English speakers. *Psychonomic bulletin & review*, *18*(3), 598-604.

Nachson, I. (1983). Directional preferences in bilingual children. *Perceptual and Motor Skills*, 56, 747—750.

Naigles, L. (1996). The use of multiple frames in verb leaning via syntactic bootstrapping. Cognition, 58, 222–251.

Newton, I. (1687). Philosophiae Naturalis Principia Mathematica. *Pan*, 510.

Núñez, R. E., & Sweetser, E. (2006). With the future behind them: convergent evidence from aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science*, *30*(3), 401–450. http://doi.org/10.1207/s15516709cog0000_62

Oakes, L., & Cohen, L. (1990). Infant perception of a causal event. Cognitive Development, 5, 193–207.

Ouellet, M., Santiago, J., Israeli, Z., & Gabay, S. (2010). Is the future the right time? *Experimental Psychology*, *57*(4), 308–314

Piaget, J. (1969). *The child's conception of time*. Routledge & K. Paul.

Pouthas, V., Droit, S., & Jacquet, A. Y. (1993). Temporal experiences and time knowledge in infancy and early childhood. Time & Society, 2(2), 199-218.

Provasi, J., Rattat, A. C., & Droit-Volet, S. (2011). Temporal bisection in 4-month-old infants. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 108.

Pun, C., Adamo, M., Weger, U. W., Black, S. E., & Ferber, S. (2010). The right time and the left time: spatial associations of temporal cues affect target detection in right brain-damaged patients. *Cognitive neuroscience*, *1*(4), 289-295.

Rosenberg, D., & Grafton, A. (2013). *Cartographies of time: A history of the timeline*. Princeton Architectural Press.

Sakurai, J. J. (1995). *Modern Quantum Mechanics, Revised Edition*. *American Journal of Physics* (Vol. 63).

Santiago, J., Lupáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512-516.

Schaeffer, B., Eggleston, V. H., & Scott, J. L. (1974). Number development in young children. *Cognitive Psychology*, *6*(3), 357–379.

Sinha, C., Sinha, V. D. S., Zinken, J., & Sampaio, W. (2011). When time is not space: the social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Language and Cognition*, *3*(1), 137-169.

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11,* 16-36.

Slife, B. D. (1993). *Time and psychological explanation*. SUNY press.

Spelke, E., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In Bowerman, M. & Levinson, S. (Eds.). Language Acquisition and Conceptual Development. Cambridge: Cambridge University Press, pp 70-100.

Srinivasan, M., & Carey, S. (2010). The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217-241.

Suddendorf, T. (2010). Linking yesterday and tomorrow: Preschoolers' ability to report temporally displaced events. *British Journal of Developmental Psychology*, *28*(2), 491 498.

Sullivan, J., & Barner, D. (2014). The development of structural analogy in number-line estimation. *Journal of Experimental Child Psychology*, *128*, 171–189.

Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science*, *30*(4), 745-757.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, *23*(4), 515–557.

Vallesi, A., Arbula, S., & Bernardis, P. (2014). Functional dissociations in temporal preparation: evidence from dual-task performance. *Cognition*, *130*(2), 141-151.

Vicario, C. M., Caltagirone, C., & Oliveri, M. (2007). Optokinetic stimulation affects temporal estimation in healthy humans. *Brain and cognition*, *64*(1), 68-73.

Wager-Smith, K., & Kay, S. A. (2000). Circadian rhythm genetics: from flies to mice to humans. Nature genetics, 26(1), 23-27.

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. Cognition , 119 (1), 10-22.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127*(3), 307-317.

Wagner, K., Tillman, K. A., & Barner, D. (2016). Inferring number, time, and color concepts from core knowledge and linguistic structure. *Core Knowledge and Conceptual Change.* Oxford University Press.

Walker, E., & Cooperrider, K. (2016). The Continuity of Metaphor: Evidence From Temporal Gestures. *Cognitive science*, *40*(2), 481.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, *7*(11), 483-488.

Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, *15*(2), 426-430.

Weist, R. M. (1989). Time concepts in language and thought: Filling the Piagetian void between two and five years. In I. Levin & D. Zakay (Eds.), Time and human cognition: A life-span perspective (pp. 63–118). Amsterdam, The Netherlands: Elsevier.

Whorf, B.L. (1956) The relation of habitual thought and behavior to language. In *Language, Thought, and Reality* (Carroll, J., ed.) MIT Press

Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209-224.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1-B11.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental science*, *8*(1), 88-101.

Chapter 1.

Learning the language of time: Children's acquisition of duration words

Katharine A. Tillman and David Barner

University of California, San Diego

Abstract

Children use time words like *minute* and *hour* early in development, but take years to acquire their precise meanings. Here we investigate whether children assign meaning to these early usages, and if so, how. To do this, we test their interpretation of seven time words: *second, minute, hour, day, week, month,* and *year.* We find that preschoolers infer the orderings of time words (e.g., hour > minute), but have little to no knowledge of the absolute durations they encode. Knowledge of absolute duration is learned much later in development – many years after children first start using time words in speech – and in many children does not emerge until they have acquired formal definitions for the words. We conclude that associating words with the perception of duration does not come naturally to children, and that early intuitive meanings of time words are instead rooted in relative orderings, which children may infer from their use in speech.

1. Introduction

Time is an ephemeral yet central dimension of human experience. The nature of time – and how it is mentally represented – has been a source of fascination for centuries, beginning with early philosophical inquiries (e.g., Augustine, 398/1992; James, 1890; Kant, 1781/2009; McTaggart, 1908), and extending to modern debates in cognitive and developmental psychology (e.g., Piaget, 1969; Whorf, 1956; Casasanto et al., 2010, Boroditsky, 2011; Bender & Beller, 2014). Many of these debates concern the role that natural language plays in mental representations of time. Linguistically, reference to time is pervasive both in speech and in written text (Brysbaert & New, 2009; Kucera & Francis, 1967). However, learning to encode time in language is difficult. Although both temporal syntax (e.g., tense marking) and time-related lexical items (e.g., *before*, *after*, *today*, *minute*) emerge early in children's language production (Ames, 1946; Brown, 1973; Dale & Fenson, 1996; DeVilliers & DeVilliers, 1978; Grant & Suddendorf, 2011; Harner, 1982; Weist & Buczomska, 1987), the meanings of time words are learned slowly in development, resulting in a long period of frequently incorrect usage and incomplete comprehension (Ames, 1946; Clark, 1971; Grant & Suddendorf, 2011; Harner, 1982; Shatz et al., 2010; Weist, Wysocka, & Lyytinen, 1991). Here, we investigate the lag between production and comprehension of duration words, such as *minute* and *hour*. We ask how children interpret these abstract words before they receive formal instruction, and thus whether they construct interim meanings for time words early in development. Characterizing the initial meanings children assign to time words may contribute both to our understanding of how time is represented in the child's mind and to our understanding of how abstract words are learned more generally.

Time words pose a difficult problem for children during language acquisition. First, although time and its passing are fundamental to experience, duration words like *second, minute,* and *hour* carve out relatively arbitrary units that cannot be directly seen or heard. Their boundaries are rarely demarcated explicitly in conversation—and, in fact, can only be precisely indicated via the use of measuring devices like clocks. Second, and relatedly, the units that define such time words are couched in a system of numerical knowledge that children take many years to master. For example, to acquire an adult-like understanding of the word *hour* (i.e., that it contains 60 minutes), children must also learn about minutes (which comprise 60 seconds), and in turn about seconds. In each case, mastery of the number words in question is difficult and protracted, and is typically not achieved until 5 or 6 years of age, if not later (for discussion of the stages of number word development, see Carey, 2004, 2009; Davidson, Eng, & Barner, 2012; Fuson, 1983, 1988; Fuson & Hall, 1983; Gelman & Gallistel, 1978; Le Corre & Carey, 2007; Schaeffer, Eggleston, & Scott, 1974; Wynn, 1990, 1992). Third, because duration words depend on numeracy, children are not explicitly taught formal definitions of these time words until very late in development, generally when they enter grade school. In standard US K-12 curricula, an introduction to clocks and time measurement begins in Grade 1, when children are 6 or 7 years of age, and instruction on time telling continues until Grade 3 (Common Core State Standards Initiative, 2010). Finally, children's experience with time words and how they are used in speech is relatively haphazard, given the variety of uses for words like *minute* and *second*. For example, a word like *minute* is frequently used informally in expressions like "just a minute" and "wait a minute," which only rarely reflect precise or accurate durations (Tare et al., 2008).

Perhaps because of these challenges, children do not acquire an adult-like understanding of duration words until quite late in development. Given this, two questions arise: When children use these words early in development, what do they mean, if anything? And, if these words have meaning for children, how are these meanings learned? Previous studies suggest that very early on, children recognize that time words are relevant to questions about time and that they form a class of lexical alternatives. For example, when asked questions regarding temporal extent, children typically respond with duration words like *minute* and *hour* despite failing to use them accurately (Shatz et al., 2010). Beyond this, however, remarkably little is known about the acquisition of time words, and what happens between the point when children begin using these words, and when they acquire their formal definitions many years later. One early study reported that at age 6, more than 50% of children had still not learned the precise meaning of the word *hour* (i.e., that an hour is 60 minutes; Ames, 1946), and no studies have documented how children acquire other duration words, like *second* and *minute*, though it is known that their ability to read this information from clocks remains imperfect until age 9 or later (e.g., Friedman & Laycock, 1989).

In the present study, we investigate children's early interpretation of duration words like *second, minute,* and *hour*, and ask whether children assign early preliminary meanings to these words before they acquire formal definitions in grade school, and if so, how these meanings are learned. Specifically, we explored two types of meanings that children might assign to duration words early in development: (1) meanings rooted in the perception of approximate duration, and (2) meanings defined by the rank ordering of time words, independent of their actual durations.

Well before acquiring the formal meanings of time words, children – like adults and many non-human animals – exhibit a robust capacity to perceive and discriminate the approximate duration of brief temporal events (e.g., Brannon et al., 2007; Droit-Volet & Wearden, 2001; Droit-Volet et al., 2004; Lewkovicz, 2003; McCormack et al., 1999; Pouthas et al., 1993; vanMarle & Wynn, 2006), much like they can represent approximate numerosity of sets (Izard et al. (2009); Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard 2005). Early work on rats and pigeons, including studies by B.F. Skinner (e.g., Ferster & Skinner, 1957), demonstrated an ability to estimate duration to guide feeding and reinforcement behaviors, governed by Weber's law (e.g., Catania, 1970; Gibbon, 1977; Stubbs, 1968). Human infants also display remarkable sensitivity to differences between brief durations: babies as young as 1.5 months of age adapt to temporal patterns in stimulus presentation and display anticipatory behaviors timed to those stimuli (e.g., Clifton, 1974; Columbo & Richman, 2002; Demany, McKenzie, & Vurpillot, 1977; Fitzgerald & Brackbill, 1968; Lewkowicz, 2003; Rivière, Darcheville & Clément, 2000; Trehub & Thorpe, 1989). At 4 to 6 months of age infants can be trained to discriminate longer vs. shorter sounds (Provasi et al., 2010; vanMarle & Wynn, 2006), a capacity which improves gradually over development. For example, performance on temporal bisection tasks steadily improves between 3 years and 10 years of age (Droit-Volet, Clement, & Wearden, 2001; Droit-Volet, Torret, & Wearden, 2004; Droit-Volet & Wearden, 2001).

Despite lacking precise meanings for time words, children's early ability to perceive and discriminate duration might allow them to associate time words with approximate duration information – what we call the *Duration Mapping* hypothesis. How might such a link between perception and language be acquired? One possibility is that children make a direct mapping between experienced duration and time words. This might occur spontaneously, as children watch events unfold over time while hearing concurrent adult speech about the durations of those events (e.g., “She’s been watching TV for an hour already!”). However, this process might be complicated by the fact that the start and endpoints of events are rarely explicitly marked, either perceptually or in speech. A second possibility is that the association between duration and time words is mediated by children’s prior knowledge of the durations of familiar events (Friedman, 1990). For instance, if a parent said to her child, “Dinner is in an hour, so we can only watch one show,” the child might constrain her hypothesis about what an hour is by associating it with the known approximate duration of watching one TV show. Thus associative mappings between duration words and temporal magnitudes could be achieved in a two-step process in which events are associated with temporal magnitudes via duration perception, and duration words are associated with events via linguistic input.

Although it is possible that children begin acquisition by associating time words with approximate durations, it is also possible that their first hypotheses about the meanings of these words do not involve nonverbal representations of duration. In addition to picking out absolute durations, time words can also be distinguished according to their rankings within a common lexical class. Hours are longer than minutes, which are longer than seconds. Children might learn this information, without having any knowledge about the precise durations denoted by such words, if they learned their initial meanings on the basis of their contrastive use in language. For example, a child who heard an utterance like, “The whole show lasts an hour, but there are only a few minutes left” might infer that an hour is longer than a minute, without having any sense of the precise duration denoted by either word.

In order to make such inferences, a child would first need to know that words like *minute* and *hour* denote temporal extent, and thus contrast along this dimension (Au & Markman, 1987; Clark, 1990). Consistent with this, children appear to know that duration words belong to a common lexical category by at least the age of 4, long before they learn their individual meanings (Tare et al., 2008; Shatz et al., 2010). For example, when asked “how long” or “how much time” an event takes, many preschoolers are able to respond using domain-appropriate expressions (i.e., using duration words paired with quantity words) despite being inaccurate in their responses (Ames, 1946; Shatz et al., 2010). Because preschoolers know which words in their lexicon refer to duration in general, this knowledge – combined with conversational cues – might allow them to infer that duration words denote contrasting durations, and that their lexical rank ordering reflects a corresponding rank ordering of relative duration – e.g., that a year is longer than a month, which is longer than a week, and so on. We call this the *Lexical Ordering* hypothesis.

Though prior work has shown that preschoolers possess incomplete comprehension of duration words, these studies cannot differentiate between the Duration Mapping and Lexical Ordering hypotheses. For example, in one recent study, children’s early comprehension of duration words was assessed by probing how well they could match familiar activities (e.g., “a boy eating breakfast”) to adult-estimated approximate durations of these activities (e.g., “ten minutes*”).* This study showed that 5-year-olds performed just above chance, demonstrating a rudimentary understanding of duration expressions and how they relate to familiar activities. Even 6- to- 7-year-old children, some of whom presumably had received instruction on time word definitions, were only 67% accurate on this binary-choice task (Shatz et al., 2010). These results clearly demonstrate that children struggle to relate conventional expressions of time to event durations. However, the findings are difficult to interpret because each prompt combined duration words, number words, and particular events (e.g., “learning to surf”). Children could succeed (or fail) at the task based on their level of understanding in any of these domains. Thus it is unclear if knowing the rank ordering of duration words would be sufficient for success, or if knowledge of their absolute durations is needed. In fact, this could vary based on a child’s knowledge of the events and number words in question. Furthermore, though prior studies have shown that children as young as 4 years old can rank-order events in terms of duration by placing markers onto a spatial representation of time (Friedman, 1990), these studies did not test conventional duration words like *minute* and *hour.* Thus it is unclear whether children can use a spatial scale to represent the rank ordering and durations of conventional time words.

In the present study, we therefore focused specifically on how children represent the rank order and absolute durations encoded by words like *second, minute,* and *hour*. In Experiment 1, we tested whether children know the temporal orderings of duration words. Children answered forced choice questions in which duration words were the only cue that could support judgments (e.g., “Farmer Brown slept for a minute. Captain Blue slept for an hour. Who slept more?”). Success on this task indicates knowledge of the rank ordering of duration words, but would also be consistent with knowing their absolute durations. Experiment 2 provided a first test of whether children are sensitive to the difference in absolute duration between terms, by testing how well children can combine duration words and number words (e.g., “Farmer Brown slept for three minutes. Captain Blue slept for two hours. Who slept more?”). While success in Experiment 2 is consistent with knowledge of approximate durations, failure could also be attributed to syntactic difficulties combining number words and duration words, or to difficulty with number words more generally. In Experiment 3, we used a number-line task to test the Duration Mapping hypothesis more directly, asking how well children can estimate the durations of familiar events and of conventional time words, on the logic that accurate estimation of duration words could not be achieved without associations between these words and approximate durations.

On the Lexical Ordering account, we predicted that children should be able to correctly compare *hour* vs. *minute* before having representations of the approximate duration represented by each word, and that mappings to duration may not arise until children have learned the precise definitions of time words. On the Duration Mapping account, which posits that each duration word is linked to an approximate magnitude, we predicted that before learning their precise definitions, children should not only know that an hour is longer than a minute, but also approximately *how much* longer it is (i.e., $hour \approx minute \times 60$).

2. Experiment 1

The goal of Experiment 1 was to test children’s knowledge of the rank ordering of time words – e.g., when they learn that an hour is longer than a minute. To do this, we tested 3- to 6-year-old children using a forced choice task in which they were asked to decide which of two verbally described events was longer – e.g., “jumping for a minute” versus “jumping for an hour.” Unlike in previous studies of time word comprehension (Shatz et al., 2010), children could not use their knowledge of the typical durations of events to guide their choices, nor could they rely on their knowledge of number words (because none were included in the prompts). To succeed at this task, children must know the rank ordering of the duration words, within their common lexical class.

2.1 Methods

2.1.1 Participants

Subjects were 92 children recruited from the San Diego area, including 25 3-year-olds (mean age = 3;6), 26 4-year-olds (mean age 4;6), 20 5-year-olds (mean age 5;5), and 21 6-year-olds (mean age 6;5). An additional 4 children participated but were excluded from analysis due to failure to complete the task. Children were either recruited by phone and tested in our UCSD laboratory, or tested in local daycares, preschools, and museums. Parents gave informed consent for their children to participate in the study, and children indicated their willingness to play a game with the experimenter before testing began. Parents who brought their children into the laboratory were compensated for their travel expenses, and children received a small gift in thanks for their participation.

2.1.2 Procedure

The child was first introduced to two action figures, Farmer Brown and Captain Blue, who were placed on a table in front of the child. On each trial, the experimenter read a short vignette of the form, “Farmer Brown [jumped] for [a minute]. Captain Blue [jumped] for [an hour].” This was followed by a two-alternative forced choice, “Who [jumped] more, [Farmer Brown or Captain Blue]?” If the child was reluctant to give a verbal response, they were encouraged to point to the character who did the action more. The experimenter then proceeded to the next trial.

Children completed two blocks of trials, each containing thirteen duration word comparisons. Seven time words were tested: *second*, *minute*, *hour*, *day*, *week*, *month*, and *year*. The specific comparisons tested were: week vs. month, day vs. week, month vs. year, hour vs. day, day vs. month, week vs. year, minute vs. hour, second vs. minute, hour vs. week, day vs. year, minute vs. day, second vs. hour, and second vs. day. Each comparison appeared once in each block, and each duration word described an activity that could in principle be done for any length of time. These activities were described using the past tense form of six high-frequency action verbs: *jump*, *sleep*, *cry*, *play*, *dance*, and *talk*. Critically, on a given trial, the same verb was paired with both duration words. Within each block, trials were presented in quasi-random order. Verbs were pseudo-randomly assigned to duration comparisons, with the stipulation that the same verb was never used in two consecutive trials. Trials were counterbalanced with respect to whether the larger duration word came first, which character performed the longer action, and which character was mentioned first in the question prompt. Half the participants received one item-order, and the other half received the reverse order. For analysis, the child's response on each trial was coded as correct (1) or incorrect (0). These numbers were then converted into proportions correct for group-wise comparisons.

2.1.3 Analyses

Linear mixed-effects analyses were performed in R (R Core Team, 2013) using the *lme4* (Bates et al., 2013) extension package. Unless otherwise noted, p-values were obtained by likelihood ratio tests of the full model in question against the model without the effect in question.

2.2 Results and Discussion

Our first analysis examined whether children knew the rank ordering of duration words. To assess children's overall understanding of the rank ordering of duration words (e.g., their knowledge of which words indicate longer times than others), we first calculated each child's proportion of correct responses, across all trials and duration word comparisons. An ANOVA revealed a significant effect of age group on overall accuracy ($F(3,84) = 23.1$, $p<0.001$). The youngest group of participants, the 3-year-olds, did not perform better than 50% accuracy, consistent with random guessing, M (SEM) = 0.48 (0.02), $t(24)= 1.2$, n.s. The 4-, 5-, and 6-year-old groups all performed significantly better than chance (t's > 2; p's < 0.05) and significantly better than each younger group: 4-year-olds, M (SEM) = 0.57 (0.02); 5-year-olds, M (SEM) = 0.67 (0.04); 6-year-olds, M (SEM) = 0.82 (0.03); t's > 2; p's < 0.05. Thus, beginning at 4 years of age, children exhibited partial knowledge of the rank ordering of duration words. However, somewhat surprisingly, even our oldest age group (6-year-olds) frequently failed to judge, for example, that a minute was longer than a second, M (SEM) = 0.68 (0.09); that a day was longer than an hour, M (SEM) = 0.65 (0.09); or that a year was longer than a month, M (SEM) = 0.60 (.08).

While this result shows that children have some knowledge of the rank ordering of duration words, it is also consistent with knowledge of their absolute durations. If children have knowledge of the absolute durations of time words (even if rudimentary), we might expect them to perform better on comparisons in which their magnitudes differ more in scale. To explore this, we next asked whether children's performance on a given trial was affected by the difference in absolute duration represented by the two terms being contrasted – e.g., if *hour vs. second* was easier than *minute vs. second*. To test this possibility, we conducted a mixed effects logistic regression testing whether the ratio between the terms in each comparison (e.g., *minute vs. second* = 60, *hour vs. second* = 360) predicted the accuracy (correct vs. incorrect) of children's responses. We entered age and duration-ratio as fixed effects in the model, and as random effects we included intercepts for subjects, as well as by-subject random slopes for the effect of duration-ratio. We observed an effect of duration-ratio that came in the form of an interaction with age, $\chi^2(2)=3.1$, $p=.04$. To investigate this interaction, we analyzed the data from each age group separately. We found no effect of duration ratio on accuracy in the 3- or 5-year-olds groups, χ^2's < 0.7, all *p*'s > 0.4, a *negative* effect in the 4-year-olds, ß = -0.0001, z = -2.9; $\chi^2=7.6$, $p < 0.01$, and a small positive effect of duration ratio in the oldest age group, the 6-year-olds, ß = 0.09, z = 2.6; $\chi^2(1)=5.8$, $p < 0.01$. In both cases, the model estimates (ß) for this factor were tiny. Thus, while increasing difference in duration between two time words did not predict better performance among children under age 6 – i.e., they were not better at comparing *second vs. day* than they were *second vs. minute* – our results suggest that knowledge of the absolute durations of these words may begin to emerge at age 6.

If children know the rank of each term in their list of known duration words, but not their approximate magnitudes, their performance should be better predicted by the difference in ranking between the two terms. For example, 'minute' and 'second' differ by a rank of 1, while 'day' and 'second' differ by a rank of 3. In an analysis similar to the one described above, we entered this rank difference factor into our model in place of the duration-ratio. Here, we again found a significant interaction of this factor with age $\chi^2(2)=11.2$, $p<0.001$. In follow-up analyses, we found that while this factor also did not predict performance in the 3- and 4-year-old groups (all $\chi^2<2.4$, all $p>0.05$), it had a marginally-significant positive effect on performance in the 5-year-old group (ß = 0.34, $z = 2.0$; $\chi^2=3.4$, $p=0.07$) and a highly-significant effect in the 6-year-old group (ß = 0.89, $z = 3.6$; $\chi^2=10.8$, $p=0.001$). This pattern of effects suggests that by age 5 children's knowledge of the rank ordering of duration words guides their ability to contrast them. However, it should be noted that none of the comparisons we tested differed more than a rank of 3, and further experiments with a wider range of rank differences would be needed to best assess the role of this factor.

Our final analysis examined whether individual words exhibited different learning trajectories. It is possible that, although children begin acquisition by ranking time words according to their relative magnitudes, they nevertheless have some absolute magnitude knowledge for some words, but not for others. To test for differences between duration words (item effects), we collapsed the data across all comparisons involving each word, yielding an accuracy score for each child for each word. Means and standard errors for each age group are depicted in Figure 1.1. A mixed-effects ANOVA with age group as the between subjects factor and duration word as the within subjects factor revealed a main effect of age ($F(3,547) = 24.3$, $p < 0.001$) and a marginal effect of duration word ($F(6,547) = 2.1$, $p = 0.05$), with no significant interaction. However, when we analyzed the data from each age group separately, an effect of duration word was found only in the 3-year-old group, who, despite showing chance performance overall, nevertheless performed better on comparisons involving the word *week* (see Figure 1). Given the overall failure of this group, this finding is difficult to interpret and is perhaps not a reliable effect. More importantly, there was no effect of duration word on accuracy within the 4-, 5-, or 6-year-olds groups (all *F*'s < 1.6, all *p*'s > 0.05). This finding suggests that as accuracy improves over these years, it may improve across the board, with equal improvement on each tested word (Figure 1.1). However, because each trial in this experiment tested children's knowledge of two duration words, differences between individual words may have been masked. We will return to this issue in Experiment 3.

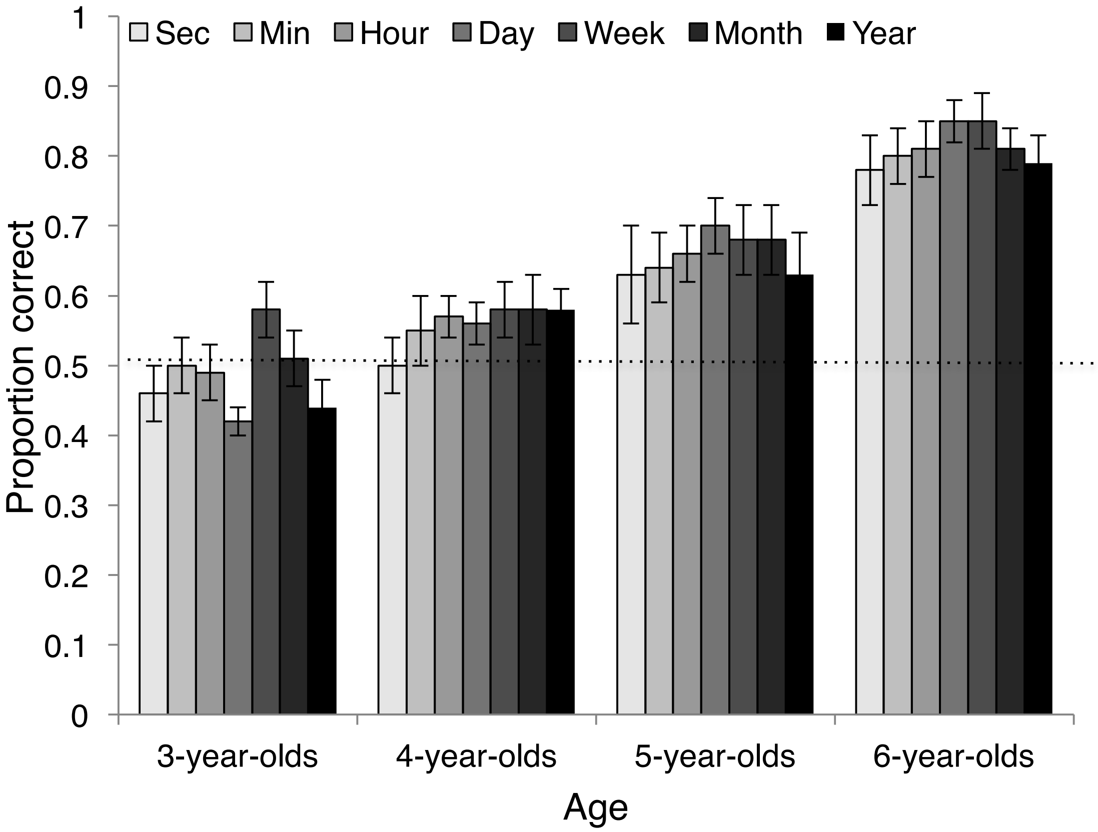

Figure 1.1. Proportion of correct responses on all trials in Experiment 1 containing the indicated time word, for each age group. Error bars represent SEM. Dashed line indicates chance performance.

3. Experiment 2

In Experiment 1, we found that by around age 4 many children know the rank order of some duration words, but that this learning process is gradual, and not yet complete in 6-year-olds. Also, surprisingly, we found no systematic differences in children's understanding of individual time words, and no evidence that children under 6 were more accurate when the time words that they compared differed more in duration.

These preliminary data suggest that children acquire the rank ordering of time words between the ages of 4 and 7, but that they may not be sensitive to their absolute durations until quite late in development. In Experiment 2 we explored this possibility further, by introducing number words into judgments like those used in Experiment 1. In addition to simply comparing *hour* to *minute*, the task required children to compare, e.g., *3 minutes* to *2 hours*. We reasoned that if children have an approximate understanding of the duration of an hour, then they should know that 2 hours is much longer than 3 minutes, so long as they also understand the meanings of 2 and 3. Although children in the US generally comprehend such numbers and can accurately compare them by the age of 3 or 4, we nevertheless verified that children were able to compare cases in which the larger number was paired with the longer duration word – e.g., *3 hours* vs. *2 minutes*.

3.1 Methods

3.1.1 Participants

We tested 93 children, including 25 4-year-olds, 22 5-year-olds, 25 6-year-olds, and 21 7-year-olds. Children were recruited from the same population as those in Experiment 1. An additional 8 children participated but were excluded from analysis due to failure to complete the study (7) and experimenter error (1). None had previously participated in Experiment 1.

3.1.2 Procedure

The procedure for Experiment 2 was identical to that of Experiment 1, except that the time words were modified by number words. For example, “Farmer Brown [jumped] for [two] [minutes]. Captain Blue [jumped] for [three] [hours]. Who jumped more?” Each child completed a total of 30 trials.

Table 1.1. Trial Types used in Experiment 2

Number comparison	Number size	Example
No numbers	None	A minute vs. an hour
Same	Small	2 minutes vs. 2 hours
	Large	6 minutes vs. 6 hours
Congruent	Small	2 minutes vs. 3 hours
	Large	6 minutes vs. 9 hours
Incongruent	Small	3 minutes vs. 2 hours
	Large	9 minutes vs. 6 hours

Trials in Experiment 2 included the same six verbs from Experiment 1. However, only five time-word comparisons were used in Experiment 2: minute vs. hour, week vs. year, day vs. year, day vs. week, and second vs. hour. Findings in Experiment 1 indicated that children’s performance on these comparisons was similar. For each time-word pair, 7 different types of number-word comparisons were made (Table 1.1). For each pair, one trial included no numbers (identical to Experiment 1), 3 included “small” numbers (2 and/or 3), and 3 three included “large” numbers (6 and/or 9). Each comparison was designated Same, Congruent, or Incongruent as defined in Table 1. Trials were presented in quasi-random order. Half the participants received one item-order while the other half received the reverse order.

3.1.3 Analyses

Linear and logistic mixed-model analyses were conducted as in Experiment 1.

3.2 Results and Discussion

To determine which experimental variables influenced children's performance in Experiment 2, we first conducted a mixed-effects logistic regression to predict the accuracy of each child's response (correct vs. incorrect) using the participant's age, the duration word comparison type (e.g., *hour/s vs. minute/s*), the number size (none, small, or large), and the number comparison type (same, congruent, or incongruent) as fixed effects. As random effects, we had intercepts for subject, as well as by-subject random slopes for word comparison, number size, and number comparison type. The only significant main effects were those of age group, $\chi^2(3) = 67.9$, $p < .001$, and number comparison type, $\chi^2(2) = 16.5$, $p < 0.001$. Because there were no significant effects of word comparison or number size, χ^2's < 5.6, p's > 0.2, the data were collapsed across these factors in all subsequent analyses.

A linear mixed effects analysis of the collapsed dataset (with age group and number comparison type as fixed effects, random intercepts for subject, and by-subject random slopes for the effect of comparison type) revealed that the effect of comparison type on accuracy was driven by an interaction with age, $\chi^2(6)=29.9$, $p < 0.001$. Mean accuracy for each age group on the Same (including no-number), Congruent, and Incongruent comparisons is depicted in Figure 1.2. As shown, the 4-, 5-, and 6-year-old groups all performed significantly worse on the critical Incongruent trials than on the Congruent trials (all t's > 2.9, all p's < 0.005), and the 4- and 5-year-old groups also performed significantly worse on Incongruent trials than on Same number trials (t's $>$

2.2, *p*'s < 0.05). 6-year-olds did not perform significantly worse on Incongruent trials than Same trials, though performance still remained far from adult-like at 74% correct. In contrast, the 7-year-old children's performance was consistently near ceiling and unaffected by comparison type, *F*(2,20) = 0.25, $p = 0.8$, n.s., indicating a much more robust understanding of the durations being contrasted.

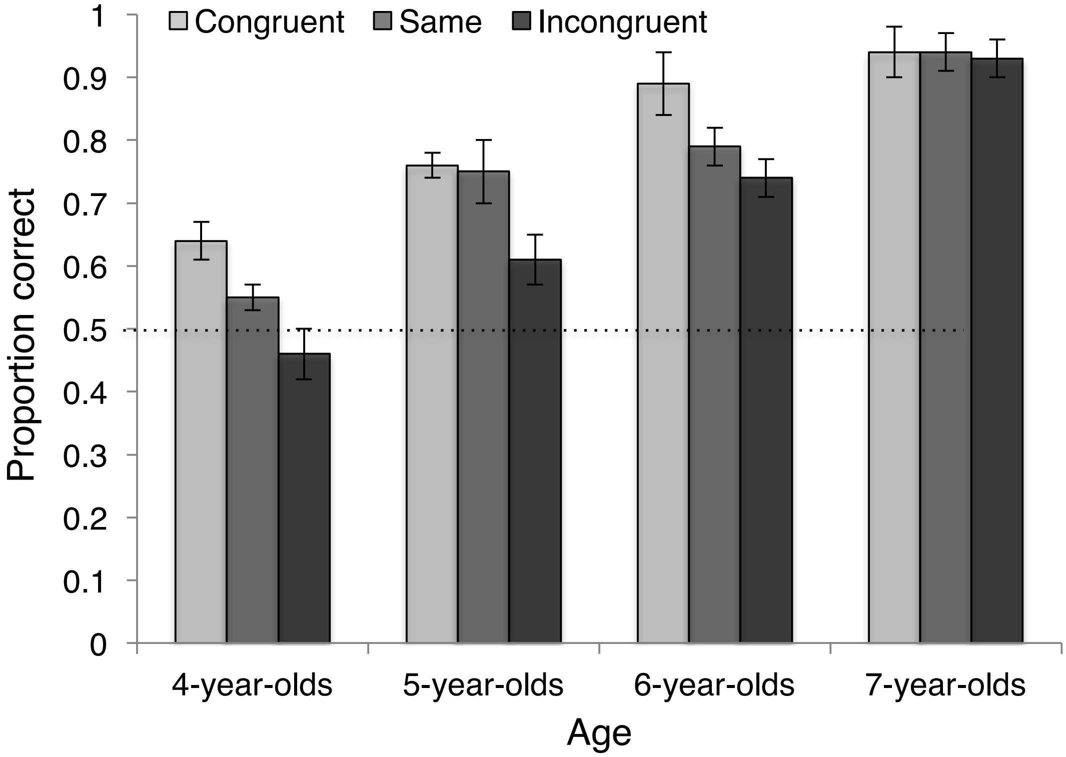

Figure 1.2. Effect of congruency between time word comparisons and number word comparisons in Experiment 2.

To understand these data it is useful to consider a concrete example. The results of Experiment 2 show that many 5-year-olds knew both that 3 is greater than 2 (76% correct overall on Congruent trials) and that an hour is longer than a minute (75% correct overall on Same trials), but performed significantly worse when judging that 2 hours is more than 3 minutes (61% correct overall on Incongruent trials). Given that the ratio between an hour and a minute greatly exceeds 3:2, this failure indicates that children often cannot combine their knowledge about duration words with their knowledge about number words, suggesting that children's early meanings for duration words may not include information about approximate duration.

4. Experiment 3

The results of Experiments 1 and 2 suggest that while some children learn the rank ordering of duration words by age 4, most children do not associate these words with even approximate durations until as late as 6 or even 7 years of age. Many 6-year-olds in Experiment 2, for example, did not know that 2 hours was longer than 3 minutes.

A limitation of the forced-choice method used in the first two experiments is that each trial probed the participant's knowledge of two different duration expressions. A problem with this approach is that data for individual words are non-independent, making it more difficult to determine whether the words have absolute meanings for children that do not depend on comparison to one another. To further probe children's knowledge of the absolute durations encoded by individual time words, Experiment 3 tested children using a task that required them to place individual duration expressions onto a time line representing increasing duration. We asked children and adults to estimate the durations of familiar events (e.g., "watching a movie"), duration words (e.g., *hour*), and duration words paired with numbers (e.g., *3 hours*), using this spatial representation of time. This allowed us to disentangle children's ability to reason about relative duration (e.g., of events) from their knowledge of linguistic labels for durations (i.e., words like *minute*). Also, because an adult-like understanding of time words requires reasoning about large precise numerosities (e.g., quantities of 60), we tested each child with a standard number line task, in which they were required to place numbers on a line representing values from 0 to 100.

Finally, we also tested how children's understanding of duration words is related to their knowledge of formal definitions. We hypothesized that children may first represent the approximate durations of time words when they are formally taught the meanings of each word – e.g., that an hour equals 60 minutes, and a minute equals 60 seconds. We tested this hypothesis by asking children follow-up questions regarding the definitions of *second, minute, hour*, and *day* – e.g., "How many seconds are in a minute?"

4.1 Method

4.1.1 Participants

We tested 64 children, including 22 5-year-olds (*M*=5;6), 21 6-year-olds (*M*= 6;6), and 21 7-year-olds (*M*=7;6). We also tested 36 adults (M=20;7). One additional child participated but was excluded from analysis due to failure to complete the task. Children were from the same population as those in Experiments 1 and 2. Twelve children had previously participated in either Experiment 1 or 2, on another day. Adults were members of the UCSD Psychology Department subject pool and received course credit for participation.

4.1.2 Procedure

Participants were given a sheet of 8.5"x11" paper with four horizontal, 17-cm lines printed in a vertical column down the center of the page. Each line had small, filled dots on both endpoints and no other markings (i.e., the midpoint of the line was not marked). The participant's task was to estimate the magnitude or duration of various stimuli by drawing vertical marks bisecting each line using colored pencils.

The first line was a standard number line. The experimenter provided the following instructions to the participant: "This is a number line. Each number has its own place on the line. You're going to show me where certain numbers go on the number line. Look, 0 goes here [to demonstrate, the experimenter drew a vertical line bisecting the left endpoint] and 100 goes here [the experimenter marked the right endpoint]." Then, for each of four number stimuli (see Table 2), the experimenter said, "The [first] number is [4]. Can you show me where [4] goes? Can you draw a line with the [blue] pencil?" A different colored pencil was used for each target, so the data could be easily interpreted later.

Next, children were presented three time lines, in order, which tested (1) event durations (e.g., "eating lunch"), (2) time word durations (e.g., "an hour"), and (3) numerically modified time words (e.g., "9 minutes"). Target items for Experiment 3 are shown in Table 2. Before the first duration estimation task, the line was explained to the participants as follows: "Now, this line is different. It shows how much time things take to do. It goes from a very short amount of time to a very long amount of time. Each amount of time has its own place on the line, and the further you go over here [experimenter gestured along the line from left to right], the more time something takes. You're going to show me how long certain things take to do on the line. Something very short, like blinking your eyes, goes here [experimenter marked left endpoint]. Something very long, like the time from waking up in the morning to going to bed at night, goes here [experimenter marked right endpoint]." For each item (see Table 2), the child was instructed to think about how long the activity takes to do and to mark the line accordingly. Participants were reminded that each subsequent line represented duration and what the endpoints represented ("short, like blinking your eyes" on the left and "long, like the time from waking up in the morning to going to bed at night" on the right) in between each of the remaining tasks and any time they indicated confusion.

Table 1.2. Number-line and Timeline Stimuli in Experiment 3

Line 1: Numbers	Line 2: Events	Line 3: Time words	Line 4: Number + time
4	Watching movie	Hour	2 hours
45	Washing hands	Second	6 hours
18	Trip to zoo	Minute	9 min
61	Eating lunch	Day	3 min

Because each child was tested with four lines – a substantial battery especially for the youngest children – questions were limited to four items per line. Each child completed the series of lines in the same order: (1) numbers, (2) events, (3) time words, and (4) time words with numbers. Within each line, half the participants received the four stimuli in the order shown in Table 2, while the other half received the reverse order. As in Experiments 1 and 2, participants were presented with time word stimuli (lines 3 and 4) in the context of events that could take variable amounts of time, e.g., "[jumping] for a minute." The same action verbs were used as in Experiments 1 and 2.

Following completion of the four number line and timeline tasks, each participant was asked 3 follow-up questions: "How many minutes are in an hour?", "How many hours are in a day?", and "How many seconds are in a minute?"

4.1.3 Coding and analyses.

To analyze the time line and number line data, we first measured the distance (measured as centimeters from the 0 point) from the left endpoint to the intersection of the number line and each participant's pencil marks. Marks falling exactly on the left endpoint were recorded as 0.1 cm (to avoid divide-by-zero errors) and those falling exactly on the right endpoint were recorded as 17.0 cm.

To assess children's knowledge of the *rank ordering* of events and time words, responses to the four stimuli for each line were rank-ordered by increasing magnitude or duration. For duration words, the order was: 1 = second; 2 = minute; 3 = hour; 4 = day. For events, the adults' modal order (N=36) was: 1 = washing hands (N=36); 2 = eating lunch (N=35); 3 = watching a movie (N=34); 4 = going on a trip to the zoo (N=35). For each estimated item which fell in the correct rank, the participant was awarded a 1, for each incorrectly ranked item, the participant was given a 0, resulting in a score out of 4 for each line.

To assess number estimation performance, participants' estimates (measured as centimeters from the 0 point) were converted to their corresponding values from 0 to 100 and plotted as a function of magnitude being estimated. The endpoints on the number line were 0 and 100, and the highest number children were asked to estimate was 61.

To assess children's knowledge of the relative durations of events and time words, we computed ratios between the estimates (termed "estimate ratios") for each possible pair of stimuli (e.g., min/sec, hour/sec, hour/min, day/sec, day/min, day/hour), and plotted the children's estimate ratios as a function of the mean estimate ratios provided by our adult sample.

As in Experiments 1 and 2, linear mixed-effects analyses were performed in R (R Core Team, 2013) using the *lme4* (Bates et al., 2013) extension.

4.2 Results and Discussion

4.2.1 Rank ordering analysis

To corroborate the findings of Experiments 1 and 2, we first assessed children's knowledge of the rank ordering of numbers, events, and time words (e.g., second < minute < hour < day). For each task (numbers, events, time words, and time words with numbers), we calculated the proportion of items each child placed in the correct rank relative to the others, yielding an accuracy score for that line. Mean proportions of correctly ordered estimates for each age group on each task are shown in Figure 3. We conducted a linear mixed effects analysis of children's accuracy in which age and task were entered as fixed effects, and random effects included random intercepts for subjects and random by-subject slopes for the effect of task. This model revealed significant effects of age group and task, driven by a significant interaction between them, $\chi^2(9)=37$, $p < 0.001$. All age groups performed best when ordering number words, indicating that children understand the number line paradigm and that their difficulty ordering duration expressions cannot be attributed to difficulty with numbers. Overall, children also demonstrated substantial understanding of the rank ordering of time words, despite the increased difficulty of this task relative to Experiments 1 and 2. Strangely, though the youngest group (5-year-olds) did not perform significantly above chance (0.25) when ranking unmodified duration words, $t(21) = 2.8$, n.s., this group performed above chance when time words were modified by (incongruent) number words, $t(21) = 2.8$, $p = .01$. However, because the stimuli were designed so that the number words gave incorrect cues to which expression denoted the longer duration (see Table 2.1; e.g., “9 seconds,” despite having the largest number word, is the second-shortest duration), we cannot attribute this result solely to number word understanding. By age 7, children rank ordered duration words both with and without number words as well as adults (*t*’s < 1.5, *p*’s > .1).

Surprisingly, we found that children in all three age groups performed relatively poorly when ranking familiar events (described without time words) by increasing duration, and that even the oldest children performed much more poorly than adults, $t(25)=-4.4$, $p<0.001$. Given their poorer ability to rank familiar events by duration, it seems highly unlikely that children’s learning of the rank ordering of duration terms is mediated by knowledge of the approximate durations of events (e.g., that children learn “an hour” by mapping it to events described as “an hour”, and noting that duration of those events).

4.2.2 Number estimation

The previous analysis tested children’s ability to order expressions, without regard to whether these estimates were accurate. We next analyzed children’s ability to accurately place numbers on the number line. Following previous number line estimation studies (e.g., Barth, Starr, & Sullivan, 2009; Booth & Siegler, 2006; Lipton & Spelke, 2005; Siegler & Opfer, 2003; Sullivan & Barner, 2014), we assessed performance by plotting subjects’ estimates as a function of the numbers being estimated. As the slopes of the best-fitting linear regression through the data for each age group (with subject as a random factor) approach 1, estimation becomes more accurate. As indicated by the steep slopes of their estimation functions (Table 1.3), we found that by age 7, children estimated numbers in this range no differently than adults, *t* = 1.8, n.s. Six-year-olds' performance was somewhat lower than that of adults, *t*=4.1, *p* < .001, but still quite high (ß=0.9). Five-year-olds performed more poorly, but also produced estimates that were correlated with the magnitudes they were asked to estimate. These data are similar to those found in other developmental studies of number-line estimation for magnitudes up to ~50 (e.g., Barth & Paladino, 2011; Barth, Starr, & Sullivan, 2009; Slusser et al., 2013; Sullivan & Barner, 2014). Critical to the present study, the results suggest that failure to

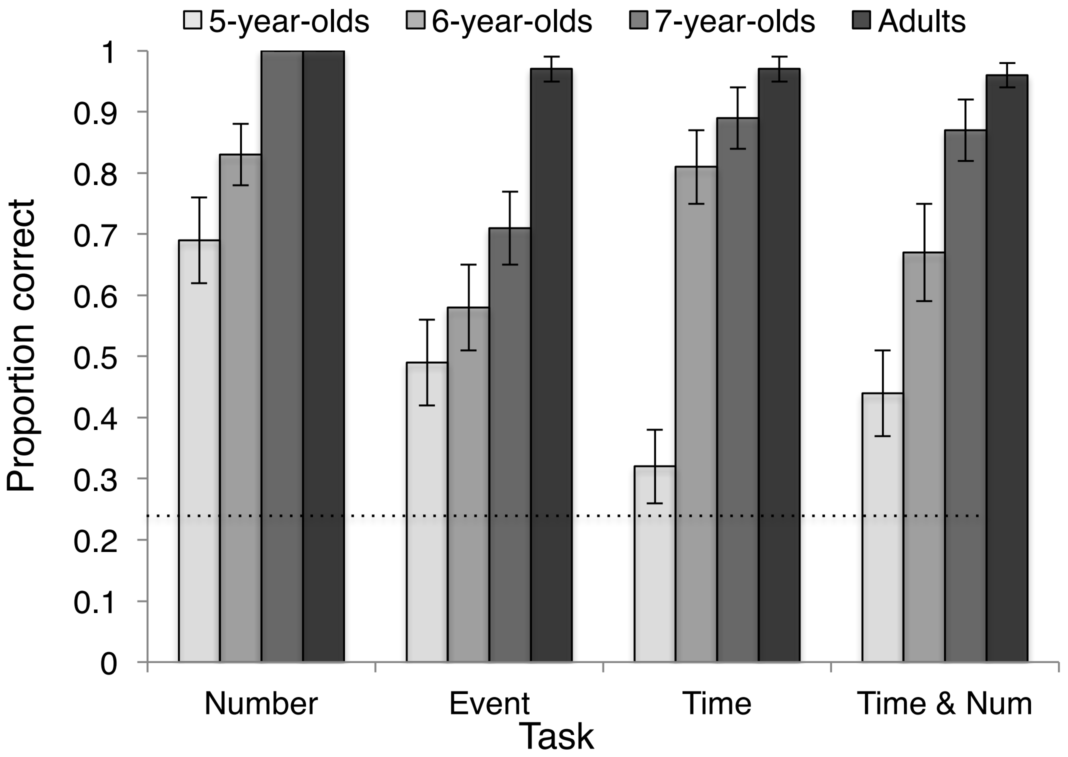

Figure 1.3. Proportion of correctly ranked estimates for each age group in Experiment 3. Tasks: *Number* = number line estimation, *Event* = familiar event duration estimation (no time words), *Time* = duration word estimation (e.g., minute), and *Time & Num* = duration words modified by number words. Error bars indicate SEM. Dashed line indicates chance performance, one item falling in the correct rank.

accurately estimate duration cannot be attributed to an inability to place quantities on a line or to associate number words with approximate magnitudes.

Table 1.3. Regression Analysis of Number-line Estimation

Task	Age group	ß	SEM	R^2
Number-line	5 YO	0.61	0.11	0.22
	6 YO	0.90	0.05	0.60
	7 YO	0.99	0.02	0.83
	Adult	1.0	0.03	0.90

Note: A mixed ANOVA on children's number estimates found main effects of the magnitude being estimated, age group ($p < 0.05$), and a significant interaction between the two (all F's > 7.6, all p's < 0.05)

4.2.3 Duration estimation

Having established that all groups were relatively competent at placing number words accurately on a number line, we next evaluated children's ability to accurately estimate the durations of events, time words, and time words with numbers on timelines. The mean estimates (distance from the left endpoint of the line) given by each age group for each temporal item are shown in Figure 1.4. Children of every age generally overestimated duration substantially. Within each of the three duration estimation tasks, linear mixed-effects analyses of participants' raw duration estimates (with age group and item as fixed effects, random intercepts for subjects, and random by-subject slopes for the effect of item) revealed highly significant main effects of age and item (e.g., 'hour'), as well as significant interactions between them, all χ^2's < 25; all p's < 0.001. Here, the effect of item indicates that participants distinguished the terms from one another on the

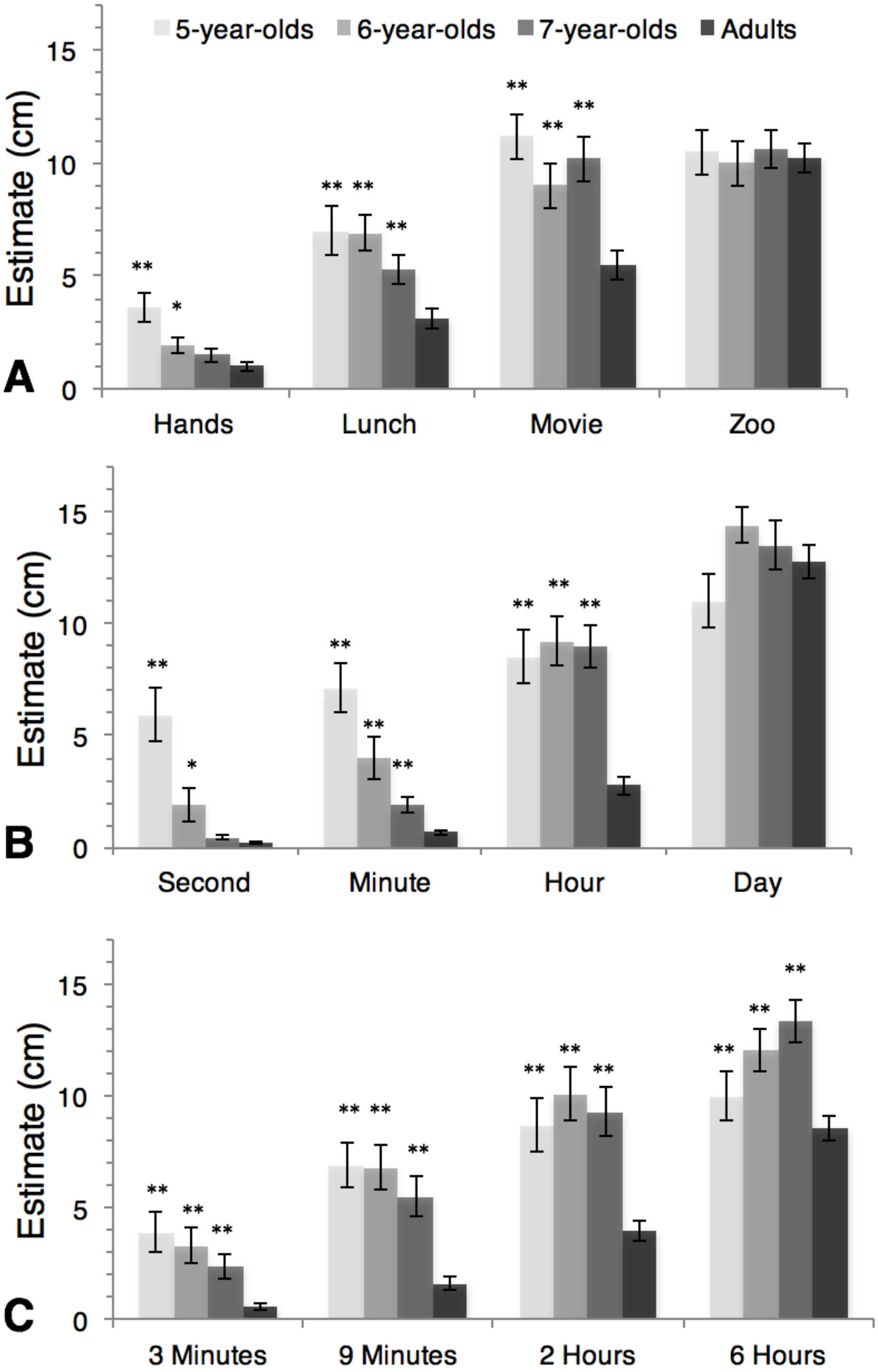

Figure 1.4. Mean duration estimates for events (A), time words (B), and time words with numbers (C). Estimates (cm) were defined as the distance from the left endpoint of the line. Asterisks indicate the significance level of the difference between children's estimates and adults', * $p<0.05$, ** $p<0.005$.

line; if they were simply guessing randomly, we would expect to see no effect. Although there were a few cases in which children's estimates did not differ significantly from adults' (see Fig. 4), overall children were poor at representing absolute durations on timelines.

To compare children's knowledge of durations of familiar events and those denoted by duration words, we divided each duration estimate given by each child by the mean estimate given by the adult sample on that item. A linear mixed-effects analysis of these normalized estimates (with age group and task as fixed effects, and random intercepts for subjects and random by-subject slopes for the effect of task) revealed effects of age group and task as well as a significant interaction, all χ^2's > 5.8, all p's < .001, and post-hoc tests showed that children's estimates of event duration were significantly more adult-like than those of duration words, both with and without numbers, t's > 4, p's < 0.001.

4.2.4 Relative duration analysis

In the previous analysis we found that children differed substantially from adults when estimating durations on timelines, and that they were poorer at estimating the durations indicated by time words like *minute* than those of familiar events. However, it is possible that some children, despite failing to represent absolute duration like adults, might nevertheless have knowledge of the approximate *relative* durations that they encode. For example, despite placing an hour at a different location on the time line than adults, they might also place a minute at a correspondingly different location, and thus show evidence of knowing that an hour is roughly 60 times longer than a minute.

To test this possibility, we calculated ratios between the durations estimated for each pair of stimuli (e.g., minute/second, hour/minute), and used these "estimate-ratios" as dependent measures in our analysis. To assess children's knowledge of the relative durations of events and time words, we plotted each group of children's estimate-ratios as a function of the adults' mean estimate-ratios (Table 4)[2]. Similar to the number line analysis described above, greater steepness in the slope of that linear model indicates more adult-like performance, in this case reflecting that children spaced their estimates on the line more similarly to adults.

As indicated by their slopes in Table 4, children demonstrated strong understanding of the relative durations of familiar events and much poorer knowledge of duration words. Despite the fact that children's raw estimates of event durations differed from adults', by age 5 children appear to understand how these durations relate to one another (e.g., that it takes roughly three times as long to visit the zoo as to watch a movie). This result suggests that preschoolers' representations of familiar events include mappings to approximate durations. In contrast, young children demonstrated poorer understanding of the relative durations of time words, though their performance increased significantly with age. In particular, a pronounced improvement was observed in the 6-year-olds relative to the 5-year-olds, suggesting that children may begin to map duration words onto approximate durations around this age. However, when numbers modified duration words, all children performed much more poorly (Table 4), and there was only

---

[2] Note that for time words it is possible to use the true ratios between items, e.g., 'day/'second' = 86,400, as predictors of performance. We chose not to use this measure since for some items, like 'second' and 'minute', a difference could not be reliably represented on the line. Nevertheless, using this metric of performance generates the same pattern of results.

marginal improvement with age. The shallow slopes for both 5- and 6-year-olds, followed by only a modest gain at 7, indicate that a more sophisticated understanding of the proportional relationships between these terms – like that required to read and interpret a clock – emerges quite late. This finding has been borne out in studies on children's ability to tell time, which remains imperfect until late in the elementary school years (e.g., Friedman & Laycock, 1989).

Table 1.4. Regression Analysis of Relative Duration Estimates (Exp. 3)

Task	Age group	ß	*SEM*	R^2
Event	5	0.86	0.15	0.19
	6	0.85	0.13	0.25
	7	0.88	0.1	0.4
	Adult	1.0	0.05	0.64
Time word	5	0.15	0.03	0.15
	6	0.68	0.09	0.31
	7	0.87	0.09	0.44
	Adult	1.0	0.04	0.73
Time word w/ number	5	0.25	0.05	0.16
	6	0.19	0.02	0.49
	7	0.46	0.04	0.5
	Adult	1.0	0.07	0.51

Notes. Estimate-ratios were regressed against the mean estimate-ratios produced by adults. The y-intercept was fixed at 0 for all regressions. Analyses of variance revealed significant effects of item in all three tasks, a significant effect of age in the time word task only, and interactions between these factors in the time word and time word with numbers task (all *F*'s > 2.5, all *p*'s <0.01).

Here, children demonstrated greater knowledge of both the absolute and relative durations of events than of time words. Note that this effect is opposite our findings in the rank ordering analysis discussed above, despite the fact that both analyses were performed on the same raw data. The rank ordering analysis indicated that children are better able to rank duration words than to rank events by increasing duration (Figure 3). These contrasting findings suggest that, although children were more likely to place duration word estimates in the correct sequence, those placements were quite poorly "spaced out" along the line. For example, a child who "bunched up" all four item placements near one end of the line, with no order errors, might demonstrate poorer knowledge of relative duration than a child who made a reversal in order but had more appropriate item spacing overall.

4.2.5 Definition knowledge

By the time children begin to show rudimentary knowledge of the absolute durations of time words (age 6 or 7), most have entered school. In Grade 1, most elementary school curricula include instruction on the meanings of time words (Common Core State Standards Initiative, 2010). Thus our final question was what role children's knowledge of these definitions might play in their ability to estimate the durations of time words and to order them. We addressed this by asking children to answer three final follow-up questions: "How many seconds are in a minute?", "How many minutes are in an hour?", and "How many hours are in a day?". The proportion of children in each age group who correctly answered each question is shown in Figure 5. We observed drastic age differences in children's knowledge of duration word definitions (Figure 5). While 75% of 5-year-olds knew no definitions and none knew more than one, 90% of 7-year-olds knew two or three definitions. The 6-year-old sample was more varied – with 55%

knowing 0 or 1 definition, and 45% knowing 2 or 3.

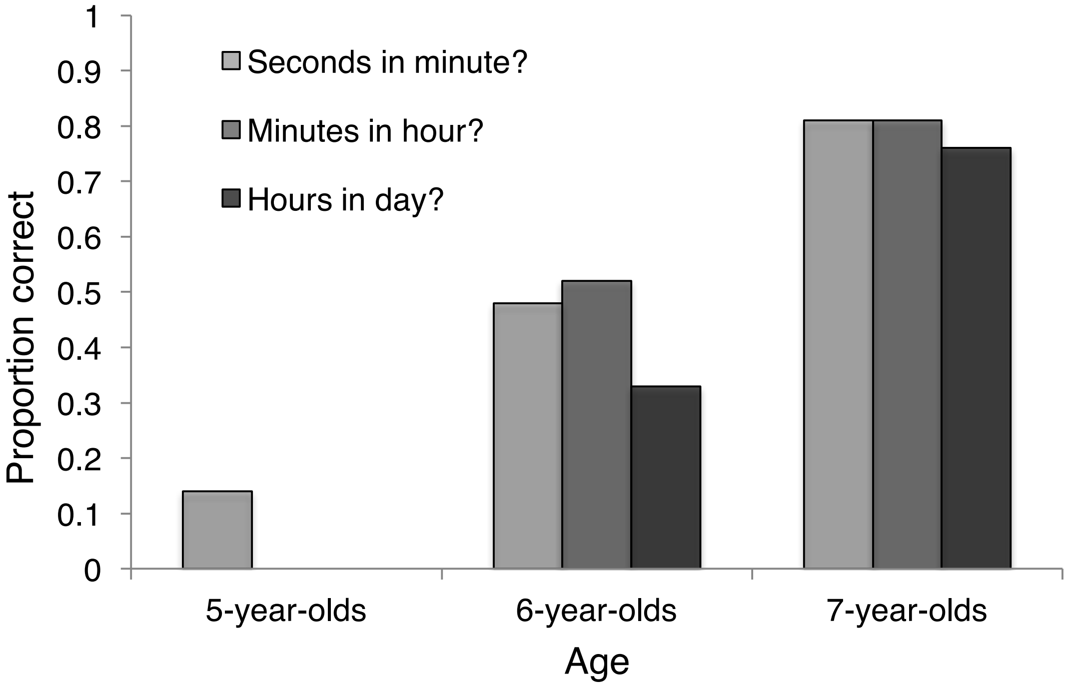

Figure 1.5. Proportion of children in each age group who correctly answered three follow-up questions about the definitions of time words: "How many seconds are in a minute?," "How many minutes are in an hour?," and "How many hours are in a day?"

Next, we tested the possibility that definition knowledge supported children's ability to estimate relative durations. In a linear mixed effects analysis of children's duration estimate-ratios, we entered definition knowledge (0-3 definitions) as a fixed effect, along with the adult mean estimate-ratio, age group, and task (events, time words, time words with numbers). We included random intercepts for subjects as well as by-subject random slopes for the effect of definition knowledge and the effect of task. This model revealed a main effect of the adult estimate and a significant interaction between definition knowledge and task, χ^2 (2)=12.9, p <0.001.

In follow-up analyses of the effect of definition knowledge specifically in the duration word estimation task, we found that when this factor was added to the model predicting estimate-ratios, there was a significant effect of definition knowledge $\chi^2(1)=12.5$, $p<0.001$. When age and definition knowledge were in the model, the effect of age was non-significant, χ^2 (2)=3.5, p=0.18, in contrast to models that did not include definition knowledge, $\chi^2(2)=12.6$, $p < 0.005$). However, even when age was included, the effect of definition knowledge remained marginally significant, χ^2 (1)=3.5, $p=0.06$. Thus, knowledge of definitions completely explained differences between age groups on this task. This finding, which was not observed for events or for time words with numbers, suggests that children's improvement on the duration word estimation task may be related to increased knowledge of time word definitions between the ages of 5 and 7.

Lastly, we assessed the relation of definition knowledge to children's ability to correctly *rank* time words and event durations. For each task, we performed a linear mixed-effects analysis of children's proportion of correctly ranked items, including age group and definition knowledge as fixed effects. As random effects, we included intercepts for subjects, and by-subject random slopes for the effect of definition knowledge. This model revealed main effects of age, but no effects of definition knowledge on the accuracy of either time word or event ranking, χ^2's < 2.3, p's > 0.1. Thus, knowledge of the formal definitions of time words is not related to children's ability to rank them, consistent with our earlier finding, in Experiment 1, that even 4-year-olds perform above chance on simple duration-word magnitude discrimination.

5. General Discussion

The purpose of these studies was to determine how children begin their acquisition of duration words, like *second, minute,* and *hour,* prior to being taught their precise definitions in grade school. First, we asked whether children assign interim meanings to these terms during the long delay between when they begin producing these words in speech, at age 2 or 3, and the time that they acquire their adult definitions, at around age 7. Secondly, we asked what these early meanings might look like, and how they might be learned. We focused on two alternatives: (1) the Duration Mapping hypothesis that children begin by individually associating duration words with approximate durations or (2) the Lexical Ordering hypothesis that they initially interpret these words based on their rank ordering within the lexical class of duration words. Consistent with the Lexical Ordering account, we found that initially children were able to contrast and rank duration words, but had little to no knowledge of their absolute durations. Our results indicate that proficiency in estimating the absolute time encoded by duration words emerges relatively late, and may even rely on formal instruction in grade school.

In Experiment 1, children completed a forced-choice task in which they were asked to identify which of two duration words represented the longer duration. For example, children were told that one character performed an action for an hour and another performed the same action for a minute, and were then asked which character did the action "more." We found that by age 4, children chose the correct character more often than expected by chance, and that there were no systematic differences across different words. As children grew older, their performance on these items improved across the board, again without significant differences between items, suggesting that their knowledge of the rank ordering emerges somewhat holistically.

However, as shown in Experiment 2, children's understanding of duration words appears to be initially limited to knowing their rank ordering. In Experiment 2, we used the same paradigm, but now inserted number words to test whether children were sensitive to the absolute durations of time words. Our logic was that if children know that an hour is roughly 50-60 times longer than a minute, then they should also know that 2 hours is more than 3 minutes. On critical trials, we found that 4-year-olds performed at chance and that even 6-year-olds – who have robust understanding of number words and counting – performed more poorly than on contrasts with congruent cues, e.g., 3 hours vs. 2 minutes. These findings indicate that, despite having knowledge of the rank ordering of these terms (e.g., day > hour > minute), children have relatively little understanding of their absolute durations.

Finally, in Experiment 3, we corroborated and extended these results by asking children to estimate durations using a spatial representation of time. We found that by age 5, children were relatively accurate when placing numbers on a number line, and, while they tended to overestimate the durations of familiar events (e.g., of 'washing your hands' and 'eating lunch'), they "spaced out" those estimates similarly to adults, indicating that they do understand their relative durations. However, children performed very poorly when asked to estimate the durations represented by time words like *minute* and *hour*. Not only were the absolute positions of children's estimates on the timeline unlike those of adults, but the relative distances between their estimates were also far from adult-like until the early school years. Furthermore, we observed that the 6- and 7-year-old children who knew the formal definitions of duration words (e.g., one minute equals sixty seconds) were much better able to represent their relative durations on timelines than those who did not.

These results indicate that prior to learning their definitions, preschoolers knew both that duration words indicate lengths of time and that some words indicate longer lengths of time than others. However, although children often understood that an hour is longer than a minute, they generally did not know *how much* longer an hour was. Thus early in development they failed on comparisons like *3 minutes* vs. *2 hours* and were unable to represent the appropriate positions or distances between these terms on timelines. Critically, none of these failures are predicted under the Duration Mapping account, in which duration words are individually mapped onto approximate durations. If such mappings were present, we would expect children to not only rank-order the words correctly, but also to estimate their approximate durations (as they could for number words). Thus, children's interim meanings for these words do not appear to be given by their individual relationships to duration perception. Instead, these meanings appear to be defined by their relations to other time words. Specifically, each word is understood by virtue of its position in a rank ordering of the other duration words known to the child, based on the inference that each of these terms denotes a different duration.

These findings are consistent with those of Shatz and colleagues (2010), who found that, before children are able to use or comprehend duration words in an adult-like way, they understand that these terms belong to a common lexical category (e.g., they answer questions about duration with duration words). Extending the findings of Shatz et al., we show that the lexical category that children form for duration words is not a simple grouping of these words, but rather a structured, ordered scale that reflects some knowledge of the relative temporal magnitudes of the words (e.g., that an hour is longer than a minute). Nevertheless, this scale is not the same as the scale eventually adopted by adults, which also includes information about the absolute duration represented by each word, and about the proportional relationships between them (e.g., that an hour is sixty times longer than a minute). Instead, it is a more coarse ordinal scale in which words are ranked by increasing duration, with little to no information regarding absolute duration or the proportional relationships among the various terms.

Duration words are only one of many ways we talk about time, but some evidence suggests that children may use similar strategies to learn other types of time words. By the time children learn duration words, they have also learned grammatical tense and several other types of time words, including sequence terms (e.g., *before* and *after*) and deictic terms (e.g., *yesterday* and *tomorrow*) (Ames, 1946; Busby & Suddendorf, 2005; Bloom, 1970; Grant & Suddendorf, 2011; Harner, 1981, 1982; Nelson, 1996; Weist, 1986; Weist & Buczowska, 1978). Though the literature indicates that these expressions are likely acquired earlier than duration words, children's production of deictic and sequence time words also precedes adult-like comprehension. In each case, there is an extended period of inaccurate use in speech and poor performance on laboratory comprehension tasks (e.g., Clark, 1973; Harner, 1975; Harner, 1982; Trosberg, 1982; Weist et al., 1991). Furthermore, within-domain errors are also observed with these other time words. For instance, a child might substitute *before* for *after* or *tomorrow* for *yesterday*, or use a term like *yesterday* to refer to the past in general rather than the specific previous day (Bloom, 1970; Harner, 1981). Such specific errors indicate that children understand that these words refer to the domains of sequential and deictic time before they know precisely what each word means. Building on this category knowledge, it is possible that children can also use lexical contrast to learn these other time words.

Prior work has also shown that young children have ordered, list-like lexical representations for larger groups of temporal words, including the days of the week and the months of the year. In both these cases, children seem to learn the list of terms by rote, and can recite them prior to being able to use them to locate events in time or reason about the temporal distances between them (Friedman, 1986, 1989, 1991). Much like we see for duration words, in early development, children seem to know simply that the days of the week and months of the year are sets of contrasting time markers, without knowing how each one relates to the passing of time. For example, a child might know that December comes after November without knowing that December is in the winter, or that it's when Christmas happens, or that it has 31 days. Similarly, a child might know that Tuesday follows Monday without having any idea whether Tuesday is today or tomorrow or yesterday (Friedman, 1976, 1982).

Of course, calendar words also differ from duration words in several important ways that make it less surprising that children initially learn calendar words as ordered lists. The days of the week and months of the year, for example, have little content beyond their relative ordering, and no apparent link to perception. The common category membership of these terms is conveyed in their lexical form – e.g., all the days of the week contain the suffix *–day*. And, perhaps most critically, children are explicitly taught to recite them in order from an early age. Like the letters of the alphabet (A, B, C…) and the count list (1, 2, 3..), the sequence of days (Monday, Tuesday, Wednesday...) is highly rehearsed, making the ordering of these words more transparent for the child. Indeed, when asked to recite the days of the week in our lab, children as young as 3 often burst into song. All of these factors make the learning problem children face when acquiring this set of time words much less complex that than involved in duration word acquisition. When catchy mnemonic devices are available, it is unlikely that children would adopt any other learning strategy.

In contrast, there are no morphological hints that English words for duration refer to time in their names, nor any indication in their morphology that they go together in any way. Children are much less likely to receive instruction from adults on the rank order of these words, because they are not defined in terms of their order. Thus it is unlikely that 4-year-olds could acquire their rank order via rote memorization. Further, given the large numerosities and complex mathematical principles (e.g., multiplying by 60) involved in their formal definitions, it is also highly unlikely that children as young as 4 actually learn the ordering of duration words by receiving instruction on their definitions in preschool. Instead, we suggest that both the common category and the rank ordering of duration words are inferred indirectly from their use in adult discourse.

Given the evidence that children form categories for abstract words such as duration words prior to learning their formal definitions, one might first ask how children initially determine that duration words refer to lengths of time. In the absence of early "word-to-world" mappings between a lexical item and its perceptual referent, it has been suggested that children rely primarily on cues from their linguistic input when learning "hard" (i.e., abstract) words (Gleitman et al., 2005). One such cue is the syntax of the sentence. In a process known as "syntactic bootstrapping," children use their knowledge about the grammatical role a novel word plays in a larger sentential context to narrow their hypothesis space when entertaining possible meanings for the term (Gleitman et al., 2005). While it is possible that this type of inference plays a role in duration word acquisition (e.g., for identifying their grammatical status as measure words), we suspect that information from the larger discourse context may play a greater role. As noted by Tare et al. (2008), duration words are used in child-directed speech in a relatively wide variety of contexts, and, although some of these usage contexts do not provide information about their precise durations (e.g., the idiomatic usage of "in a minute"), they nevertheless may help children to establish their common category. In particular, information about the rank ordering of duration words could be gleaned from contrastive usages in speech (e.g., "I thought he'd only be 15 minutes late, but it's been an hour already"). Studies in progress are exploring this possibility via corpus analyses of child-directed speech.

The model of time word learning suggested by these studies is consistent with a broader theory of how children learn the meanings of abstract domains of words. In the cases of number words, color words, and emotion words, for example, children initially identify the conceptual domain to which a set of terms relate, then infer how those terms interrelate, and only later map them onto their extra-linguistic (e.g., perceptual) referents (Davidson et al., 2012; Le Corre & Carey, 2006; Shatz et al, 2010; Wagner, Dobkins & Barner, 2014; Widen & Russell, 2003; Wynn, 1990). In each of these cases, children quickly identify a set of words that refer to a common conceptual domain, and begin using the words in speech in a non-adult-like way. According to Carey (2009), such case studies are examples of what she calls "Quinian bootstrapping," whereby "an explicit structure is learned initially without the meaning it will eventually have, and at least some relations among the explicit symbols are learned directly in terms of each other” and “the ordering … exhausts their initial representational content” (pp. 329).

In the case of number words, for example, there is evidence that though children have access to approximate nonverbal representations of number from birth (e.g., Izard et al., 2009; Xu & Spelke, 2001; Xu, Spelke, & Goddard, 2005), they do not initially map number words onto these representations (Davidson, Eng, & Barner, 2012; Le Corre & Carey, 2006; but see Wagner & Johnson, 2012). Instead, they first memorize a subset of their count list, then gradually learn the individual meanings of a subset of those items (e.g., 1-2-3) over the course of many months, and eventually learn that the count list can be used to identify the cardinality of large sets (see Carey, 2009; Davidson, Eng, & Barner, 2012; Le Corre & Carey, 2006; Sarnecka & Lee, 2012; Wynn, 1990, 1992). Only after learning to count do they begin showing systematic evidence of mapping number words onto approximate magnitudes (Davidson et al., 2012; Le Corre & Carey, 2006), a process that likely involves creating a global structure mapping between the count list and the structure of nonverbal number representations (Sullivan & Barner, 2012, 2014). As with duration words, children appear to learn the structural relationships between number terms, prior to associating these terms with nonverbal representations, despite the fact that those representations are present from birth.

Like duration words and number words, children produce color words and recognize that they belong to a lexical class prior to learning their adult meanings (e.g., Backscheider & Shatz, 1993; Sandhofer & Smith, 1999; Wagner et al., 2014). Also, while early studies argued that children produce color words for many months before learning what the domain of words refers to, more recent work indicates that children know that color words represent color from the time they begin using them. Similar to what we’ve shown here for time, children begin by acquiring proto-meanings for color words that differ from the meanings understood by adults, and generally are broad overextensions (Wagner et al., 2014). According to Wagner et al., children acquire the precise adult-like meanings of color words by learning how color words relate to each other. Early, overly broad meanings for words like *green* and *red* are narrowed as children add words like *blue, pink* and *orange* to their vocabularies. Thus, much like the cases of time and number, color word learning involves not simply mapping words to perception, but also learning about how words within a lexical class relate to one another.

In conclusion, although the mapping between duration words and approximate durations does not appear to be intuitive for children, we find that children nevertheless learn quite a bit about the relationships between these terms in the preschool years, and assign proto-meanings to them based on the inference that they denote different durations. The three experiments presented here support the Lexical Ordering hypothesis, indicating that children learn the lexical category for time words as well as the ordered structure of that category prior to learning their formal definitions. However, our results also suggest that many children do not map these words onto precise representations of duration until after they learn their formal definitions.

Acknowledgments

We thank Eleanor Chestnut, Katherine Kimura, Saima Malik Moraleda, and Karen Bejar for their assistance with subject recruitment and data collection. Tyler Marghetis, Esther Walker, Jessica Sullivan, and members of the Language and Development Lab provided helpful comments. We are grateful to the Montessori Institute of San Diego Children's House, Hawthorne Children's Development Center, Kinderhouse Montessori School, La Jolla United Methodist Church Nursery School, Miramar College Child Development Center, World of Wonders Preschool, Honey Bear Preschool and Child Care Center, Del Mar Montessori, St. James Lutheran, and the Reuben H. Fleet Science Center, where data was collected. Most of all, we thank all the children and families who participated, without whom this research would not be possible. This work was funded by a grant to D.B. from the James S. McDonnell Foundation.

References

Ames, L. B. (1946). The development of the sense of time in the young child. The Pedagogical Seminary and Journal of Genetic Psychology,68(1), 97-125.

Au, T. K. F., & Markman, E. M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development*, *2*(3), 217-236.

St. Augustine (398/1992) *Confessions*. (H. Chadwick, Trans.) Oxford: Oxford University Press.

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

Barth, H., Starr, A., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. Cognitive Development, 24(3), 248-264.

Bates, D., Maechler, M., Bolker, B. & Walker, S (2013). *lme4*: Linear mixed-effects models using Eigen and S4 classes. R package version 1.0-5.

Bender, A., & Beller, S. (2014). Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition, 132,* 342-382.

Bloom, L. (1970). Language development: Form and function in emerging grammars. Cambridge, MA: MIT press.

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

Boroditsky, L. (2011). How Languages Construct Time. In Dehaene and Brannon (Eds.) *Space, time and number in the brain: Searching for the foundations ofmathematical thought.* Elsevier. ISBN: 978-0-12-385948-8

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental science*, *10*(6), 770-777.

Brown, R. (1973). *A first language: The early stages*. Boston: Harvard U. Press.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, 41(4), 977-990.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press: New York.

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and Time in the Child's Mind: Evidence for a Cross-Dimensional Asymmetry. Cognitive Science 34(3), 387-405.

Catania, A.C (1970). Reinforcement schedules and psychophysical judgments: a study of some temporal properties of behavior. In: W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules.* New York: Appleton-Century-Crofts.

Clark, E. V. (1971). On the acquisition of the meaning of 'before' and 'after'. *Journal of verbal learning and verbal behavior*, *10*(3), 266-275.

Clark, E. V. (1990) On the pragmatics of contrast. *Journal of Child Language 17,*417-431.

Colombo, J., & Richman, W. A. (2002). Infant timekeeping: Attention and temporal estimation in 4-month-olds. *Psychological Science*, 13, 475-479.

Common Core State Standards Initiative (2010). Common Core State Standards for Mathematics. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*(6), 803-808.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125-127.

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction?. *Cognition*, *123*(1), 162-173.

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, 266 (5604), 718–719.

De Villiers, J. G., & De Villiers, P. A. (1978). *Language acquisition*. Harvard University Press.

Droit-Volet, S., Clément, A., & Wearden, J. (2001). Temporal generalization in 3-to 8-year-old children. *Journal of Experimental Child Psychology*, *80*(3), 271-288.

Droit-Volet, S., & Wearden, J. H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology*, *80*(2), 142-159.

Droit-Volet, S., Tourret, S., & Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *The Quarterly Journal of Experimental Psychology, 57A,* 797-818.

Ferster C.B & Skinner B.F (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Fitzgerald, H.E., & Brackbill, Y. (1968). Interstimulus interval in classical pupillary conditioning. *Psychological Reports, 23* (2), 369–370.

Fitzgerald, H. E., & Brackbill, Y. (1976). Classical conditioning in infancy: development and constraints. *Psychological Bulletin*, *83*(3), 353.

Friedman, W. J. (1977). The development of children's understanding of cyclic aspects of time. *Child development*, 1593-1599.

Friedman, W. J. (1986). The development of children's knowledge of temporal structure. *Child development*, 1386-1400.

Friedman, W. J. (1989). The representation of temporal structure in children, adolescents and adults. *Advances in Psychology*, *59*, 259-304.

Friedman, W. J. (1990). Children’s representations of the pattern of daily activities. *Child Development, 61*, 1399–1412.

Friedman, W. J. (1991). The development of children’s memory for the time of pastevents. *Child Development*, 62(1), 139–155.

Friedman, W. J., & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development*, 357-371.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23-64.

Grant, J. B. & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*, 87-96.

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. *Developmental Psychology*, *11*(6), 864.

Harner, L. (1981). Children's understanding of time. *Topics in language disorders*, *2*(1), 51-66.

Harner, L. (1982) Talking about the Past and the Future. In Friedman, W. J. (Ed.) *The developmental psychology of time*. Academic Press.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. Proceedings of the National Academy of Sciences,106(25), 10382-10385.

James, W. (1890). *The principles of psychology* (Vol. 1). *New York: Holt*.

Kant, I. (2009/1781). *Critique of pure reason*. P. Guyer (Ed.). Cambridge University Press.

Kucera, H, & Francis, W. N,.(1967). Computational Analysis of Present-day American English. Providence: Brown University Press.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395-438.

Lewkowicz, D. J. (2003). Learning and discrimination of audiovisual events in human infants: the hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. *Developmental psychology*, *39*(5), 795.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, *76*(5), 978-988.

MacWhinney, B., & Snow, C. (1990). The child language data exchange system: An update. *Journal of child language*, *17*(02), 457-472.

McCormack, T., Brown, G. D. A., Maylor, E. A., Darby, R. J., & Green, D. (1999). Developmental changes in time estimation: comparing childhood and old age. *Developmental Psychology*, 35, 1143–1155.

McTaggart, J. (1908) The unreality of time. *Mind* 17, 457–474

Nelson, K. (1998). *Language in cognitive development: The emergence of the mediated mind*. Cambridge University Press.

Piaget, J. (1969). *The child's conception of time*. Routledge & K. Paul.

Provasi, J., Rattat, A. C., & Droit-Volet, S. (2011). Temporal bisection in 4-month-old infants. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 108.

R Core Team (2013). R: A language and environment for statistical computing.

Rivière, V., Darcheville, J. C., & Clément, C. (2000). Rapid timing of transitions in interreinforcement interval duration in infants. Behavioural processes, 52(2), 109-115.

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11,* 16-36.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. Psychological science, 14(3), 237-250

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General*, *142*(1), 193.

Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. Journal of the Experimental Analysis of Behavior, 11(3), 223–238.

Sullivan, J., & Barner, D. (2012). How are number words mapped to approximate magnitudes? *Quarterly Journal of Experimental Psychology, 66,* 389-402.

Sullivan, J., & Barner, D. (in press). Children's use of structure mapping in number line estimation. *Journal of Experimental Child Psychology.*

Tare, M., Shatz, M., & Gilbertson, L. (2008). Maternal uses of non-object terms in child directed speech: Color, number, and time. *First Language, 28,* 87–100.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *43*(2), 217.

VanMarle, K., & Wynn, K. (2006). Six‐month‐old infants use analog magnitudes to represent duration. *Developmental science*, *9*(5), F41-F49.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127* (3), 307-317.

Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of child language*, *18*(01), 67-92.

Whorf, B.L. (1956) The relation of habitual thought and behavior to language. In *Language, Thought, and Reality* (Carroll, J., ed.) MIT Press

Widen, S., & Russell, J. (2003). A closer look at preschoolers’ freely produced labels for facial expressions. *Developmental Psychology*, 39, 114–128.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155-193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220-251.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1-B11.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental science*, *8*(1), 88-101.

Chapter 1, in full, is a reprint of the material as it appears in Learning the language of time: Children's acquisition of duration words. *Cognitive Psychology, 78*, 57-77. Tillman, K. A., & Barner, D. (2015). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier.

Chapter 2.

Today is tomorrow's yesterday:
Children's acquisition of deictic time words

Katharine A. Tillman[a], Tyler Marghetis[c], David Barner[a,b], and Mahesh Srinivasan[d]

[a]Department of Psychology [b]Department of Linguistics
University of California, San Diego

[c]Department of Psychological and Brain Sciences
Indiana University, Bloomington

[d]Department of Psychology
University of California, Berkeley

Abstract

Deictic time words like “yesterday” and “tomorrow” pose a challenge to children not only because they are abstract, and label periods in time, but also because their denotations vary according to the time at which they are uttered: Monday’s “tomorrow” is different than Thursday’s. Although children produce these words as early as age 2 or 3, they do not use them in adult-like ways for several subsequent years. Here, we explored whether children have partial but systematic meanings for these words during the long delay before adult-like usage. We asked 3- to 8-year-olds to represent these words on a bidirectional, left-to-right timeline that extended from the past (infancy) to the future (adulthood). This method allowed us to independently probe knowledge of these words’ deictic status (e.g., “yesterday” is in the past), relative ordering (e.g., “last week” was before “yesterday”), and remoteness from the present (e.g., “last week” was about 7 times longer ago than “yesterday”). We found that adult-like knowledge of deictic status and order emerge in synchrony, between ages 4 and 6, but that knowledge of remoteness emerges later, after age 7. Our findings suggest that children’s early use of deictic time words is not random, but instead reflects the gradual construction of a structured lexical domain.

K, age 4: “Yesterday … that’s last night’s morning.”

1. Introduction

To learn their first words, children must rely primarily on the extra-linguistic context in which those words are used, since they are not yet able to understand the sentences in which the words are embedded. Consequently, children’s first words often label concrete referents that can be ostensively identified, like “mama,” “doggie,” and “cup.” Other words, however, are more difficult to learn through observation of the world and may require children to recruit their knowledge of the linguistic context in which those words are embedded (e.g., Gillette, Gleitman, Gleitman & Lederer, 1999; Gleitman, 1990; Gleitman et al., 2005; Snedeker & Gleitman, 2004). For example, the meanings of *deictic time words,* such as “yesterday” and “tomorrow,” cannot be gleaned solely from extra-linguistic situations. These words are abstract and describe periods in time, which are difficult to reference ostensively. Further, due to their deictic functions, these words do not have fixed denotations and cannot be understood without information about the time at which they are uttered (Fillmore, 1979/1999): Tuesday’s “tomorrow” is different from Wednesday’s “tomorrow.” Acquiring words like these is one of the greatest challenges that English-learning children face, as evidenced by the massive gap between their first use of deictic time words around age 3 and their eventual mastery of adult-like meanings in elementary school (Ames, 1946; Busby Grant & Suddendorf, 2011; Harner, 1975, 1981). However, the process through which these words are ultimately acquired—and thus the roles of linguistic and referential context—remains mysterious. Here, as a case study of abstract word learning, we explore children's gradual construction of deictic time word meanings between ages 3 and 8.

While many children produce words like "yesterday" and "tomorrow" as early as age 2 or 3 (Ames, 1946; Busby Grant & Suddendorf, 2011; Dale & Fenson, 1996), they do not use them as adults do for several subsequent years (Ames, 1946; Busby Grant, & Suddendorf, 2011; Harner, 1975, 1981; Nelson, 1996; Weist, 1989; Weist et al., 1991). According to parental report, two thirds of 3-year-olds produce the word "yesterday," but fewer than 20% use the word in adult-like ways; by age 5, more than 80% of children produce "yesterday," but, still, fewer than 60% use it like adults (Busby Grant & Suddendorf, 2011). Children struggle not just with production but also with comprehension: When asked to name an event that occurred "yesterday" or one that will occur "tomorrow," only about a quarter of 3-year-olds can provide reasonable answers (Busby & Suddendorf, 2005; Suddendorf, 2010). These difficulties persist even later in acquisition, as 5-year-olds can correctly generate an event from "yesterday" only 66% of the time, and an event that will occur "tomorrow" only 63% of the time (Busby & Suddendorf, 2005).

Although children differ from adults in how they use time words, it remains possible that they nevertheless use them systematically, and that they construct their meanings gradually and in stages over the first 6 or 7 years of life. Consider an anecdote: When 21-month-old Franny tried to remove dirty dishes from the dishwasher, her mother stopped her and said, "We can empty it *tomorrow.*" Upon hearing this, Franny ran to her bedroom, climbed under her blanket, closed her eyes, and after a brief delay returned to the kitchen to begin the chore. For Franny, "tomorrow" seemed to mean something like "after waking up." Just a few months later, Franny began producing the word "yesterday," but used it to refer not only to events that happened the previous day, but also to events that happened two days ago, five minutes ago, and even several months earlier. Productions like Franny's are thought to be quite common (Ames, 1946; Friedman, 1990; Harner, 1981; 1982; Nelson, 1996; Weist, 1989) and suggest that although young children do not use deictic time words in adult-like ways, they may have partial knowledge of their meanings.

Critically, knowledge of partial word meanings may not have been detectable by the comprehension measures used in previous studies (i.e., parental report and event naming). For example, although adult English speakers may judge that Franny fails to use "yesterday" correctly, Franny may nonetheless know that "yesterday" refers to a period of time, and that it refers to periods prior to the time at which it is uttered. Further, even if Franny were to develop an adult-like meaning of "yesterday" and understand that it refers to a specific period exactly one day ago, she might not be able to name an event that occurred "yesterday." The ability to associate time words with life events depends not only on knowledge of these words' meanings, but also on the ability to recall, order, and anticipate events (e.g., a capacity for "mental time travel"; Suddendorf & Corballis, 2007). These abilities develop slowly (Busby & Suddendorf, 2005; Suddendorf, 2011; Schachner, Addis, & Buckner, 2007). These considerations suggest that other methods are required to probe children's knowledge—or partial knowledge—of deictic time word meanings.

Understanding the nature of children's early uses of deictic time words—and the partial word meanings they may implicate—could provide critical insight into the inductive hypotheses children make about these words' meanings. While there has not been systematic study of children's partial knowledge of deictic time words during the long delay between initial production and adult-like usage, there are hints that children may acquire information about different facets of their meanings independently, with some acquired before others. These facets include a word's *deictic status* (e.g., "yesterday" is in the past; "tomorrow" is in the future), its sequential *order* relative to other time words (e.g., "next week" is a time after "tomorrow"), and its *remoteness* from the present (e.g., "yesterday" is exactly one day from today). For instance, 3-year-olds appear to understand that "yesterday" and "tomorrow" refer to a non-present time, without knowing that they refer specifically to the past and future, respectively (Harner, 1975). Further, children struggle to grasp the differing causal implications of events from "yesterday" vs. "tomorrow" on the present until at least age 5, also suggesting that their understanding of the distinction between past and future is incomplete (Busby & Suddendorf, 2010). Together, these results suggest that children may first learn that deictic time words label periods in time, without understanding much about their deictic past/future status, order, or remoteness. Furthermore, children's over-extension errors *within* the past or future, like Franny's use of "yesterday," suggest that at some stage, children may understand a word's deictic status without understanding its remoteness (e.g., Harner, 1981; Nelson, 1996).

One reason to think that children may acquire information about a word's deictic status, order, and remoteness separately is that there is substantial variation in how these facets of time are expressed across languages. In English, for instance, all time words refer to either the past, present, or future. By contrast, in Urdu, "kal" refers to a period exactly one day from the present—whether in the future or the past—and thus does *not* encode deictic status but *does* encode temporal remoteness. Other languages include terms that encode degrees of temporal remoteness that are not lexicalized in English. For example, German's "übermorgen" and Georgian's "zeg" label a period that is in the *future*, much like English "tomorrow," except that they refer to the period exactly *two* days away. Meanwhile, German's "vorgestern" and Japanese's "ototoi" pick out a time in the past, like English "yesterday," except that it refers to the period exactly *two* days ago. The fact that deictic time words vary across languages according to factors like deictic status and remoteness is consistent with the idea that these facets of meaning are dissociable and may be learned independently by children.

The goal of the present study was to explore whether English-learning children have systematic but partial meanings of deictic time words during the long delay between their initial production of these words and eventual adult-like usage. Critically, the nature of children's partial meanings—i.e., the developmental sequence in which information about deictic status, order, and remoteness are acquired—could constrain theories of the process through which these words are learned and the informational sources that children might draw upon. Broadly, there are two sources of information a child might use to learn the meanings of these terms: The *events* that time words refer to (e.g., a birthday party), and the *linguistic context* in which these words appear. As we describe below, these sources of information are differentially suited to supporting children's inferences about deictic status, order, and remoteness.

Mappings between deictic time words and the events they are used to describe could plausibly help children learn the deictic status and remoteness of these words. For example, by noting whether deictic time words are used to describe events that are anticipated (e.g., a birthday party *tomorrow* / *next week* / etc.) or in the past (e.g., a birthday party *yesterday / last week* / etc*;* Johnson et al., 1988) children could learn the deictic status of these words. Children could also generate hypotheses about the approximate temporal remoteness indicated by these words, e.g., by using the strength of a memory trace to estimate the remoteness of a party that occurred "last week" (Hinrichs, 1970; Friedman, 2003). From an understanding of deictic status and remoteness, children could then make inferences about relative order: e.g., *last year* is before *yesterday* because it refers to a more remote time in the past. Thus, if children were to rely primarily on *event mappings*, then they might acquire knowledge of deictic status and remoteness in tandem, and later infer information about order.

If children leverage the broader linguistic context to learn deictic time word meanings, they might exhibit a different developmental trajectory, in which knowledge of deictic status and/or order is constructed prior to knowledge of remoteness. To begin, even before children have learned anything about the meanings of deictic time words, they could use the linguistic context to infer that deictic time words belong to a common lexical class. For example, children might observe that deictic time words often appear in similar sentence frames (e.g., "The party [happened/will happen] [yesterday/tomorrow/last week/next year]."), or that they are often used in response to "When" questions, and from this infer that these words have similar kinds of meanings (Backshneider & Shatz, 1993; Tare, Shatz, & Gilbertson, 2008; Shatz, 1993).

After grouping deictic time words into a common class, children could use other linguistic cues to make inferences about the specific semantic content of these words. For example, children could use their early knowledge of English tense markings (Brown, 1973; De Villiers & De Villiers, 1978; Harner, 1976; Weist et al., 1991) to infer whether deictic expressions refer to events in the past (e.g., "He **danced** *last year*") or in the future (e.g., "He **will dance** *tomorrow*"). This process, in which grammatical cues are used to restrict hypotheses about meaning, is known as *syntactic bootstrapping* (Brown, 1957; Gleitman, 1990; Gleitman et al., 2005; Landau & Gleitman, 1985; Naigles, 1996). Further, children could also leverage cues from *discourse structure* to infer the relations among deictic time words. For example, contrastive uses of these words such as "The package isn't coming *tomorrow*, it's coming later, *next week*" could provide information about relative order. Moreover, since order-of-mention in discourse typically reflects temporal order (Jakobson, 1966), children could use input such as, "Bobby danced at his birthday party *last year*, but probably won't dance at his friend's party *tomorrow,*" to infer the sequential order of individual lexical items like "last year" and "tomorrow." In sum, if children rely primarily on the linguistic context to learn deictic time word meanings, then they might acquire knowledge of deictic status (supported by syntactic bootstrapping) and order (supported by discourse structure) prior to knowledge of remoteness.

The present study

Here, to explore whether children might have systematic but partial meanings for deictic time words during the gap between initial production and later adult-like usage, we asked 3- to 8-year-old children to place these words onto a spatial timeline. As described below, this method allowed us to separately assess knowledge of these words' deictic status (i.e., past vs. future), sequential order, and remoteness from the present.

Spatial scales have been used extensively to study children's mental representation of number (e.g., Barth, Starr, & Sullivan, 2009; Booth & Siegler, 2006; Ebersbach et al., 2008; Kolkman et al., 2013; Laski & Siegler, 2007; Moeller et al., 2009; Siegler & Booth, 2004; Siegler & Opfer, 2003; Slusser at al., 2013; Sullivan & Barner, 2014), and a smaller number of studies have used spatial scales to assess children's understanding of temporal sequence (Busby Grant & Suddendorf, 2009; Friedman, 2000, 2002; Friedman & Kemp, 1989; Hudson & Mayhew, 2011). However, most timeline studies involving preschoolers have explicitly *avoided* deictic time words, and, furthermore, have been limited in their ability to tease apart children's understanding of the different semantic facets of these terms. For instance, all timeline studies of preschoolers have used scales depicting either the past or the future, but not both simultaneously (Busby Grant & Suddendorf, 2009; Friedman, 2000, 2002; Friedman & Kemp, 1998). Thus these studies could not gauge children's knowledge of deictic status, because children never had to decide whether events came from the past or the future. Furthermore, most timelines used to test preschoolers were divided categorically into distinct regions that represented broad periods of time, such as "a short time ago" or "a long time ago" (e.g., Busby Grant & Suddendorf, 2009; Friedman, 2002; Friedman & Kemp, 1998). Since terms placed inside the same region are not distinguished from each other, it is difficult and sometimes impossible to use categorical timelines to probe children's knowledge of sequential order or remoteness.

We thus developed a new timeline task that allowed us to independently assess children's knowledge of deictic status, sequential order, and remoteness. In our task, 3- to 8-year-old children and adult controls used colored pencils to mark where deictic time words (e.g., “yesterday”) and events (e.g., the participant’s last birthday) should go on horizontal timelines that extended, left-to-right, from the past (“when you were a baby”) to the future (“when you’ll be a grown-up”), with the present moment (“right now”) indicated by a dividing line between past and future. Knowledge of the deictic status of a word was assessed by its placement to the left or right of the midpoint, regardless of its placement relative to other words. Knowledge of sequential order was assessed by the ordering of words along the line (e.g., whether “last week” was placed before “yesterday”), ignoring their relation to the present and the distances between them. And knowledge of remoteness from the present was assessed by the spacing of terms along the line (e.g., the distance between “last year” and “now,” compared to the distance between “yesterday” and “now”). Finally, to confirm that timeline performance was a valid measure of children’s developing semantic knowledge, we investigated whether it correlated with children’s ability to answer verbal questions about the meanings of deictic time words (e.g., “Which will happen first: Tomorrow or next week?”).

2. Experiment 1

2.1 Methods

2.1.1 Participants

Children from the greater San Diego, CA, (n=93) and Berkeley, CA, (n=25) areas participated in this experiment, along with 38 young adult controls from the UC San Diego Psychology Department subject pool. Data collection continued until we reached our target of sixteen children in each age category. Since recruited participants also participated in a related study on temporal gesture, this target sample size was based on past studies on children’s gesture (e.g., Sauter et al, 2013). The total child sample[3] included 17 3-year-olds, 20 4-year-olds, 18 5-year-olds, 26 6-year-olds, 19 7-year-olds and 17 8-year-olds. An additional 15 children participated but were excluded from analysis due to failure to complete the task (n=6), illegible timelines (n=6), experimenter error (n=1), and clerical error (n=2). Children were tested in local preschools, elementary schools, museums, or our laboratories at UC San Diego and UC Berkeley. Parents gave informed consent for their children to participate in the study, and children indicated their willingness to participate before testing began. Parents who brought their children into the laboratory were compensated for their travel expenses, and children received a small gift in thanks for their participation. Adult controls were awarded course credit for their participation

2.1.2 Materials

Materials included one pack of colored pencils and one 8.5in. × 11in. sheet of paper with three 13.5 cm horizontal timelines printed on it. Timelines were positioned vertically down the center of the page. Small dots indicated the left and right endpoints of the line and a vertical line indicated the midpoint (see Fig. 1). An icon of a baby was placed to the left of the timeline and an adult figure to the right, to remind children of the timeline’s interpretation. Each timeline was associated with a preset list of items (Table 1), selected to test for a range of possible error patterns. Using preset lists also maintained

---

[3] In some cases, a child’s video data—recorded for a related study on temporal gesture—was impossible to analyze (e.g., background noise was too loud; file was corrupted). In these cases, an additional child was recruited. All data were analyzed.

timeline complexity across individuals,[4] allowing direct comparison of error patterns across individuals and age groups.

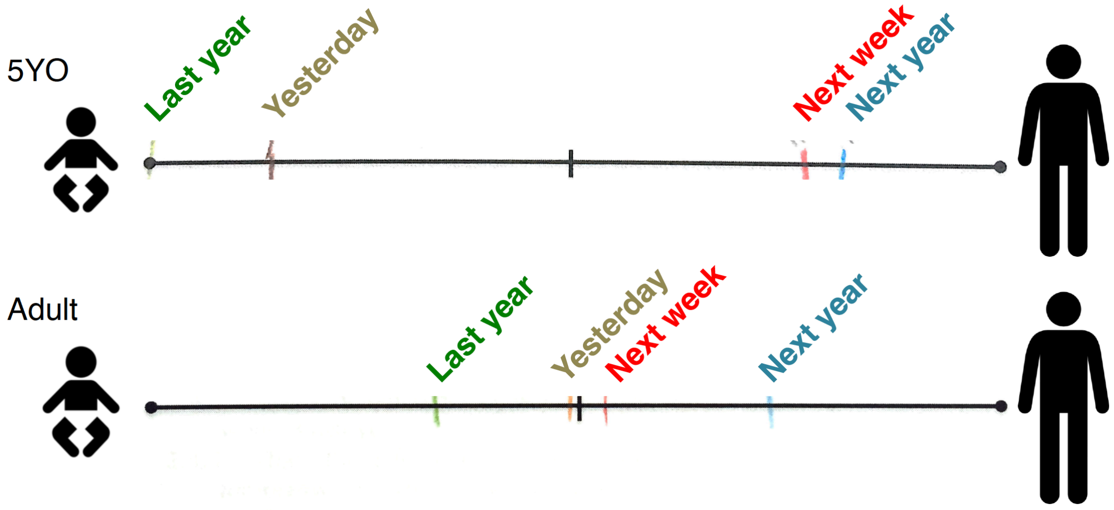

Figure 2.1. Timelines used by participants to indicate the relative location of deictic time words (e.g., “tomorrow”) and events (e.g., “breakfast today”), with example data from a 5-year-old and an adult. The adult placed “last year” and “next year” equidistant from the present moment, “next week” considerably closer to the present, and “yesterday” just barely in the past. By contrast, the child’s placements did not track the relative remoteness of each term: “yesterday” is farther away from now than “next week,” which is nearly as remote as “next year.” This child thus appears to exhibit knowledge of deictic status and order, but not remoteness.

2.1.3 Timeline procedure

The paper and colored pencils were placed in front of the child. The experimenter explained that a timeline “shows *when* different things happen,” from the past to the

---

[4] One downside of this approach—controlling the complexity of each timeline using a preset item list—is that it limits the interpretation of item analyses. The preset lists likely made some comparisons more difficult than others: “tomorrow,” for instance, was always on a timeline with two items that occur within a day, but items placed with “yesterday” were all a week or a year away. While this study was not designed to investigate mastery of particular words or concepts, we nevertheless report some exploratory item analyses in Appendix B.

future, with the present represented in the middle, and each time having its own place on the line. The child was instructed to indicate *when* different events or time periods happen by marking their location on the line, using colored pencils to draw a vertical mark (see Appendix A for complete script). The experimenter then introduced the first target item (e.g., “breakfast today”), asked the child to think of the time or event, and asked them to draw a line using a colored pencil to indicate when the target item occurred (Fig. 1). To facilitate subsequent coding, each item was associated with a particular color of pencil. This was repeated for the rest of the items, on all three timelines. In a few cases children were unwilling to mark the timeline but were willing to point to where they thought the target should go. In these cases, the experimenter drew the lines at the point where the child’s finger met the timeline. Participants always received the Events line first (to orient them to the task), but the order of the other two lines was counterbalanced between subjects. For each line, half of the subjects received the items in the order shown in Table 1, and the rest received the reverse order. If children asked questions about where they should put the lines, they were told, “It’s up to you! Put it wherever you think [e.g., tomorrow] goes on the line.” If participants asked to be reminded what their marks from prior trials referred to, the experimenter provided the answer.

Table 2.1. Target items used in timeline tasks.

Timeline	Item 1	Item 2	Item 3	Item 4
Events	breakfast	next birthday	dinner	last birthday
Time Words 1	last week	tomorrow	tonight	this morning
Time Words 2	next week	next year	yesterday	last year

2.1.4 Timeline coding

All analyses were conducted using the distance from the midpoint (0) to each mark on the timeline, which was used to determine the relative ordering of items along the timeline and whether the item was placed in the past (negative values) or the future (positive values). For our analysis of remoteness, we standardized distances from the midpoint by dividing each raw distance by the maximum distance at which any item was placed on that timeline (see Results, Section 3.1.3). Thus, distances ranged from 0 (midpoint) to 1 (farthest mark along the timeline). This was done to control for possible age-related differences in absolute placements (e.g., given that 3-year-olds might systematically place their last birthday closer to when they were a baby than adults would). Analyses were conducted using the *R* software package (R Core Team, 2013).

2.1.5 Verbal forced-choice questions

To gauge basic event and time word comprehension in a non-spatial task, participants were also asked 8 verbal forced choice questions. These questions were of two types. *Event* questions asked about when an everyday event occurred (e.g., Can you think about when you ate breakfast? Was that *this morning* or *tonight*?). *Order* questions asked about the relative ordering of time words (e.g., Which will happen first: *tomorrow* or *next week*?). Event questions preceded Order questions. For each type, two past-related and two future-related questions were asked. The experimenter introduced the past and future sets of Order questions by saying, "Now I'm going to tell you about things that *have already happened*" or *"… are going to happen,"* respectively. Within each block, half the children received the questions in the order listed in Table 2, and half received the reverse order. Half of the participants answered the forced-choice questions before performing the timeline task, and the other half answered them afterward. Whether the correct answer was mentioned first or second within a question was counterbalanced between questions.

Table 2.2. Verbal forced-choice questions.

Question Type	**Time**	**Prompt**	**1st alternative**	**2nd alternative**
Event		*Think about when you…*	*Was that…*	*Or…*
	Past	*…ate breakfast.*	this morning?	tonight?
		…were [age-1] years old.	last year?	next year?
	Future	*…are going to eat dinner.*	this morning?	tonight?
		…are going to be [age+1] years old.	last year?	next year?
Order				*Or…*
	Future	*Which will happen first:*	tomorrow?	next week?
			next year?	next week?
	Past	*Which happened first:*	last week?	last year?
			yesterday?	last week?

2.1.6 Unreported measures

As part of an ongoing project on spontaneous temporal gestures, all participants (except the 3- and 4-year-olds) also completed a structured interview that was designed to elicit temporal gestures. The interview consisted of open-ended questions that contrasted animal names ("What's the difference between a *cat* and a *dog*?"), vehicles (*car* vs. *motorcycle*), and several pairs of time words (e.g., "yesterday" vs. "tomorrow"; "next week" vs. "last year"). Children received positive encouragement but no feedback regarding their responses. The order of the gesture interview and timeline tasks was counter-balanced across subjects. Since task order did not affect any of our dependent measures (all *F*s < 0.4, *p*s > 0.4), subsequent analyses collapsed across order. No other measures were collected.

3. Results

We report four analyses of the timeline data: First, we assessed comprehension of the deictic status, sequential order, and temporal remoteness of the deictic time words. Second, we determined the typical ages of acquisition of these facets of meaning, pinpointing the age at which the majority of children displayed adult-like comprehension of each facet. Third, we calculated the contingencies between adult-like knowledge of these three facets of meaning: i.e., the degree to which adult-like knowledge of deictic status, order, and temporal remoteness predicted one another within an individual child. Finally, we asked whether children's performance on the timeline task predicted their ability to answer non-spatial, verbal forced-choice questions about time word meanings.

3.1 Facets of meaning

3.1.1 Deictic status. As an index of knowledge of deictic status, we calculated the average accuracy for all items' placement relative to "now" (e.g., "tomorrow" should be in the future) for each timeline and subject, and then calculated a mean deictic status accuracy for each subject and each type of timeline (i.e., Deictic vs. Event). We then analyzed deictic status accuracy with a mixed ANOVA, with Timeline Type (Deictic *vs.* Event) as a within-subjects factor and Age (3 through 8 years old, and adults) as a between-subjects factor. There was no effect of Timeline Type, suggesting that children were equally able to represent the past/future status of time words and events on the timelines. The only effect to reach significance was the main effect of Age, $F_{(6,\ 146)}$ =

32.6, $p < .001$ (Fig. 2A). While 3-year-olds performed at chance (0.5) overall, $t_{16} = 0.36$, $p = 0.73$, 4-year-olds were better than chance, $t_{19} = 4.5$, $p < .001$. Mean accuracy improved monotonically among 3- to 7-year-olds, $M_3 = .51 < M_4 = .65 < M_5 = .75 < M_6 = .85 < M_7 = .94$, and but did not differ between 7- and 8-year-olds, $M_8 = .92$, $t_{32} = 0.44$, $p = .67$. Six-year-olds performed significantly worse than adults, $t_{62} = 4.2$, $p < .001$, while 7-year-old performance did not differ significantly from adults, $M_{\text{adults}} = .97$, $t_{54} = 1.9$, $p = .07$.

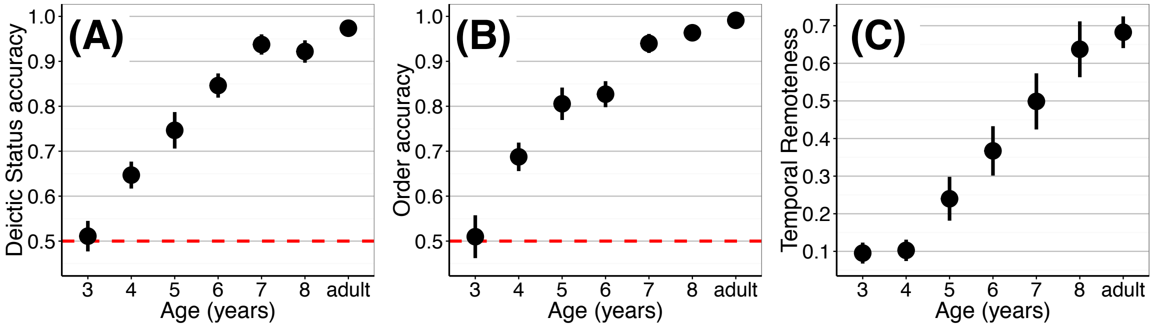

Figure 2.2. Developmental time-course of (A) deictic status, (B) order, and (C) remoteness knowledge. For all three facets, knowledge improved with age, although improvement was delayed for remoteness**.** Error lines = SEM; dashed line = chance performance.

3.1.2 Order. We next assessed knowledge of relative sequential order, separately from deictic status. Because children had to place four items on each timeline, this could have taxed their working memory, leading them to forget which items they had already placed on a timeline when placing other items. To control for this possibility, we compared the placement of each item relative to the placement of the immediately preceding item, rather than to the timeline as a whole (e.g., if “last week” is tested just after “tomorrow,” it should be placed to the left of “tomorrow”). This measure of pairwise order knowledge places minimal working memory demands on children, since they need only recall the meaning of an immediately preceding item. We calculated mean accuracy on this 1-back measure of order knowledge, for each participant and Timeline Type. There was only a main effect of Age, $F_{(6,\ 146)} = 27.8$, $p < 0.001$ (Fig. 2b). Just as with deictic status, 4-year-olds performed significantly above chance, $t_{19} = 4.9$, $p < 0.001$, while three-year-olds did not, $t_{16} = 0.18$, $p > .8$. Eight-year-olds' orderings were indistinguishable from those of adults, $t_{52} = -1.6$, $p = .10$, but 7-year-olds' were significantly different, $t_{54} = -2.5$ $p = .02$.

We also conducted analyses involving two additional measures of order knowledge, reported in Appendix C. The first was a whole-timeline measure, which assessed children's rank ordering of all four items on each timeline. Analyses using this measure yielded the same pattern of results as those using the 1-back measure of order knowledge. The second additional measure of order knowledge evaluated participants' relative placement of the two items involving the past for a given timeline (e.g., *yesterday*, *last year*) separately from their placement of the two items involving the future (e.g., *next week*, *next year*). Although this measure relied on only two comparisons from each timeline, it ensured that participants who only understood the past/future status of the items – e.g., that *yesterday* is in the past and *next week* in the future – wouldn't also be credited with knowledge of their order. Analyses using this measure converged with those using the other two order measures, with the exception that children did not reliably perform above chance until age 5 (likely due to a loss of statistical power).

3.1.3 Remoteness. We next evaluated knowledge of the temporal remoteness of each item—i.e., its relative distance from "now." To account for absolute differences in the amount of space used by participants, we first standardized distances from the midpoint by dividing each raw distance by the maximum distance at which any item was placed on that timeline. This approach controls for possible age-related differences in absolute placements (e.g., adults might place “last year” farther from “when you were a baby” than do 3-year-olds), and focuses instead on remoteness relative to the placement of other items on the same timeline. The median distance of each item placed by participants in each age group is shown in Figure 3.

To characterize the maturity of children’s representations of remoteness, for each child, we used multiple regression to see how well each item's remoteness was predicted by its mean remoteness among adults. A child’s knowledge of temporal remoteness was measured by the strength of the relationship between their placements and adult-like placements, after factoring out the child’s knowledge of order (i.e., semi-partial correlation squared). Thus, this measure assessed how much children understood about the remoteness of these terms *above and beyond* knowledge of their order.

Mean remoteness knowledge for each age group is shown in Fig. 2c. A linear regression revealed that children’s knowledge of remoteness improved gradually with age, $b = 0.12$, $t_{113} = 7.934$, $p < .001$. In contrast to our analyses of deictic status and 1-back order knowledge, we found that 4-year-olds performed no differently than 3-year-olds on our measure of remoteness knowledge, $t_{35} = 0.18$, $p = .86$. Five-year-olds’ understanding of remoteness was significantly more adult-like than that of 4-year-olds, $t_{36} = 2.2$, $p = .04$. Seven-year-olds performed significantly differently from adults, $M_7 = .50$ vs. $M_{\text{adult}} = 0.68$, $t_{54} = 2.3$, $p = .025$, but 8-year-olds did not, $M_8 = 0.64$, $t_{52} = 0.6$, $p = .58$ (Fig. 3).

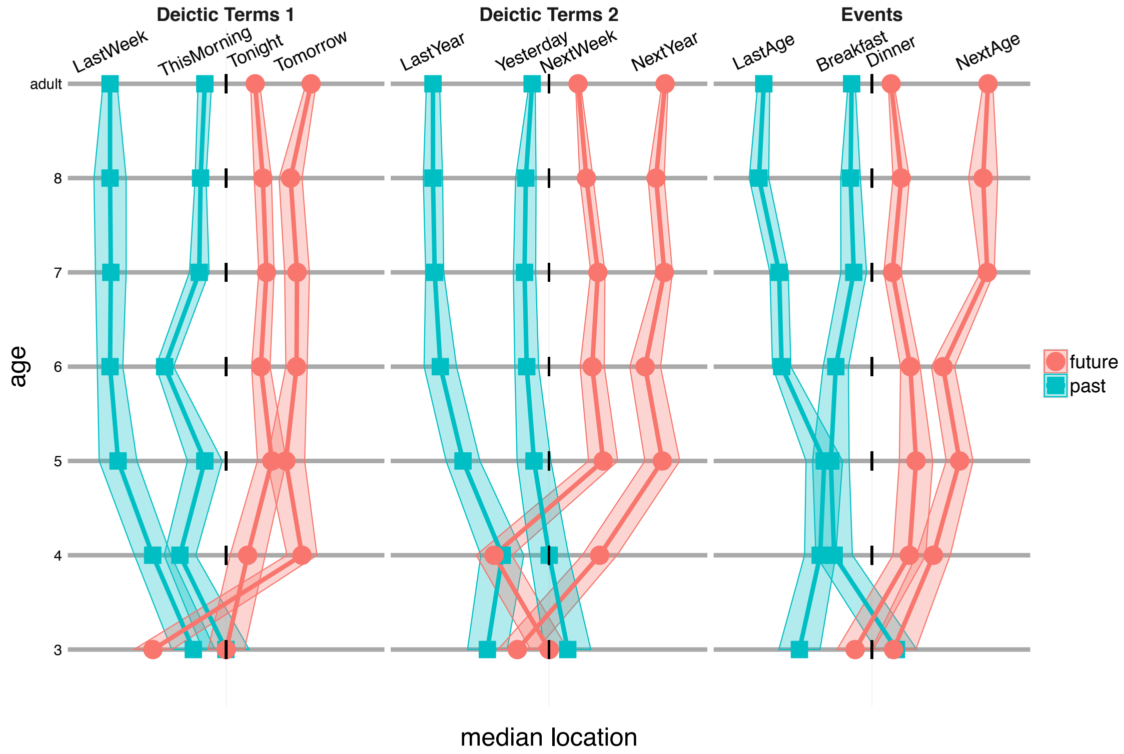

Figure 2.3. Median placements of each item on the timeline. Only the youngest children reliably made errors representing the deictic status and relative ordering of deictic time words (bottom timelines). Knowledge of relative temporal remoteness continued to develop until quite late (cf., the placement of "next week" vs. "next year" in 8-year-olds vs. adults). In order to control for effects of age on raw timeline placements, locations were rescaled for each participant, such that the location of the most distant item was equal to ±1. Error bands = SEM.

3.2 Order of acquisition

Together, the analyses described above suggest different time courses for the acquisition of different facets of deictic time word meaning. On the one hand, mean levels of performance on our deictic status and 1-back order measures were above chance by age 4. Further, by age 7, deictic status accuracy reached adult levels, and order accuracy exceeded 90% correct. On the other hand, performance on our remoteness measure accelerated more slowly, and continued to increase through age 8.

Next, we directly investigated the sequence in which these different facets of meaning are acquired. In order to make direct comparisons across our three different measures, we used a threshold approach, comparing the ages at which the majority (50%) of children had made the transition to adult-like understanding of each facet of meaning. First, based on the continuous measures of deictic status, order, and remoteness described above, we characterized each child, via k-means clustering, as a "knower" or "non-knower" of that facet of meaning. ("Knowers" were those children who clustered with adult participants; Fig. 4, A-C.) Next, for each facet, we modeled the transition to adult-like knowledge using a Weibull function, allowing us to estimate the age at which the majority of children exhibit adult-like knowledge of each facet, independently of the other two facets.

Our analyses reveal that the majority of children transitioned to adult-like knowledge of deictic status and order before their sixth birthday (deictic status: 5;4, bootstrapped 95% CI [4;10, 5;9]; order: 5;8, [5;1, 6;1]). By contrast, most children did not transition to adult-like knowledge of temporal remoteness until nearly two years later (7;3, [6;11, 7;9]). Since the confidence interval for remoteness does not overlap with the other confidence intervals, this delay is statistically significant.

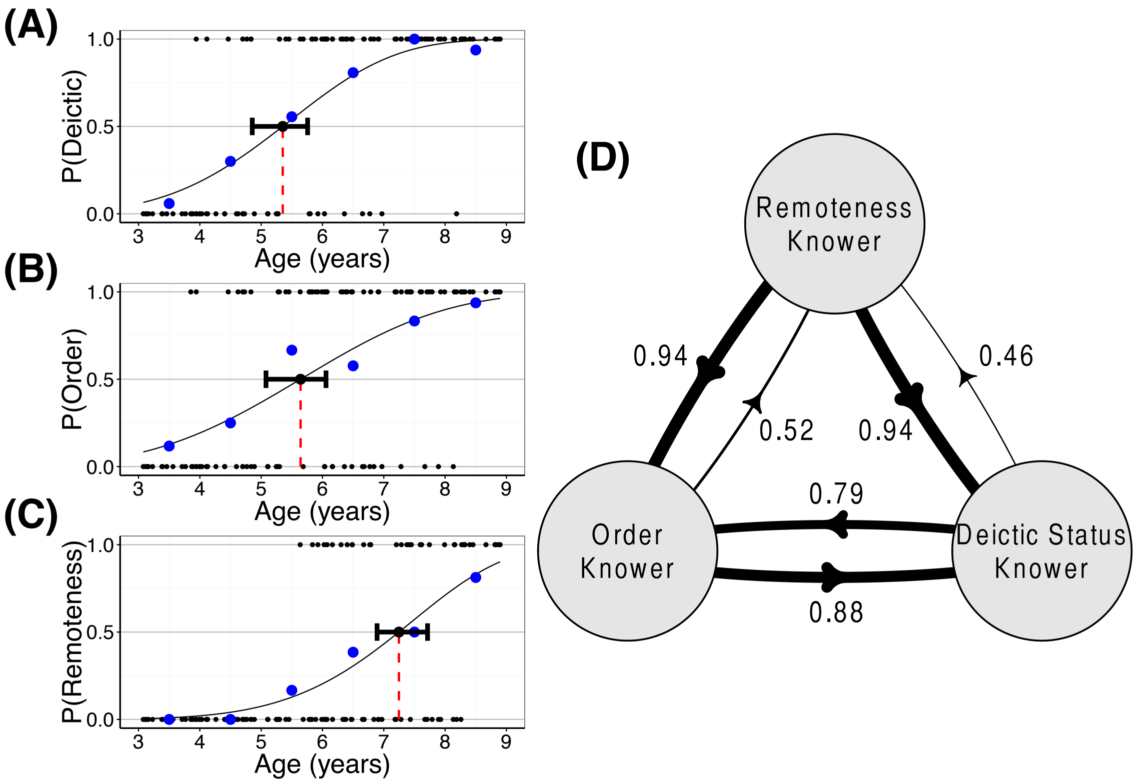

Figure 2.4. (A) Age of acquisition for three facets of deictic time word meaning. Each black dot represents an individual child, categorized as either a "knower" or "non-knower" of that dimension. While adult-like knowledge of deictic status and order emerged around the same time, before age 6, adult-like knowledge of remoteness was delayed for nearly another two years. Blue dots indicate the mean probability of being a knower in each age group. Block dots and red vertical lines indicate age-thresholds after which the majority of children are knowers (i.e., $p \geq 0.5$); error bars indicate bootstrapped confidence intervals on those age-thresholds. **(B)** Contingencies among facets of meaning. Arrows denote direction of influence; numbers indicate the conditional probability of knowing the target facet of meaning, given knowledge of the source facet; line widths visualize these conditional probabilities (scaled from .4 to 1).

3.3 Learning contingencies

The previous analysis (section 3.2) suggests that on their sixth birthday, only a quarter of children exhibit an adult-like understanding of the remoteness of deictic time words, while the majority of children know their deictic status and their relative order. But this analysis does not tell us whether the *same children* who understand one facet of deictic terms' meanings also understand other facets. To address this, we calculated the conditional probability of being each type of "knower," given one's "knower" status on each other facet of meaning (Fig. 4d).

Interestingly, we found that knowledge of deictic status and order were highly linked: Among deictic-status-knowers, 79% also exhibited adult-like knowledge of order (95% CI [69%, 88%]); conversely, among order-knowers, 88% also exhibited adult-like knowledge of deictic status ([79%, 96%]. However, while it was extremely common for a child who was a remoteness-knower to also be a knower of deictic status or order (both 94%, [87%, 100%]), the reverse was not true. Children who were deictic-status-knowers had only a 46% chance of being a remoteness-knower ([35%, 58%]); if they were order-knowers, they had only a 52% chance ([39%, 64%]). This is exemplified by the top timeline in Fig. 1b, produced by a 5-year-old who appears to understand both deictic status and order, but does not exhibit adult-like knowledge of remoteness. Compare this to the bottom timeline, from an adult, where items' placements reflect not just their deictic status and order, but also their relative remoteness (i.e., "yesterday" is close to now; "next week" is farther away; and "next year" even farther). Together with the cross-sample age-of-acquisition data described in section 3.2, these results reveal a clear developmental trajectory in which deictic status and order emerge early and in synchrony, while knowledge of temporal remoteness is developed independently and often much later.

3.4 Forced-choice questions

Because timelines are complex spatial artifacts, young children's ability to use them could be affected by factors other than their semantic knowledge of the temporal items, e.g., their spatial reasoning or motor skills. To validate our timeline measure, we investigated the relation between children's ability to express their temporal knowledge using the timeline and a purely verbal measure of time word knowledge: i.e., children's answers to verbal forced-choice questions about the relative ordering of time words and their relations to events (Table 2). Overall performance on the Verbal Forced Choice task [5] improved with age throughout childhood, $b = .09$, $p < .001$, increasing monotonically ($M_3 = 0.52 < M_4 = 0.71 < M_5 = 0.83 < M_6 = 0.93 < M_7 = 0.96 < M_8 = 0.97 < M_{\text{adults}} = 0.98$). Four-year-olds performed significantly above chance, $t_{19} = 3.5$, $p < .01$, but 3-year-olds did not, $t_{16} = 0.4$, $p = 0.70$. While 6-year-olds performed significantly differently from adults, $t_{62} = -2.0$, $p < .05$, 7-year-olds were indistinguishable from adults, $t_{55} = -1.24$, $p > .2$.

Critically, the results of our non-spatial task provide evidence that the timeline task is a valid measure of semantic knowledge. Measures of each of the three facets of deictic time word meaning from the timeline (i.e., deictic status, order, and remoteness) were predictive of children's performance on the verbal forced-choice task, ($\beta_{\text{deictic}} = 0.12$, $p < .001$; $\beta_{\text{order}} = 0.13$, $p < .001$; $\beta_{\text{remoteness}} = 0.09$, $p < .001$). Even after controlling for age, knowledge of deictic status and order remained reliable predictors of verbal forced-choice performance ($\beta_{\text{deictic}} = 0.19$, $p = .01$; $\beta_{\text{order}} = 0.14$, $p = .02$), although knowledge of remoteness was only a marginally significant predictor ($\beta_{\text{remoteness}} = 0.18$, $p = .06$). When all three measures of timeline performance were included in a multiple

---

[5] Data from two forced-choice questions about the relative ordering of past items were excluded due to ambiguity in the instructions. These ambiguous instructions produced two distinct response patterns in both children and adults: some participants interpreted the "first event" to be the one closest to the present moment, while others interpreted it to be the event that was more distant in the past.

regression model of performance on the forced-choice task, knowledge of deictic status ($\beta_{deictic} = 0.08$, $p < .01$) and order ($\beta_{order} = 0.06$, $p = .04$) each accounted for variance in forced-choice accuracy, but knowledge of remoteness did not ($\beta_{remoteness} = 0.01$, $p = .59$). This finding—that deictic status and order knowledge account for unique variance in forced choice accuracy—confirms that, despite the very similar developmental trajectories we observed for these facets of meaning, the two measures themselves are distinct.

4. Discussion

Our study investigated what children know about deictic time words during the long delay between when they begin producing these words, around age 3, and eventual adult-like usage in elementary school. Specifically, we asked whether children assign systematic, preliminary meanings to these words during this period, by using a timeline measure to independently assess children's knowledge of three facets of their meanings: deictic status, order, and remoteness. Consistent with previous studies (e.g., Ames, 1946; Busby Grant, & Suddendorf, 2011; Harner, 1975, 1981; Weist et al., 1991), we found that learning deictic time words is a slow and difficult process for children. However, we also found evidence that many of children's early "errors" in speech may belie a nascent, though partial, understanding of these words. Our findings suggest that children assign systematic, partial meanings to deictic time words even at ages when they are said to use them in "incorrect" ways.

Together, the results of this study reveal a trajectory of deictic time-word learning that spans four or more years from the time that children begin using the words in speech. Although children struggle to map deictic time words to experienced or anticipated events until at least age 5 (Busby & Suddendorf, 2005; Busby & Suddendorf, 2010; Suddendorf, 2010), our timeline measures revealed that some 4-year-olds understood these words' deictic status (e.g., that "yesterday" is in the past) and order (e.g., that "yesterday" is before "this morning"). Further, children who understood one of these facets of meaning almost always understood the other, suggesting that they emerge in synchrony. However, in contrast to the relatively early emergence of deictic status and order knowledge, we found that adult-like knowledge of remoteness (e.g., how much further away from the present "next week" is than "tomorrow") emerges nearly two years later, after age 7.

4.1 Implications for how deictic time words are learned

As noted in the Introduction, children could in principle rely on event memories to make inferences about the meanings of deictic time words that refer to those events. An event mapping strategy most plausibly predicts a trajectory in which knowledge of deictic status and remoteness emerge in tandem, *followed* by knowledge of order. This is not what we found, as children rarely demonstrated adult-like knowledge of remoteness prior to demonstrating mastery of deictic status and order (Fig. 4B). Instead, our finding that children understand deictic status and order before remoteness is consistent with accounts of word learning in which children initially draw on the linguistic context to constrain their early hypotheses about word meanings. For example, children may learn deictic status early in life by using their knowledge of tense markings to make inferences about the past or future status of time words (e.g., "He **danced** *last year"*; "He **will dance** *tomorrow*"), a process known as syntactic bootstrapping (Brown, 1957; Gleitman, 1990; Gleitman et al., 2005; Landau & Gleitman, 1985; Naigles, 1996). Meanwhile, to acquire knowledge of order, children might learn from contrastive uses of these words ("The package isn't coming *tomorrow*. It's coming later, *next week.*"), or exploit the fact that order-of-mention in discourse typically respects temporal order ("We will go to school *tomorrow* and can go to the zoo *next week*"; Jakobson, 1966). Discourse structure and tense marking, in contrast, are unlikely to provide children with information about remoteness, which could help explain its late emergence in our study.

Although our findings are consistent with an account in which children exploit linguistic cues to learn deictic time word meanings, they do not provide direct evidence for such an account. Also, this account leaves open how children acquire the semantics of tense in the first place—and what types of conceptual representations this might draw upon. One way to explore both of these questions is to investigate how tense markers and deictic words are acquired in other languages. For example, cross-linguistic differences in temporal morphology suggest that a reliance on tense marking will be differentially viable across languages. Languages like Inuktitut (spoken in Alaska and northern Canada) and Zulu (spoken in South Africa) have a metrical tense system that encodes different degrees of remoteness in the past and/or future in addition to deictic status (Comrie, 1985; Chung & Timberlake, 1985; Dahl, 1983), while other languages, like Mandarin Chinese, lack a morphological tense system altogether (Comrie, 1985). This type of variation presents the possibility of asking whether tense marking plays a causal role in deictic word learning, and also whether the late acquisition of remoteness is restricted to deictic terms, or also found in the acquisition of tense systems, too (see Swift, 2004, for discussion).

Regardless of what informational sources support learning about deictic status and order in English, our findings suggest that these two facets of knowledge are tightly linked. In particular, although our timeline method could have documented distinct developmental trajectories for deictic status and order, we found that they appear to arise simultaneously, around age 4, and at the same developmental moment within individual children. In practice, an understanding of deictic status may scaffold understanding of order, and vice versa. For example, if a child understands that "tomorrow" is in the future (deictic status) and that "next week" is after tomorrow (order), she might infer that "next week" must also be in the future (deictic status). Conversely, if a child knows that "yesterday" is in the past (deictic status) and "tomorrow" is in the future (deictic status), she can then infer that "tomorrow" must be after yesterday (order).

To learn about remoteness, children might require particularly structured and explicit information—e.g., "Next week is seven days away"—which they may only receive in formal educational settings. Consistent with this, one previous study suggests that children may have to be taught explicit definitions to work out the meanings of duration words like "week" and "year," which are components of deictic terms like "next week" and "last year" (Tillman & Barner, 2015). Interestingly, the acquisition of duration words follows a similar trajectory to the one presented here for deictic time words. At age 4, children grasp that duration terms indicate durations (Shatz et al., 2010); at age 5, they have begun to work out their order (i.e., *year* > *week* > *hour* > *minute*; Tillman & Barner, 2015); but children have little understanding of their proportional relations (e.g., how much more time an hour is than a minute) until early in grade school. Further, the transition to mature knowledge of duration words, and the remoteness of deictic time words, occurs around the same time that clock reading becomes a major focus in standard elementary school curricula, in Grade 2 (Common Core State Standards Initiative, 2012). These considerations suggest that knowledge of the metric structure of time—both for duration and for deictic time—may require exposure to explicit definitions, and knowledge of one set of definitions (e.g., 1 year = 52 weeks) may support learning about the other (e.g., next year is up to 52 times more remote than next week).

4.2 Implications for abstract word learning

Our findings bear similarity to several other case studies of word learning, including the acquisition of color words (e.g., Backscheider & Shatz, 1993; Sandhofer & Smith, 1999; Wagner, Dobkins & Barner, 2013), emotion words (e.g., Widen & Russell, 2003), and duration words (e.g., Shatz et al, 2010; Tillman & Barner, 2015). In each of these cases, there is a gap between children's initial production of a set of words and their eventual adult-like usage of these words. Further, within this gap, children do not use words in haphazard ways, but instead in systematic ways that reflect their partial meanings for these words.

By some accounts, such findings reveal a bootstrapping process through which children build the meanings of words not simply by understanding how they relate to the world, but also by understanding how they relate to other words (Carey, 2009). For example, children may begin acquisition for a particular lexical domain by grouping words from the domain together (e.g., by observing that these words appear in similar distributional profiles; Tare et al., 2009). These words may initially serve as placeholders with little semantic content, but might gain content as children directly learn about how each word relates to the others. As children begin to relate some of these words to experience, they may use such information to enrich the meanings of other words within the same placeholder structure. Number word learning provides an illustrative example: Children appear to learn the count list (“one, two, three, …”) before they learn exact meanings for any of the individual number words or can reliably map them to approximate numerical magnitudes (Wynn, 1990; 1992; Davidson et al., 2012; Le Corre & Carey, 2007). Later in acquisition, children can use the structure of the count list to learn the successor principle and bootstrap richer, exact meanings (e.g., such that each successive number in the list denotes a set with one more individual in it; see Carey, 2009 for discussion).

The bootstrapping account described above may provide a useful framework for understanding the acquisition of deictic time words. Children may initially group these words into a common semantic class because they appear in similar discursive contexts (e.g., that “yesterday”, “last week”, and “last year” are used in response to “When did that happen?”), and may then learn about their contrastive relations to one another by leveraging cues from the linguistic context (e.g., from morphosyntax and discourse structure). Finally, as children gradually link some of these words to experience, they could bootstrap richer meanings for other words.

4.3 Spatial tools for time

Our study used spatial timelines as a tool to characterize children’s knowledge of deictic time words, and to track the development of different facets of time-word meaning. Although this method requires children to use an external spatial representation, we showed that it is sensitive to children’s semantic knowledge, since children’s ability to use the timeline predicted their ability to answer verbal questions about time words. By allowing us to independently assess knowledge of deictic status, order, and remoteness, the timeline method was more sensitive than previous studies to the possibility that children could have partial meanings for deictic time words.

The question of how readily children can use timelines is interesting in its own right. Although timelines are now widespread and commonplace, their use and comprehension has been relatively understudied. In particular, there has been little empirical research on effectiveness of timelines as learning aids for children, despite their prevalence in elementary school textbooks (Burny et al., 2008). Precursors of an ability to use timelines are well-documented: Even pre-linguistic infants associate spatial length and temporal duration (de Hevia et al, 2014; Srinivasan & Carey, 2010; Winter, Marghetis, & Matlock, 2015), and, by the time children enter kindergarten, they overwhelmingly use conventional linear arrangements to represent the order of temporal events (e.g., left-to-right for English-speakers; Tversky et al, 1991).

Building on these previous findings, our studies indicate that children as young as 4 can represent the temporal locations of both events and time words using a canonical timeline – i.e., one that is horizontally-oriented, bidirectional (encompassing the past and future), and continuous. To our knowledge, this is the most sophisticated use of a canonical timeline by preschoolers yet recorded. Previous studies have relied primarily on modified, non-standard timelines, which were also less sensitive to partial knowledge. For example, in one set of studies, time was represented by a road taking a diagonal, upward path, with distance conveyed using visual perspective (Friedman 2000; 2002), and in another, children placed items along ruler-like wooden boards extending away from the body into either the past or future (Friedman & Kemp, 1989; Friedman, 2000; Busby Grant & Suddendorf, 2009). Our findings have immediate implications for pedagogy, and indicate that, as early as preschool, timelines can be used both as a learning aid and as a measure of children's comprehension of temporal words and concepts.

4.4 Conclusion

Acquiring mature meanings for deictic time words can take children four years or more. Our findings suggest that, during the long delay between children's initial production of these words and eventual adult-like usage, children construct systematic, partial meanings, including information about deictic status and order. One intriguing possibility is that these partial meanings are built through a gradual inductive process in which children construct an ordered, semantic domain for these words based on cues found in natural language, in the structure of both morphosyntax and discourse. Ultimately, to learn deictic time words, children must make the insight that there is an invisible and highly-structured dimension of time, which can be described by a rich system of linguistic labels, but is nonetheless separable from the events that occupy it. This represents a profound conceptual breakthrough, providing a transformational framework for organizing events, interpreting the past, and planning for the future.

Acknowledgements

We thank Sara Al Mughairy, Katherine Kimura, Miriam Rubenson, Karen Bejar, Sean Burdeaux, Kaitlyn Peck, Nestor Tulagan, Tilly Sirota, and Eren Fukuda for help with subject recruitment, data collection, and data entry. We are most grateful to the daycares and museums where data were collected, including the Reuben H. Fleet Science Center in San Diego, and to all the children and families who participated in this research. Thanks to Laura Wagner and two anonymous reviewers for constructive comments which improved the manuscript. This work was funded by a UC Faculty Startup award to M.S. and a grant to D.B. from the James S. McDonnell Foundation.

References

Ames, L. B. (1946). The development of the sense of time in the young child. Journal of Genetic Psychology, 68, 97–125.

Backscheider, A. G. & Shatz, M. (1993). Children's acquisition of the lexical domain of color. In K. Beals et al. (Eds), What we think, what we mean, and how we say it (Papers from the parasession on the correspondence of conceptual, semantic, and grammatical representation, CLS 29), Vol. 2 (pp. 11–21). Chicago: The Chicago Linguistic Society.

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development*, *20*(3), 362-372.

Busby, J., & Suddendorf, T. (2010). Young children's ability to distinguish past and future changes in physical and mental states. *British Journal of Developmental Psychology*, *28*(4), 853-870.

Busby Grant, J., & Suddendorf, T. (2009). Preschoolers begin to differentiate the times of events from throughout the lifespan. *European Journal of Developmental Psychology*, *6*(6), 746-762.

Busby Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly*, *26*(1), 87-95.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Chung, S., & Timberlake, A. (1985). Tense, aspect, and mood. Language typology and syntactic description: Grammatical categories and the lexicon, Vol. by Timothy Shopen, 202-258.

Common Core State Standards Initiative (2012). *Common core state standards for mathematics*.

Comrie, B. (Ed.). (1985). *Tense* (Vol. 17). Cambridge University Press.

Dahl, Östen. 1985. Tense and Aspect Systems. Oxford: Blackwell.

De Villiers, J. G., & De Villiers, P. A. (1978). *Language acquisition*. Harvard University Press.

Dahl, Ö. (1983). Temporal distance: Remoteness distinctions in tense-aspect systems. *Linguistics*, *21*(1), 105-122.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children.

Behavior Research Methods, Instruments, & Computers, 28(1), 125–127.

Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? Cognition, 123(1), 162–173.

Fillmore, C. J. (1997). Lectures on deixis. Stanford, CA: CSLI. (Original work published 1977)

Friedman, W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Journal of Cognition and Development*, *3*(3), 333-356.

Friedman, W. J. (2003). The development of a differentiated sense of the past and the future. *Advances in child development and behavior*, *31*, 229-269.

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development*, *13*(3), 335-367.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135-176

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, *1*(1), 3-55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. Language Learning and Development, 1, 23–64.

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. *Developmental Psychology*, 11(6), 864–865

Harner, L. (1976). Children's understanding of linguistic reference to past and future. *Journal of Psycholinguistic Research*, *5*(1), 65-84.

Harner, L. (1981). Children's understanding of time. Topics in language disorders, 2(1),51-66.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, 111, 4809-4813.

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. Psychological Review, 77, 223-233.

Hudson, J. A., & Mayhew, E. M. (2011). Children's temporal judgments for autobiographical past and future events. *Cognitive Development*, *26*(4), 331-342.

Jakobson, R. (1966). Question for the essence of language. *Diogenes*, 51, 21-37.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. Journal of Experimental Psychology: General, 117, 371–376.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. Cognition, 105(2), 395–438.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, *92*, 199-227.

Naigles, L. (1996). The use of multiple frames in verb leaning via syntactic bootstrapping. Cognition, 58, 222-251.

Nelson, K. (1998). Language in cognitive development: The emergence of the mediated mind. Cambridge University Press.

Núñez, R. E., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in Cognitive Sciences*, 17, 220–229.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. Developmental Psychology, 35(3), 668.

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge in early childhood. Journal of Experimental Child Psychology, 103, 325–337.

Sauter, A., Uttal, D., Alman. A. S., Goldin-Meadow, S., & Levine, S., C. Learning what children know about space from looking at their hands: The added value of gesture in spatial communication. *Journal of Experimental Child Psychology,* 2012, *111(4)*, 587-606

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*,*8*(9), 657-661.

Shatz, M. (1993, July). Ways to words: Early strategies for building a lexicon. Paper presented at the Sixth International Congress for the Study of Child Language, Trieste, Italy.

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development*,*11*(1), 16-36.

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

Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. *Weaving a lexicon*, 257-294.

Srinivasan, M., & Carey, S. (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition*, 116, 217-241.

Suddendorf, T. (2010). Linking yesterday and tomorrow: Preschoolers' ability to report temporally displaced events. *British Journal of Developmental Psychology*, *28*(2), 491-498.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans?. *Behavioral and Brain Sciences*, *30*(03), 299-313.

Swift, M. D. (2004). *Time in child Inuktitut: a developmental study of an Eskimo-Aleut language* (Vol. 24). Berlin: Walter de Gruyter.

Tare, M., Shatz, M., & Gilbertson, L. (2008). Maternal uses of non-object terms in child-directed speech: Color, number and time. *First Language*, *28*(1), 87-100.

Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children’s acquisition of duration words. *Cognitive psychology*, *78*, 57-77.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127*(3), 307-317.

Weist, R. M. (1989). Time concepts in language and thought: Filling the Piagetian void between two to five years. In I. Levin & D. Zakay (Eds.), Time and human cognition: A life-span perspective (pp. 63–118). Amsterdam: North-Holland.

Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of child language*, *18*(01), 67-92.

Widen, S., & Russell, J. (2003). A closer look at preschoolers’ freely produced labels for facial expressions. Developmental Psychology, 39, 114–128.

Winter, B, Marghetis, T., & Matlock, T. (2015). Of metaphors and magnitudes: Explaining cognitive interactions between space, time, and number. *Cortex,* 64, 209-224.

Wynn, K. (1990). Children’s understanding of counting. Cognition, 36(2), 155–193.

Wynn, K. (1992). Children’s acquisition of the number words and the counting system. Cognitive Psychology, 24(2), 220–251.

Chapter 2 Appendix A: Full procedure scripts for Timeline task

The paper and colored pencils were placed in front of the child. The experimenter (E) explained the task by pointed out the top timeline and stating: “Look, this is a timeline. It shows *when* different things happen. The line starts in the past [E points to left endpoint] and it goes to the future [traces the line with her finger, ending on the right endpoint]. So, it goes from when you were a baby [E to points to left endpoint] all the way to when you’re going to be a grown up [E gestures along line to right endpoint]. And here in the middle is right now [E points to vertical line at midpoint]. Each time has its own place on the line. You're going to show me *when* different things happen by showing me *where* they go on the line. Look, when you were a baby goes here [E draws a vertical line on the left end point to demonstrate the procedure] and when you are going to be a grown up goes here [E draws a vertical line at right endpoint]. And right now goes here [E draws line at midpoint]. I’m going to give you a pencil, and your job will be to draw an up-and-down line to show me where each thing goes. Ready?”

At this point, the experimenter introduced the first target item, “When [did you] [eat breakfast today]? Think about when you [ate breakfast today]? Draw a line for when you [ate breakfast today].” Each item was associated with a particular color of pencil. The child marked the line, as shown in Figure 1, and the experimenter proceeded to the next trial. The items relating to the child’s first and last birthday used the phrasing “draw a line for when you [turned/are going to turn] [child’s age-1/child’s age+1].” In a few cases children were unwilling to mark the timeline, but were willing to point to where they thought the target should go. In these cases the experimenter drew the lines herself. This procedure was repeated for the remaining two timelines, with trials in the form, “Now you’re going to show me where [last week] goes. Where does [last week] go? Can you draw a line for [last week]?” Participants always received the Events line first (to orient them to the task), but the order of the other two lines was counterbalanced between subjects. For each line, half of the subjects received the items in the order shown in Table 1, and the other half received the reverse order. If children asked questions about where they should put the lines, they were told, “It’s up to you! Put it wherever you think [tomorrow] goes on the line.” If participants asked to be reminded what their marks from prior trials referred to, the experimenter provided the answer.

Chapter 2 Appendix B: Item Effects

Item effects for Deictic Status were assessed with a mixed ANOVA, with Age as a between-subjects variable and Time Word (e.g., “yesterday”, “next year”, etc.) as a within-subjects variable. The main effect of Time Word was not significant, *F*(7, 1022) = 1.6, *p* = 0.1), though there was a marginal interaction with Age, *F*(42, 1022) = 1.3, *p* = 0.09. Table S1 shows the mean Deictic Status accuracy for each time word for each age group.

Table 2.3. Deictic status accuracy for 8 temporal terms. Percentage of participants at each age (in years) who correctly assigned past items to the left or future items to the right of “now.”

Age	Percentage of participants demonstrating correct deictic status (SEM)							
	Last year	Last week	Yesterday	This morning	Tonight	Tomorrow	Next week	Next year
3	71 (11)	59 (12)	41 (12)	59 (12)	47 (12)	29 (11)	53 (12)	35 (12)
4	60 (11)	70 (11)	45 (11)	80 (9)	70 (11)	75 (9)	40 (11)	65 (11)
5	82 (10)	76 (11)	59 (12)	71 (11)	88 (9)	76 (11)	82 (10)	82 (10)
6	95 (7)	85 (8)	83 (8)	92 (6)	88 (7)	88 (7)	79 (8)	75 (9)
7	95 (5)	85 (8)	90 (7)	100 (0)	95 (5)	90 (7)	90 (7)	95 (5)
8	94 (6)	94 (6)	89 (8)	89 (8)	100 (0)	94 (6)	89 (8)	94 (6)
Adult	97 (3)	100 (0)	100 (0)	97 (3)	92 (4)	100 (0)	100 (0)	100 (0)

Chapter 2 Appendix C: Confirmatory analyses of order knowledge

To confirm our findings related to order knowledge, we constructed two additional measures. The first was a whole-timeline measure that looked at ordering errors across the timeline as a whole, not just errors within the past and future in isolation. The second assessed knowledge of relative sequential order, independent of deictic status, by considering the past and future separately.

We assessed the total order error for each timeline as follows: For each item from that timeline, we calculated the absolute deviation between its correct rank (i.e., 1st, 2nd, etc.) and its actual position along the timeline, and then summed across all items. For example, on the Event line, if the mark for "next birthday" – which should have been the *fourth and final* item – was actually the *first* mark on the timeline, this would contribute $|1-4| = 3$ to the total order error for that timeline. A perfectly ordered timeline would thus receive a score of 0, while a maximally disordered timeline would receive a score of 8 (chance performance was 5). For each participant, we also calculated the mean Order error for each type of timeline (i.e., Deictic vs. Event).

Order error was analyzed with a mixed ANOVA, with Timeline Type (i.e., Deictic *vs.* Event) as a within-subjects factor and Age (3 through 8 years old, and adults) as a between-subjects factor. The only effect to approach significance was the main effect of Age, $F(6, 146) = 31.8$, $p < .001$ ($p > .19$ for all other effects). Indeed, performance improved monotonically with age, $M_3 = 4.7 > M_4 = 3.5 > M_5 = 2.4 > M_6 = 1.8 > M_7 = 0.9 > M_8 = 0.4 > M_{\text{adults}} = 0.1$, and, among children, a linear regression revealed that order error was predicted significantly by age, $b = -0.86 \pm 0.09$ SEM, $p < .001$, $r^2 = .45$, $p < .001$. Three-year-old children performed at chance, $M_3 = 4.7$, $t_{16} = -0.88$, $p = .39$. By contrast, 4-years-olds had significantly lower error than 3-year-olds, $t_{35} = 2.4$, $p = .02$, and their performance was significantly better than chance, $M_4 = 3.5$, $t_{19} = -4.0$, $p < .001$. Seven-year-olds still had significantly higher error scores than adults, $t_{54} = 3.0$, $p < .01$, but 8-year-olds were indistinguishable from adults, $t_{52} = 1.6$, $p = .13$.

We next assessed knowledge of relative order independently from deictic status, by evaluating the relative placement of the two items that were in the past, and, separately, the relative placement of the two items that were in the future. Thus, simply knowing the past status of ‘yesterday’ and the future status of ‘next year,’ for instance, was insufficient to succeed on this measure of relative order knowledge. Instead, if ‘yesterday’ were placed to the right of ‘last year,’ this would count as a correct ordering, regardless of where the items were placed relative to either the ‘now’ midpoint or the future items. This measure was thus designed to completely isolate order knowledge from knowledge of deictic status—at the cost of power, however, since it involves only two comparisons per timeline.

We calculated mean accuracy on this measure of order knowledge, for each participant and Timeline Type. There was only a main effect of Age, $F_{(6, 146)} = 11.95$, $p < 0.001$, with knowledge of order increasing monotonically with age, $M_3 = .57 < M_4 = .58 < M_5 = .69 < M_6 = .74 < M_7 = .86 < M_8 = .93$. Just as with deictic status, 3-year-olds were not above chance, $t_{16} = 0.68$, $p > .5$ — but neither were 4-year-olds, $t_{19} = 1.60$, $p = .13$. It was not until children were five years old that they were reliably above chance on this measure of order knowledge, $t_{17} = 3.29$, $p < .01$. Eight-year-olds were indistinguishable from adults, $M_{\text{adults}} = .98$, $t_{52} = -1.58$, $p = .12$, but 7-year-olds had significantly more error than adults, $t_{54} = -2.76$, $p < .01$.

Chapter 2, in full, is a reprint of the material as it appears in Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology, 92*, 87-100. Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier.

Chapter 3.

Constructing the mental timeline:
Flexibility and conventionality in children's spatial representations of time

Katharine A. Tillman[1], Nestor Tulagan[1,3], Eren Fukuda[1], & David Barner [1,2]

[1]Department of Psychology

[2]Department of Linguistics

University of California, San Diego

[3]School of Education

University of Calfornia, Irvine

Abstract

When reasoning about time, English-speaking adults often invoke a “mental timeline” stretching from left to right (LR). Although the direction of the timeline varies across cultures, linear representations of time have been argued to be ubiquitous and primitive. On this hypothesis, we might predict that children also spontaneously invoke a spatial timeline when reasoning about time. However, little is known about how and when the mental timeline develops, or to what extent it is variable and malleable in childhood. Here, we used a sticker placement task to test whether preschoolers and kindergarteners spontaneously produce linear representations of temporal events (breakfast, lunch, and dinner) and deictic time words (*yesterday*, *today*, *tomorrow*), and to what degree those representations are adult-like. We found that, at age 4, preschoolers were able to make linear mappings between time and space with minimal spatial priming. However, unlike kindergarteners and adults, most preschoolers did not adopt linear representations spontaneously, in absence of priming, and did not prefer LR over right-to-left lines. Furthermore, unlike most adults, all children could be easily primed to adopt an unconventional vertical timeline. Our findings suggest that mappings between time and space in children are initially flexible, and become increasingly automatic and conventionalized in the early school years.

1. Introduction

Although all humans have a sense of the passage of time, as events emerge and then fade into memory, the perception of time is both ephemeral and subjective.

However, precise coordination of activities across large groups of people requires a means of timekeeping that is available to all. To deal with this problem, many cultures have devised symbolic systems to describe and measure the passage of time (e.g., Barnett, 1998; Gell, 1992; Whitrow, 1989). These include a variety of spatial tools, such as clocks and calendars, and graphical representations of time such as charts and timelines. Space and time are also linked via linguistic descriptions ("a *long* time") and practices such as reading and writing, which associate the unfolding of narrative with the eye's progress along a spatial pathway. In many cultures, these phenomena appear to reflect a linear, spatial framework for time — a "mental timeline" that supports temporal cognition. However, while most studies report a strong spatial component to adults' understanding of time (see Bender & Beller, 2014; Bonato et al., 2012), there is substantial cross-cultural variability in how time is spatially represented (Bergen & Lau, 2012; Fuhrman & Boroditsky, 2010; Ouellet et al., 2015; Nachson, 1983; Tversky, Kugelmass, & Winter, 1991). From these observations, a central question has emerged: Are linear space-time mappings in the mind forged exclusively through the prolonged exposure to cultural practices? Or, alternatively, do cultural conventions emerge from pre-existing—possible innate—associations between space and time? Current theoretical accounts do not reach consensus about the origin and nature of space-time mappings (see Bender & Beller, 2014; Nunez & Cooperrider, 2013; Winter, Marghetis, & Matlock, 2015). Furthermore, the degree of malleability in space-time mappings—and the precise effects of cultural practices and artifacts in shaping cross-cultural differences—are difficult to pin down in adult populations. For this reason, we sought here to test the emergence of space-time mappings in children, and how they initially use space to represent sequential and deictic temporal relations.

In adult speakers of English there is a strong association between leftward space and past or earlier events and between rightward space and later or future events (for a review, see Bonato et al., 2012). Evidence of a horizontal, left-to-right (LR) “mental timeline” (MTL) has also been observed in speakers of many other modern Western languages that are also read and written from left-to-right, including Dutch (e.g., Casasanto & Bottini, 2014), French (e.g., Droit-Volet & Coull, 2015), German (e.g., Ishihara et al., 2008), and Spanish (e.g., Santiago et al., 2007). This linear, LR structuring of time can be observed in subjects’ overt spatial arrangements of items representing temporal events and concepts (e.g., Bergen & Lau, 2012; Fuhrman & Boroditsky, 2010; Maas & Russo, 2003), in implicit reaction time differences on temporal judgment tasks related to the spatial locations of the stimuli or response keys (Gevers et al., 2003, 2004; Ishihara et al., 2008; Santiago et al., 2007, 2010; Torralbo et al., 2006; Vallesi et al., 2014; Weger & Pratt, 2008), and in spontaneous behaviors such as gesturing while talking about time (Casasanto & Jasmin, 2012; Cooperrider & Nunez, 2009). English-speaking adults with brain lesions resulting in left hemi-spatial neglect appear to also neglect the “left side” of time (i.e., have impaired reasoning about the past; Saj et al., 2013). These findings have been argued to reflect the neural underpinnings of the LR mental timeline in adults.

Some researchers have argued that time-space associations are an innate feature of human cognition, and may suggest the existence of a general, inherited system of magnitude representation (Fabbri et al., 2012; Walsh, 2003). Supporting this hypothesis, evidence of an implicit mapping between spatial length and temporal duration has been found in prelinguistic infants, in some cases as early as the first weeks of life (e.g., de Hevia et al., 2014; Brannon et al., 2007; Laurenco & Longo, 2010; Srinivasan & Carey, 2010). Also suggesting that space-time mappings may be innate, polysemous space-time words and metaphors (e.g., *long* and *short* in English) as well as linear artifacts representing time are pervasive across cultures (Haspelmath, 1997; Grady, 1999). While some linguistic and cultural systems employ diverse spatial representations of time, including both linear and cyclical ones, it is rare for societies to lack linear representations of time entirely. Although some have claimed that a few small indigenous communities lack a linear concept of time, this remains controversial (e.g., Le Guen, Balam, & Ildefonsa, 2012; Sinha et al, 2011). Together, such findings are consistent with the idea that humans have a predisposition toward linear representations of time that is molded, rather than formed, by cultural practices and artifacts.

While most infant studies argue that there could be an innate association between spatial and temporal magnitudes (e.g., long lines and long sounds), others make the stronger claim that the ordinal structure of the MTL (or the related "mental number-line") and its common LR direction may also be biological defaults. On these accounts, "mental lines" are a product of neurological constraints, such as hemispheric asymmetry in the brain or low-level visuo-attentional biases (e.g., Chatterjee, Southwood, & Basilico, 1999; Chatterjee, 2001; Vicaro et al., 2007). Consistent with this, there is significant overlap in the brain regions subserving time perception and shifts of visuospatial attention, including the right parietal lobe and temporal-parietal junction (e.g., Corbetta & Shulman, 2002; Pun et al, 2010). LR associations between spatial position and numerical magnitude have also been attested in non-human animals such as newborn chicks, leading to the proposal that LR mappings in humans have early evolutionary origins (Rugani et al., 2015), though this finding is controversial. Critically, on any account suggesting that the directionality of the MTL is an innate default, LR representations should be present in all human children before they gain extensive exposure to cultural input.

However, the LR timeline is far from universal among adults, suggesting that, if there is a biologically determined default of LR, this can be overridden by externally-imposed cultural conventions (Maass & Russo, 2003). Rather than positing a single, innate representational system, some accounts of time-space relationships in the mind emphasize the role of culture and experience in promoting particular mappings between time and space (e.g., Lakoff and Johnson, 1980). Importantly, both the direction and the axis onto which time is mapped vary according to factors such as writing direction (Bergen & Lau, 2012; Fuhrman & Boroditsky, 2010; Ouellet et al., 2015; Nachson, 1983; Tversky, Kugelmass, & Winter, 1991), the manner in which space is encoded in language (e.g., Boroditsky et al., 2011; Boroditsky & Gaby, 2010; Brown, 2012; Dehaene et al., 1993), and particular linguistic space-time metaphors (Boroditsky, 2001; Lai & Boroditsky, 2011; Miles et al., 2011; Nunez & Sweetser, 2006). Though the LR timeline is the most studied, both right-to-left (RL) and top-to-bottom (TB) timelines are also known to be present in several linguistic and cultural groups (e.g., Bergen & Lau, 2012; Fuhrman & Boroditsky, 2010; Scott, 1989; Tversky, Kugelmass, & Winter, 1991). Further, increased variability in space-*number* mappings has been observed among individuals who are illiterate, suggesting that this may also be the case for space-time mappings (Bergen & Lau, 2012; Zebian, 2005). Even in adulthood, experiences such as reading mirror-reversed text can modify adults' associations between time and space, at least temporarily (Casasanto & Bottini, 2014). All of these findings suggest that learning to read and write, to comprehend temporal language, and to use spatial artifacts like calendars may have important effects on mature spatial representations of time.

Current theories of how space represents time in the mind have struggled to account for both infant and adult data (see Winter, Marghetis, & Matlock, 2015, for discussion). Understanding the development of space-time mappings may provide a unique window into their nature and origin. Although questions of how linear representations of time emerge or change with cultural exposure are fundamentally developmental, and research on children's temporal reasoning dates back to Piaget (1927/1969), the development of space-time associations between infancy and adulthood remains relatively mysterious. In Western culture, many of the practices and artifacts that have been argued to shape space-time mappings (e.g., reading/writing, clock and calendar use) are introduced during the early school years, suggesting that environmental influences on space-time mappings may emerge during this period. Past studies, which we review below, have investigated children's ability to use specific spatial scales to order temporal events, following instruction from adults (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman, 2000, 2002; Hudson & Mayhew, 2011; Tillman & Barner, 2015). However, these studies do not test whether children make linear space-time associations when a template is not provided. Other studies have tested whether spatial properties of stimuli, such as their length or position, influence children's estimates of their temporal duration (Casasanto, Fotakopoulou, & Boroditsky, 2010;

Droit-Volet & Coull, 2015; Piaget, 1927/1969; 1946/1970), but do not address the role of space in children's reasoning about temporal sequences. Few studies have tested whether children spontaneously represent temporal events and sequences linearly, with a culture-specific orientation and direction (Dobel, Diesendruck, & Bolte, 2007; Tversky, Kugelmass, & Winter, 1991), which is the topic of the present study.

If present, how might children's spatial representations of time differ from those of adults? One possibility is that, even if children are capable of making mappings between spatial position and temporal location early in development (e.g., because representations of space and time exhibit similar structure), they may not do so spontaneously until relevant cultural conventions are internalized. In this case, we might expect that, unlike adults, children require external prompting such as instruction, spatial priming, or explicit linear templates (e.g., timelines) in order to produce such space-time mappings. Another possibility is that children possess associations between space and time, but with some features that differ from those of adults. For example, children may initially possess a spatial model of time that is linear but does not yet have a specific, culture-dependent directionality (Casasanto & Bottini, 2014). In this case, we might expect to see evidence of linear representations of time in young children that are more variable and/or malleable than those of adults. The present study seeks to answer each of these questions: (1) Do children spontaneously deploy linear representations of time in the years before they receive extensive exposure to cultural practices linking time and space in conventionalized ways? (2) If present, to what extent are children's spatial representations of time adult-like, e.g., LR for English-speakers? and (3) Are children's representations of time more malleable than those of adults?

Suggesting that linear representations of time are intuitive for children, several prior studies have found that preschoolers can be trained to use spatial scales to express their knowledge of the temporal ordering of events (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman, 2000, 2002; Hudson & Mayhew, 2011; Tillman & Barner, 2015; Tillman et al., 2017). For instance, children demonstrate a rudimentary ability to differentiate the times of past autobiographical events using a scale extending outward from the body as early as 3 years of age (Busby Grant & Suddendorf, 2009), and using an LR timeline by 4 years of age (Tillman et al., 2017). Around 4 or 5 years of age, children can also differentiate the times of future events using spatial scales (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman 2000, 2002), and can use timelines to order conventional time words and expressions of time (e.g., “yesterday”, “next week”; Hudson & Mayhew, 2011; Tillman et al., 2017). Although the fact that such mappings can be produced by preschoolers is striking, these abilities are initially limited, and continue to improve until at least age 7 (Hudson & Mayhew, 2011; Tillman et al., 2017), suggesting that conventional, adult-like associations between space and time may nonetheless require cultural conditioning, and may rely in part on formal education. Critically, however, while success on timeline-type tasks shows that children are *capable* of mapping time to linear space, they do not show that children make such mappings by default. In particular, the timelines themselves imposed constraints on what type of mappings children could deploy, and experimenters also provided instructions or training on how to use them (i.e., which spatial direction should represent increasing time). These studies test neither whether implicit, spontaneous associations between time and space are present, nor whether such spatial representations, if present, are more variable or malleable than those of adults.

To our knowledge, only two previous studies have tested whether children produce left-to-right representations of time in the absence of a timeline, other linear artifact, or pre-training in the LR convention. One study (Dobel, Diesendruck, & Bolte, 2007) showed that while adults and literate children tended to represent the agent, object, and recipient in a verbally-described event congruently with their reading/writing direction, preschoolers did not show this bias. However, in this study participants only illustrated a single event, such as “The mother gives the boy a ball,” rather than a series of events over time. The most convincing demonstration that children might spontaneously represent the temporal relations between events linearly, in the same direction in which their language is written, comes from a study by Tversky, Kugelmass, and Winter (1991). In this study, children were asked to place stickers on paper to represent the relative positions of three events. For example, the experimenter placed a sticker in the center of the page to represent “lunch,” and the child placed two other stickers to represent “breakfast” and “dinner.” Critically, children could place the stickers anywhere they chose: No timeline or template was given and children were not told that the stickers should be arranged in a line. Despite this, over 80% of kindergarteners placed the stickers in an ordinal line. Further, 70% of English-speaking school-children across grades K-5 placed the stickers in order from left-to-right, while only 30% of Hebrew-speakers (who read from right-to-left) did so, suggesting that school-aged children are already sensitive to the conventions of spatial depictions of time in their cultures.

While this study provides the strongest evidence thus far for the existence of a horizontal MTL in children, it has several limitations. First, because the youngest children in the study, kindergarteners, were already producing adult-like linear representations at a high rate, and all were in school, it leaves open the question of when and how this tendency develops, and whether it depends on formal education or other cultural learning. Secondly, although overt instructions regarding how to arrange event-denoting stickers were not given, children did perform a spatial "warm-up" task using physical objects arranged in a horizontal line, which, as we show in the studies below, could have primed them to adopt horizontal representations of time they might not have otherwise used. Third, because the orientation of spatial priming (horizontal) matched that of cultural conventions, the study did not address whether unconventional representations can be induced in children, and if so, whether children are more flexible than adults. Therefore, it remains an open question whether children are equally able to form space-time associations along different axes and directions—e.g., left-to-right, top-to-bottom—and thus the degree to which such associations are (1) flexible, and (2) determined by cultural conventions rather than, e.g., default preferences that are independent of culture. Finally, because the items in Tversky and colleagues' (1991) task comprised highly familiar daily sequences of events, it remains unclear how broadly children can use space to represent temporal relations, e.g., if they can also represent past/future relationships between abstract time words.

In the present study, we adapted the method used by Tversky and colleagues (1991) to test whether preschoolers, kindergarteners, and adult controls spontaneously represent time linearly[6]. Further, we tested whether these representations are conventional (i.e., LR), and how flexibly children and adults use space to represent time. Participants placed stickers on a square card to depict the relative positions of temporal items: Events (breakfast, lunch, dinner) and time words (yesterday, today, tomorrow). In the first condition, to assess whether linear representations of time can be produced, we replicated the original paradigm in which children were exposed to minimal horizontal (LR) spatial priming prior to the test trials. Next, to test whether participants make linear representations spontaneously, in the second condition they received no spatial priming prior to performing the test trials. Finally, in the third condition, to test the flexibility of spatial representations of time, participants received unconventional vertical (top-to-bottom) priming.

2. Method

2.2 Participants

Because only participants who create linear arrangements of stickers could be included in our planned analyses of line direction, we set a criterion number of "line-makers" (20 for Meals trials, 20 for Days trials, in each condition) *a priori*, and continued data collection until we reached this number in each age group. Because the proportion of participants who put their stickers in ordered lines varied across age groups and conditions, the total sample sizes varied as well. As a result, a total of 152 4-year-old preschoolers (*M* age = 4.5 years, range 4.0 - 5.0 years), 88 kindergarteners (*M* age = 5.9

---

[6] Prior the present study, we conducted another study with a very similar method (Tillman, Tulagan, & Barner, 2015; *Proc. CogSci*). The present experiment yields the same pattern of findings as this pilot work, in a larger sample of children, using improved methodological details.

years; range 5.0 - 6.9 years), and 62 adults (*M* age = 20.8 years, range 18.2 - 32.4 years) were tested. This sample included 39 preschoolers, 24 kindergarteners, and 20 adults in the Horizontal Prime condition; 67 preschoolers, 34 kindergarteners, and 21 adults in the No Prime condition; and 46 preschoolers, 31 kindergarteners, and 21 adults in the Vertical Prime condition. In cases where an additional participant was tested inadvertently, all data were analyzed.

All participants spoke English as a primary language. Children were tested in preschools, elementary schools, museums, and in the laboratory. A subset of kindergarteners (*n* = 43, distributed across conditions) was recruited and tested in the Comox Valley (British Columbia), and all other participants were recruited and tested in the San Diego, CA, area. The study protocol was approved by the UC San Diego Institutional Review Board. Adults and parents of child subjects provided informed consent, and children provided verbal assent before the study began. Adults were members of the UCSD Psychology Department subject pool, and were awarded course credit for participation. Children received a small gift (e.g., a sticker) in thanks for their participation.

An additional 25 participants were excluded from analysis due to being outside the target age range (*n* = 3), not speaking English as a primary language (*n* = 4), having a second language with non-LR orthography (*n* = 15), developmental delay (*n* = 2), or clerical error (*n* = 1).

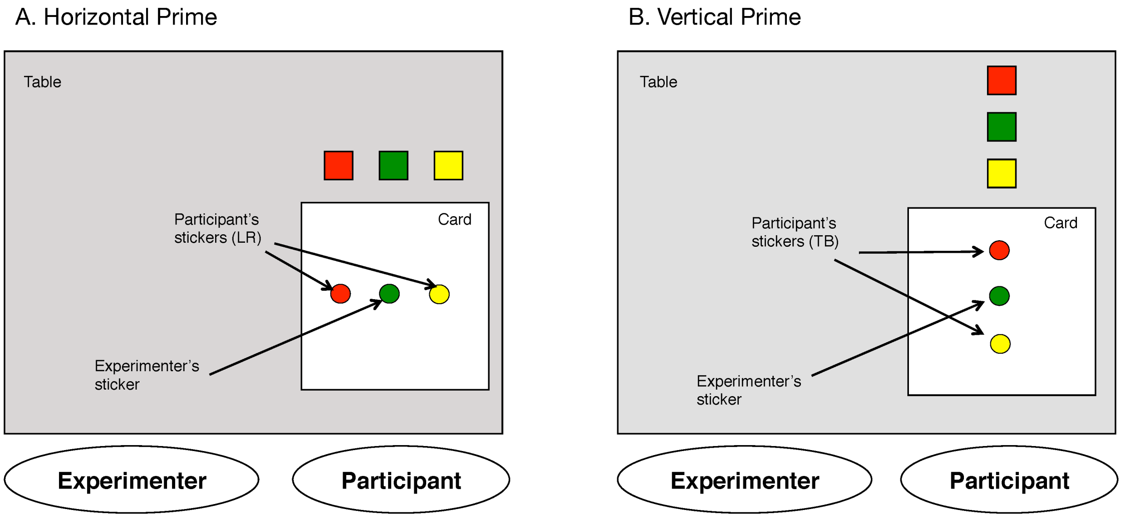

Figure 3.1. *Sticker placement task.* On each trial, the experimenter placed one sticker in the center of the card, and the participant sequentially placed two others. As shown, on spatial priming trials, stickers represented the positions of colored blocks placed on the table. On test trials (not shown), the blocks were removed and stickers represented (invisible) temporal items: Meals (breakfast, lunch, dinner) and Days (yesterday, today, tomorrow).

2.3 Procedure

Figure 1 shows a schematic of the sticker task, which was modeled on that used by Tversky, Kugelmass, and Winter (1991). To begin, a square white (6 in. × 6 in.) card was placed in front of the participant. The experimenter introduced the task by saying, "This is a sticker game! Every time we play the sticker game, I'm going to put one sticker on the card and you're going to put two stickers on the card. Listen carefully, because I'm going to tell you where to put the stickers. Let's start!"

2.3.1 Horizontal Prime condition.

Prior to the two test trials, participants in the Horizontal Prime condition completed a spatial priming trial, similar to the "warm-up" used by Tversky, Kugelmass, and Winter (1991). In this *blocks trial* (shown in Figure 3.1)**,** the experimenter placed a green (1in × 1in × 1in) wooden block on the table, approximately 1-in above the center of the top edge of the card, and asked, “What color is this?” After the participant responded (always correctly), the experimenter removed the block from the table. This was repeated for two other blocks, identical to the first except in color. The order of presentation of the three blocks was counterbalanced across participants. Next, the experimenter placed all three blocks (simultaneously) in a horizontal line parallel to the top edge of the card, approximately 1-in from the edge of the card, and with approximately 1-in spacing between blocks (see Fig. 3.1). She continued by saying, “Good job! I’m going to put a [green] sticker on the paper in the place of the [green] block, then I want you to put red and yellow stickers in the places of the red and yellow blocks.” She handed the first sticker to the child saying “I want you to put the [red] sticker in the place of the [red] block,” paused while the child placed the sticker, and then handed the child the other sticker. The blocks and the completed card were then removed from the child’s sight, and the test trials began.

On the *meals trial***,** the experimenter gave the following instructions: “I want you to think about the times of day we eat meals: breakfast, lunch, and dinner. I’m going to put a sticker down for *lunch* time, and I want you to put stickers down for [*dinner*] time and [*breakfast*] time. Here’s where I’m putting the sticker for *lunch* time.” The experimenter then placed a sticker in the center of the card and said, “Now you put a sticker down for *dinner* time…” The experimenter handed the participant a sticker and paused while the participant placed it on the card. The experimenter handed the child another, different-colored sticker, saying “…and another sticker down for *breakfast* time.” The order in which the two stickers were placed on the card was counterbalanced across participants. After the participant placed the second sticker, the experimenter took the completed card and replaced it with a new blank card. On the *days trial*, the experimenter began, “Now, I want you to think about the times when different days happen: yesterday, today, and tomorrow,” and the task proceeded identically to the first, using differently colored stickers for coding purposes. The order of the meals and the days trials was counterbalanced across participants.

2.3.2 No Prime condition. Procedures in the No Prime condition were identical to those of the Horizontal Prime condition, except that the blocks trial was omitted. After the game was introduced, participants completed only the meals and days trials, counterbalanced across subjects.

2.3.3 Vertical Prime condition. Procedures in the Vertical Prime condition were identical to those of the Horizontal Prime condition, except that the three blocks in the priming trial were organized in a vertical line, perpendicular to the top of the index card. Note that “vertical” blocks were not stacked on top of one another. All three were flat on the table, with the “top” block farthest from the participant, and bottom block nearest to her (see Fig. 3.1).

2.4 Coding

To facilitate offline coding, the experimenter drew a small mark on the back of each card to indicate its orientation during testing. Each target item (e.g., ‘breakfast’, ‘tomorrow’) was associated with a particular color of sticker. The ordinality and directionality of each sticker arrangement was coded as described below. Two independent coders achieved 98% inter-coder reliability, and a third coder resolved discrepancies.

2.4.1 Ordinality. Each sticker arrangement created by each participant was coded as *ordinal* if the largest angle the three stickers created was between 140 and 180 degrees, *and* the stickers were placed on opposite sides of the experimenter's central sticker, creating an ordinal temporal sequence along any axis. The leniency in acceptable angles was intended to avoid penalizing participants for immature motor control.

2.4.2 Directionality. Arrangements that met the criteria for ordinality were also coded for line direction. Any linear arrangement falling within 30 degrees of the cardinal directions — LR, RL, top-to-bottom (TB), or bottom-to-top (BT) — was assigned to that category. Ordinal arrangements that did not meet this requirement were classified as diagonal (D).

In cases where a participant created a line in only one of the two test trials, the nonlinear arrangement was not included in analyses of directionality, and an additional child was recruited. All participants completed both the meals and days trials.

2.5 Analysis software

Analyses were conducted using R (R Core Team, 2013) and the *lme4* package.

3. Results

The purpose of this experiment was to assess whether children produce linear representations of time, either: (1) following spatial priming consistent with conventional space-time associations in their culture, in the Horizontal Prime condition; (2) spontaneously, in the No Prime condition, or (3) with unconventional spatial priming, in the Vertical Prime condition.

3.1 Mixed-effects models

First, using mixed-effects logistic regression, we modeled the likelihood of producing an ordinal arrangement of stickers (in any direction). As main effects, we included Age Group (preschool, kindergarten, adult), Priming Condition (Horizontal, No Prime, Vertical), Trial Type (Meals, Days), and Trial Order (Meals first, Days first). Both Age Group, $X^2(2) = 26.0$, $p < 0.001$, and Priming Condition, $X^2(2) = 21.1$, $p < 0.001$, significantly improved the fit of the model, but neither Trial Type, $X^2(1) = 0.02$, $p = 0.9$, nor Trial Order, $X^2(1) = 0.21$, $p = 0.6$, did so. Separately, we modeled the likelihood of producing a conventional LR line. In this case, Age Group, $X^2(2) = 16.6$, $p < 0.001$, Priming Condition, $X^2(2) = 19.3$, $p < 0.001$, *and* Trial Type, $X^2(1) = 9.8$, $p = 0.002$, were significant predictors, though Trial Order was not, $X^2(1) = 1.4$, $p = 0.2$. This model indicated that LR arrangements were more frequent on Days trials than on Meals trials overall. Nonetheless, we did not detect significant differences in the proportion of LR lines of meals vs. days within conditions. Therefore these data were combined for ease of exposition below. The Supporting Information includes full frequency distributions of sticker arrangements for each trial type and additional details on modelling.

3.2 Horizontal Prime condition

To test whether participants would produce linear representations of time following a spatial prime, participants in the Horizontal Prime condition first used stickers to represent the relative positions of three blocks, which had been placed in the horizontal line by the experimenter. This task was similar to the “warm-up” used by Tversky, Krugelmass, and Winter (1991). On the horizontal blocks trial (i.e., the priming task), we found that all adults and kindergarteners created ordinal, LR arrangements.

Preschoolers created linear arrangements 92% of the time, and, of those, 86% were LR. Data from children who did not produce LR lines on the blocks trial were excluded from further analysis.

Next, we asked whether participants created ordered linear arrangements of stickers on the critical trials, in which stickers represented meals and days, after priming. Arrangements were classified as ordinal if the stickers were arranged in a line (> 140 deg.) that had events in the correct temporal sequence (e.g., breakfast, lunch, dinner), with any direction (e.g., LR, RL, TB; see Methods). As shown in Figure 3.2A, on the test trials, 100%, (95% CI [91%-100%]), of adults', 88% percent of kindergarteners' [75%-94%], and 73% of preschoolers' [60%-82%] arrangements were ordinal. The proportion of ordinal sticker arrangements on test trials increased marginally between preschool and kindergarten, $X^2(1) = 3.6$, $p = 0.06$, and significantly between kindergarten and adulthood, $X^2 (1) = 5.4$, $p = 0.02$.

Secondly, we asked whether ordinal arrangements also had a conventional LR direction. As shown in Figure 3.2B, we found that 90% (95% CI [76%-96%]) of adults' lines, 64% [49%-77%] of kindergarteners' lines and 64% [50%-77%] of preschoolers' lines were LR. Though kindergarteners made proportionally fewer LR lines than did adults, $X^2(1) = 7.6$, $p = 0.01$, there was no significant difference in the proportion of LR lines made by preschoolers and kindergarteners, $X^2(1) = 0.0002$, $p = 1$. As shown in Fig. 3.3, most lines created in the Horizontal Prime condition that were *not* LR were RL (20% of lines, across age groups), while the vertical axis was rarely used (6% of lines, across ages).

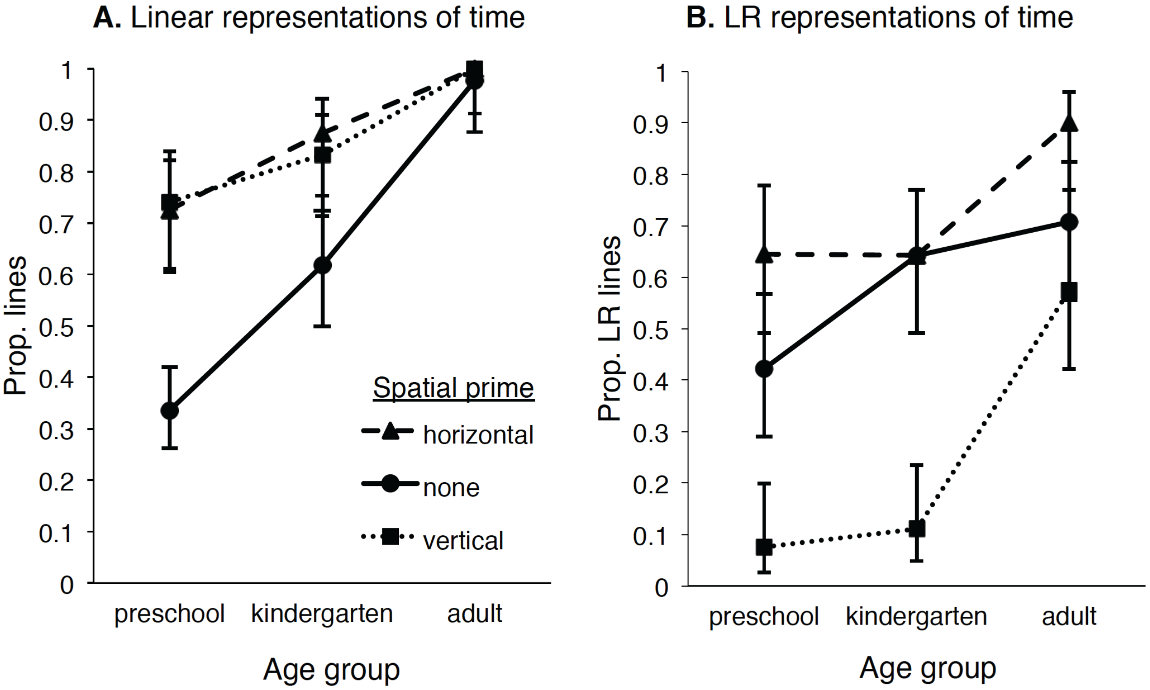

Figure 3.2. *Ordinal and left-to-right representations of time.* A. Proportion of participants in each age group who produced ordered linear arrangements of stickers representing temporal items, in the Horizontal Prime, No Prime, and Vertical Prime conditions. B. Proportion of linear arrangements in each condition that were ordered from left-to-right, consistent with conventional representations of time in the participants' culture. Error bars = 95% CI on the proportion.

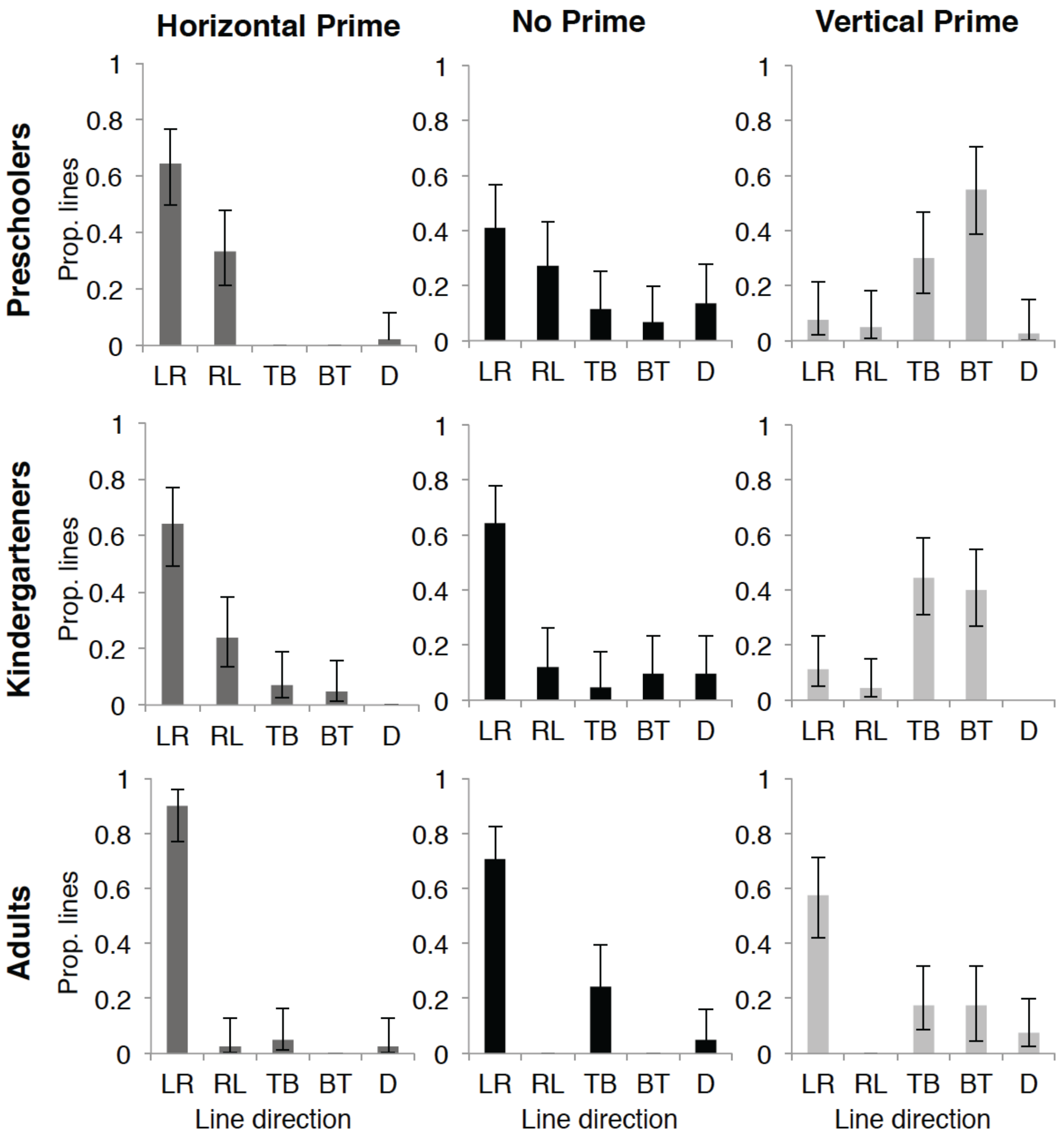

Figure 3.3. *Distributions of line directions.* Proportion of ordinal representations of time in each direction, for each age group (rows) in each priming condition (columns). LR = left-to-right, RL = right-to-left, TB = top-to-bottom, BT = bottom-to-top, D = diagonal. Error bars = 95% confidence intervals on the proportions.

Together, the results of the Horizontal Prime condition indicate that a majority of English-speaking children — including preschoolers — were able to make linear mappings between space and time with minimal priming, and most did so in a culturally-conventional manner. These findings indicate that linear, LR representations of time are relatively intuitive for children. They are compatible with either the possibility that LR mappings are a default, or that, though formal schooling is not required, other environmental cues may shape time-space mappings from an early age. However, it is also possible that the conventional representations of time children produced in this condition were a result of the priming they received.

3.3 No Prime condition

Next, to determine whether the horizontal priming provided by the blocks trial was necessary to induce linear representations of time, we removed it for participants in the No Prime condition. As shown in Fig 3.2A, without priming, 98% (95% CI [88%-100%]) of the sticker arrangements made by adults were ordinal, no fewer than in the Horizontal Prime condition, $X^2(1) = 1.0$, $p = 0.3$. Without a prime, 62% [49%-72%] of kindergarteners' arrangements were ordered lines. Thus we found that most of the kindergarteners created lines even without priming. However, this was proportionally fewer lines than adults created in the No Prime condition, $X^2(1) = 18.0$, $p < 0.001$, and fewer lines than kindergarteners created with the Horizontal Prime, $X^2(1) = 9.3$, $p = 0.002$. Without a prime, only 34% [26%-42%] of preschoolers' arrangements were ordered lines, a much lower percentage than we observed in kindergartners, $X^2(1) = 14.6$, $p < 0.001$. Relative to those in the Horizontal Prime condition, less than half as many preschoolers created lines without a prime, $X^2(1) = 26.0$, $p < 0.001$. This result suggests that although most preschoolers are *capable* of making linear mappings between time and space (i.e., with a spatial prime), they are unlikely to do so spontaneously.

Next, we asked whether participants created conventional LR lines without a spatial prime. As shown in Figure 3.2B, fewer LR lines were produced in the No Prime than in the Horizontal Prime condition. Without priming, 71% (95% CI [56%-82%]) of adults',[7] 64% [49%-77%] of kindergarteners', and 42% [29%-57%] of preschoolers' lines were LR. The difference in the proportion of LR lines between preschool and kindergarten was significant, $X^2(1) = 4.2$, $p = 0.02$, but that between kindergarteners and adults was not, $X^2(1) = 0.39$, $p = 0.53$. Further, as shown in Figure 3.3, the preference for LR over RL lines was highly significant in kindergarteners, $X^2(1) = 24.4$, $p < 0.001$, but did not reach significance in preschoolers, $X^2(1) = 2.4$, $p = 0.12$. Although the vertical axis was very rarely used to represent time after Horizontal priming, without the prime, 19% of lines (across ages) were vertical rather than horizontal.

Together, the findings of the No Prime condition suggest that both the tendency to create ordinal representations of time spontaneously and the conventionalized directionality of those lines is shaped in early childhood, inconsistent with the view that these behaviors are the result of innate mappings between time and space.

3.4 Vertical Prime condition

Because the horizontal prime was consistent with the dominant space-time mapping used by English-speaking adults, that manipulation did not address the degree to which space-time mappings are flexible in children (or in adults). In the Vertical Prime

---

[7] This relatively low proportion was driven by the meals trials. As shown in the Supporting Information, 86% of adults created LR representations of days.

condition, we probed this question asking whether priming participants with vertical block arrangements, placed perpendicularly to the top of the paper (Fig. 3.1), makes them more likely to represent time vertically. On this vertical priming trial, 95% of adults, 90% of kindergarteners, and 76% of preschoolers placed their stickers representing blocks in a TB line. Data from participants who did not do so were excluded from further analysis.

On the critical time trials, 100% of adults' (95% CI [91%-100%]), 83% of kindergarteners' [71%-91%], and 74% of preschoolers' arrangements [61%-84%] were ordinal (Fig. 2A, dotted lines). Virtually all adults in the Horizontal, No Prime, and Vertical Prime conditions produced lines, and there were therefore no differences between conditions, all $X^2(1) < 1.0$, all $p > 0.3$. The difference between kindergarteners and adults in the Vertical Prime condition was significant, $X^2(1) = 7.4$, $p < 0.01$, while that between preschoolers and kindergarteners was not, $X^2(1) = 1.4$, $p = 0.24$. Both the preschoolers and kindergarteners who received vertical priming were more likely to produce ordinal lines as those in the No Prime condition (for preschoolers, $X^2(1) = 25.4$, $p < 0.001$; for kindergarteners, $X^2(1) = 6.8$, $p < 0.01$; Fig 2A, dotted vs. solid lines). However, preschoolers and kindergarteners who in the Vertical Prime condition were equally likely to produce lines as those in the Horizontal Prime condition (for preschoolers, $X^2(1) = 0.03$, $p = 0.9$; for kindergarteners, $X^2(1) = 0.35$, $p = 0.6$, Fig. 3.2A dotted vs. dashed lines). Thus both types of linear spatial priming prompted children to also represent time linearly, and they were equally effective.

Next, we asked whether changing the orientation of the prime also changed the orientation of participants' spatial representations of time. We found that, after vertical priming, 58% (95% CI [42%-71%]) of lines created by adults were LR. The percentage of LR lines made by adults was lower in the Vertical Prime condition than in the Horizontal Prime condition, $X^2(1) = 10.9$, $p > 0.001$, but did not differ significantly from that in the No Prime condition, $X^2(1) = 1.5$, $p = 0.21$ (Fig. 3.2B). In contrast to adults, who often made conventional LR arrangements even after vertical priming, a mere 11% of kindergarteners' lines [5%-23%] and 8% of preschoolers' lines [3%-19%] were LR. These proportions of LR arrangements did not differ between kindergarteners and preschoolers, $X^2(1) = 0.3$, $p = 0.6$. In both groups, the percentage of LR lines in this condition was far lower than that observed in the Horizontal and in the No Prime conditions (all $X^2(1) > 12.5$, all p's < 0.001). Furthermore, after vertical priming, 85% of preschoolers' [71%-93%], 84% of kindergarteners' [71%-92%], and 35% [21%-52%] of adults' lines were vertical, inconsistent with cultural conventions (Fig. 3.3). Preschoolers were more likely to create BT (55%) than TB (30%) lines, $X^2(1) = 5.1$, $p = 0.02$,[8] while both kindergarteners and adults used the two directions equally often, both $X^2(1) < 0.2$, p's > 0.7.

The results of the Vertical Prime condition show that children are much more flexible than adults with respect to how they represent time in space. In fact, with appropriate priming, English-speaking children, unlike adults, were just as willing to organize temporal events vertically as horizontally, a result at odds with the idea that children have a strong default preference for LR mappings.

---

[8] This unexpected pattern is also consistent with the use of a *sagittal* timeline, in which time proceeds forward from the body, rather than downward from the top of the paper.

4. Discussion

We explored children's developing associations between time and space, and found that even preschoolers were capable of creating linear representations of time following modest spatial priming. Furthermore, with a horizontal prime, both preschoolers and kindergarteners represented time from LR, consistent with conventions in their culture. They did so without explicit instruction from adults or a visual template, such as a printed timeline. However, unlike kindergarteners and adults, most preschoolers did not *spontaneously* produce linear representations of time, despite having the ability to do so. Absent *any* spatial priming, linear representations of time in preschoolers were rare, and they were no more likely to produce LR as RL lines. We also found that children's linear representations of time were highly malleable. Unlike adults, both preschoolers and kindergarteners were just as willing to organize temporal events vertically as horizontally after being primed to do so. Taken together, these findings suggest that although preschoolers can represent time linearly, deployment of a "mental timeline" (MTL) when reasoning about time is initially neither automatic nor directionally biased. Although it remains possible that the ability to align space and time is innate, both the spontaneous deployment and the conventionalized direction of the mental timeline appear to be shaped by experience in early childhood.

The present study provides convergent evidence with cross-cultural studies suggesting that the MTL is a learned convention. Prior work with school-aged children and adults found that specific associations between space and time vary across cultures, reflecting corresponding differences in language, reading/writing direction, and artifacts (e.g., Bergen & Lau, 2012; De Sousa, 2012, Fuhrman & Boroditsky, 2010; Ouellet et al., 2015; Nachson, 1983; Tversky, Kugelmass, & Winter, 1991). For example, Tversky and colleagues (1991) showed that, while most English-speaking elementary school children produced LR representations of time, Arabic-speaking schoolchildren, who read from right-to-left, were more likely to create RL lines. Moving beyond these results, our findings in preschoolers suggest that cross-cultural differences are likely shaped during the early elementary school years, when children begin to receive increasing exposure to cultural conventions such as reading, writing, and calendar use, as we discuss further below. While we tested children from only one language group here, ongoing studies are investigating whether cross-cultural differences on this task exist in preschoolers. If younger children have not yet internalized conventional space-time mappings, we expect to observe less difference between the language groups earlier in development.

Although our study leaves open the idea that infants have an innate ability to align space and time in particular contexts, our findings in preschoolers suggest that space and time are not associated to one another automatically, and that instead this cultural convention is learned gradually in early childhood. Also, they suggest that an LR direction is not a biological default (Chatterjee et al., 1999). If such a default existed, we might expect 4-year-olds to use linear, LR mappings to guide behavior, as adults do, in tasks that *require* them to create spatial representations of time. However, preschoolers did not spontaneously deploy linear representations *even though* most were able to do so when primed. These findings in preschoolers are in contrast to the preference for ordinal, LR representations found in the majority of kindergarteners, even in the absence of priming, suggesting that the onset of automaticity and directional biases in most children may be linked to entering school.

Although the results of the No Prime condition argue against automatic deployment of an MTL in preschoolers, the results of the spatial priming conditions suggest that they may nonetheless find linear representations of time relatively intuitive. Spatial priming was strikingly effective, given how minimal the intervention was. Though past studies have shown that preschoolers can use explicit timelines and other physical artifacts to coarsely differentiate the times of events, this typically required adult demonstrations and extensive explanation (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman, 2000, 2002; Tillman et al., 2017). Here, the physical objects children observed during priming were no longer visible during the test trials, no linear template was used, and experimenters gave no overt instructions on how temporal events should be spatially arranged. Furthermore, the space of possible responses was large: Children could put their stickers anywhere on the cards. Nevertheless, priming increased linear representations of time in preschoolers by a factor of 2, suggesting that space and time are easily mapped by children. This finding is compatible with the idea that mental representations of time and space may be similar in structure, even if they are not automatically aligned (see Srinivasan & Carey, 2010, for discussion). Beyond this, the ease with which unconventional vertical mappings could be elicited suggests that alternative structure mappings between space and time can be dynamically generated on the fly, in response to external input (e.g., Gentner, 1983; Gentner & Markman, 1997; Murphy, 1996; see also Casasanto & Bottini, 2014, for evidence of flexibility in adults).

How can the current finding that children are slow to *spontaneously* associate time with an ordinal line be reconciled with previous reports that prelinguistic infants, in some cases only a few hours old, already make associations between spatial and temporal magnitudes (e.g., Brannon et al., 2007; de Hevia et al., 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010)? These striking findings have helped to fuel an active debate over the extent to which environmental factors, such as language and literacy, are required for space-time mappings (e.g., Bender & Beller, 2014; Bonato et al., 2012; Boroditsky, 2011; Bottini & Casasanto, 2010; 2013; Walsh, 2003; Lakoff & Johnson, 1980; Walsh, 2003; Winter, Marghetis, & Matlock, 2015). More specifically, the infant studies indicate that cultural and linguistic experience are not required to associate temporal magnitudes (e.g., a long sound) with spatial ones (e.g., a long line). Importantly, in the infant studies, both a spatial extent and a temporal duration were presented simultaneously, as two properties of a single stimulus, providing the opportunity to align them. One possibility is that the present results in preschoolers are not incompatible with the infant work, and that infants *only* align space and time if they are presented together. A second possibility is that the ability to associate spatial and temporal magnitudes, as demonstrated in infancy, simply does not require the directional spatial model of time that we explored here, and which has been explored more extensively in adults (for further discussion, see Winter, Marghetis, & Matlock, 2015). We are currently investigating whether preschoolers exhibit a preference when given a choice between ordered and unordered representations of events (i.e., if temporal and spatial features are provided in tandem; Tillman et al., 2017).

It is possible that preschoolers in the current study may have already had LR mappings, but simply not realized that producing such lines was the goal of the task in the No Prime condition. In other words, although children did not produce linear representations of time spontaneously, it is nevertheless possible that they mentally represent time in this way, and that spatial priming worked not by inducing a new strategy (i.e., the use of lines), but instead by indicating to children the nature of the task, and which of their existing knowledge to deploy. However, if linear priming served only to prompt participants to use an existing LR-specific MTL to guide their behavior in this task, we might expect a prime in any orientation to induce the use of LR mappings, which was not the case. Unlike adults, preschoolers and kindergarteners rarely made LR lines following a vertical prime. Furthermore, only a small proportion of children failed to use stickers appropriately in the blocks task (and were thus excluded), suggesting that children's production of non-linear arrangements on critical trials was not due to incomprehension of the task.

Interestingly, though kindergarteners were more likely to spontaneously adopt conventional LR representations of time than were preschoolers, our findings also suggest that the directionality of the MTL is still not fully ingrained in children this age. Specifically, kindergarteners remained much more flexible than adults in their use of space to represent time, and very rarely produced conventional LR lines after experiencing a vertical spatial prime. Consistent with the idea that development of space-time mappings continues later into the elementary school years, other recent work suggests that mappings between spatial position and temporal duration (e.g., between leftward space and shorter durations) may not be fully formed until as late as age 8 (Droit-Volet & Coull, 2015). Similarly, although children's ability to use explicit spatial timelines emerges during the preschool years (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman, 2000, 2002; Tillman et al., 2017), it also continues to improve until at least age 7 (Hudson & Mayhew, 2011; Tillman & Barner, 2015; Tillman et al., 2017).

While our study suggests that the direction of the MTL is learned slowly by young children, it leaves open the mechanism by which such learning might occur. Previous studies have pointed to both language and non-linguistic cultural practices as potential sources of cross-cultural differences in space-time mappings. In the case of the LR timeline in particular, language may play a relatively limited role. Although some have argued that space and time are associated via words like *behind* and *ahead,* which can describe either type of phenomenon, English does not express the horizontal timeline linguistically (see Clark, 1973; Cienki, 1998; Evans, 2004). Perceptual experience related to literacy, physical calendars, or other factors may be more critical to the development of the LR timeline than linguistic experience. Because reading and writing LR orthography repeatedly associates temporal progress through a narrative with rightward movement on the page, literacy has been argued to play a crucial role in the formation of the LR timeline (e.g., Bergen & Lau, 2012; Casasanto & Bottini, 2014; Tversky, Kugelmass, & Winter, 1991). Recent evidence suggests that blind participants who read Braille from left-to-right also develop LR space-time associations (Bottini et al., 2015). We found that LR representations of time increased substantially at around the age that children enter school and receive direct instruction on reading and writing. This suggests that early experiences such as passive exposure to LR orthography in children's books or counting objects from left-to-right (Shaki, Fischer, & Gobel, 2012) may not be sufficient for the formation of an adult-like LR timeline. Interestingly, LR representations of days exceeded those of meals in both kindergarteners and adults, but not in preschoolers (see

Supporting Information), a pattern that suggests that exposure to calendars may create specific associations between the ordering of the days of the week and the LR axis. However, we did not measure children’s emergent literacy skills, print exposure, or ability to use a calendar here. Consequently, future work should test these factors more directly, to determine the precise factors that trigger the mapping of time to space in early development.

Acknowledgements

We thank Junyi Chu, Lara Haines, Ana Magallanes, and Vaishnavi Nandanam for their assistance with subject recruitment and data collection. Tyler Marghetis, Mahesh Srinivasan, Vic Ferreira, Ben Bergen, Lera Boroditsky, members of the Language and Development Lab at UCSD, and three anonymous reviewers provided helpful discussion and comments on prior versions of the manuscript. We are most grateful to the children and families who participated in this research, and to the daycares, preschools, and museums where testing was conducted, particularly the Fleet Science Center in San Diego. Results of a prior version of this experiment are reported in Tillman, Tulagan, and Barner (2015). Some of the present results also appear in an unpublished undergraduate honors thesis by E.F. This work was funded by a grant to D.B. from the James S. McDonnell Foundation.

References

Barnett, J. E. (1999). *Time's pendulum: From sundials to atomic clocks, the fascinating history of timekeeping and how our discoveries changed the world*. Houghton Mifflin Harcourt.

Bender, A., & Beller, S. (2014). Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition*, *132*(3), 342-382.

Bergen, B. K., & Lau, T. T. C. (2012). Writing direction affects how people map space onto time. *Frontiers in psychology*, *3*.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2257-2273.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive psychology*, *43*(1), 1-22.

Boroditsky, L. (2011). How languages construct time. In Dehaene & Brannon (Eds.), Space, time and number in the brain: Searching for the foundations of mathematical thought. 978-0-12-385948-8. Elsevier

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently?. *Cognition*, *118*(1), 123-129.

Boroditsky, L., & Gaby, A. (2010). Remembrances of times east: Absolute spatial representations of time in an Australian aboriginal community. *Psychological Science*, *21*(11), 1635-1639.

Bottini, R., & Casasanto, D. (2010). Implicit spatial length modulates time estimates, but not vice versa. In *Spatial Cognition VII* (pp. 152-162). Springer Berlin Heidelberg.

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMic?. *Frontiers in psychology*, *4*.

Bottini, R., Crepaldi, D., Casasanto, D., Crollen, V., & Collignon, O. (2015). Space and time in the sighted and blind. *Cognition*, *141*, 67-72.

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. Developmental Science, 10(6), 770—777.

Busby Grant, J., & Suddendorf, T. (2009). Preschoolers begin to differentiate the times of events from throughout the lifespan. *European Journal of Developmental Psychology*, *6*(6), 746-762.

Casasanto, D., & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General*, *143*(2), 473.

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and time in the child's mind: Evidence for a cross-dimensional asymmetry. *Cognitive Science*, *34*(3), 387-405.

Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics*, *23*(4), 643-674.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, *37*(4), 395-402.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphorical expressions. In J. Koenig (ed.), Discourse and Cognition: Bridging the Gap, 189–204.Stanford: CSLI Productions.

Clark, H. (1973). Space, time, semantics, and the child. In T. E. Moore (ed.), Cognitive Development and the Acquisition of Language, 27–63. New York: Academic Press.

Cooperrider, K., & Núñez, R. (2009). Across time, across the body: Transversal temporal gestures. *Gesture*, *9*(2), 181-206.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002).Neural systems for visual orienting and their relationships to spatial working memory. Journal of Cognitive Neuroscience, 14(3), 508–523.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, *111*(13), 4809-4813.

De Sousa, H. (2012). Generational differences in the orientation of time in Cantonese speakers as a function of changes in the direction of Chinese writing. *Frontiers in psychology*, *3*.

Dobel, C., Diesendruck, G., & Bölte, J. (2007). How writing system and age influence spatial representations of actions a developmental, cross-linguistic study. *Psychological Science*, *18*(6), 487-491.

Droit-Volet, S., & Coull, J. (2015). The Developmental Emergence of the Mental Time Line: Spatial and Numerical Distortion of Time Judgement. *PloS one*,*10*(7), e0130465.

Evans, V. (2004). The Structure of Time: Language, Meaning and Temporal Cognition. Amsterdam: John Benjamins.

Fabbri, M., Cancellieri, J., Natale, V., (2012). The A Theory Of Magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica* 139, 111—123

Friedman, W. J. (2000). The development of children's knowledge of the times of future events. *Child development*, 913-932.

Friedman, W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Journal of Cognition and Development*, *3*(3), 333-356.

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development*, *13*(3), 335-367.

Fuhrman, O., Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: evidence from an implicit nonlinguistic task. *Cognitive Science* 34, 1430—1451.

Gell, A. (1992). The anthropology of time: cultural constructions of temporal maps and images.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*(2), 155-170.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American psychologist*, *52*(1), 45.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition*, *87*(3), B87-B95.

Grady, J. (1999). A typology of motivation for metaphor: Correlations vs. resemblances. In R. Gibbs & G. Steen (Eds.), *Metaphor in cognitive linguistics* (pp. 79—100). Amsterdam, the Netherlands: Benjamins

Haspelmath, M. (1997). *From space to time: Temporal adverbials in the world's languages*. Lincom Europa.

Hudson, J. A., & Mayhew, E. M. (2011). Children's temporal judgments for autobiographical past and future events. *Cognitive Development*, *26*(4), 331-342.

Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex*, *44*(4), 454-461.

James, W. (1890). The principles of psychology (Vol. 1). *New York: Holt*.

Kant, I. (2009/1781). In P. Guyer (Ed.), Critique of pure reason. Cambridge University Press

Lai, V. T., & Boroditsky, L. (2013). The immediate and chronic influence of spatio-temporal metaphors on the mental representations of time in English, Mandarin, and Mandarin-English speakers. *Frontiers in psychology*, *4*.

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. *Chicago: U of Chicago P*.

Le Guen, O., Balam, P., & Ildefonsa, L. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. Frontiers in psychology, 3, 271.

Lourenco, S. F., & Longo, M. R. (2010). General Magnitude Representation in Human Infants. *Psychological Science*, *21*(6), 873-881.

Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events nature or culture?. *Psychological Science*, *14*(4), 296-301.

McTaggart, J. E. (1908). The unreality of time. *Mind*, 457-474.

Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (2011). Can a mind have two time lines? Exploring space—time mapping in Mandarin and English speakers. *Psychonomic bulletin & review*, *18*(3), 598-604.

Murphy, G. L. (1996). On metaphoric representation. *Cognition*, *60*(2), 173-204.

Nachson, I. (1983). Directional preferences in bilingual children. *Perceptual and Motor Skills*, 56, 747—750.

Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in cognitive sciences*, *17*(5), 220-229.

Núñez, R. E., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive science*, *30*(3), 401-450.

Ouellet, M., Santiago, J., Israeli, Z., & Gabay, S. (2015). Is the future the right time?. *Experimental psychology*.

Piaget, J. (1927/1969). *The child's conception of time*. New York: Ballantine Books.

Piaget, J. (1946/1970). *The child's conception of movement and speed*. New York: Basic Books.

Pun, C., Adamo, M., Weger, U. W., Black, S. E., & Ferber, S. (2010). The right time and the left time: spatial associations of temporal cues affect target detection in right brain-damaged patients. *Cognitive neuroscience*, *1*(4), 289-295.

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science*, *347*(6221), 534-536.

Saj, A., Cojan, Y., Vocat, R., Luauté, J., & Vuilleumier, P. (2013). Prism adaptation enhances activity of intact fronto-parietal areas in both hemispheres in neglect patients. *Cortex*, *49*(1), 107-119.

Santiago, J., Lupáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512-516.

Shaki, S., Fischer, M. H., & Göbel, S. M. (2012). Direction counts: A comparative study of spatially directional counting biases in cultures with different reading directions. *Journal of Experimental Child Psychology*, *112*(2), 275-281.

Sinha, C., Sinha, V. D. S., Zinken, J., & Sampaio, W. (2011). When time is not space: the social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Language and Cognition*, *3*(1), 137-169.

Spelke, E., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In Bowerman, M. & Levinson, S. (Eds.). Language Acquisition and Conceptual Development. Cambridge: Cambridge University Press, pp 70-100.

Srinivasan, M., & Carey, S. (2010). The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217-241.

Tillman, K. A., & Barner, D. (2015). Learning the language of time: Children's acquisition of duration words. *Cognitive psychology*, *78*, 57-77.

Tillman, K. A., Marghetis, T, Barner, D., & Srinivasan, M. (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology*, 92, 87-100.

Tillman, K. A., Tulagan, N., & Barner, D. (2015). Building the mental timeline: Spatial representations of time in preschoolers. *Proceedings of the 37th Annual Conference of the Cognitive Science Society.*

Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science*, *30*(4), 745-757.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, *23*(4), 515-557.

Vallesi, A., Arbula, S., & Bernardis, P. (2014). Functional dissociations in temporal preparation: evidence from dual-task performance. *Cognition*, *130*(2), 141-151.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, *7*(11), 483-488.

Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, *15*(2), 426-430.

Whitrow, G. J. (1989). *Time in history: Views of time from prehistory to the present day*. Oxford University Press, USA.

Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. Carroll (Ed.), Language, thought, and reality. MIT Press.

Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209-224.

Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: The SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *Journal of Cognition and Culture*, *5*(1), 165-190.

Chapter 3 Supporting Information

1. Frequency of sticker arrangements

The following tables report counts for each type of sticker arrangement produced in the blocks priming tasks (Table S1) and the meals and days tasks given in the Horizontal (Table S2), No Prime (Table S3), and Vertical Prime (Table S4) conditions. Note that "line" here indicates only that the largest angle created by the three stickers was > 140 deg, while *ordinal* lines ("Ord. lines") indicates that the arrangement was both linear and ordered. While only ordered lines were considered to be "linear representations of time" for the purposes of the primary text of the paper, it is interesting to note that younger children sometimes put the stickers in a line even when they failed to represent the correct temporal ordering of the items the stickers were meant to denote, suggesting that some preschoolers may have a general tendency to arrange items in lines.

Lines were only coded for direction if they were both linear and ordered. Lines were coded as "diagonal" when the best fitting line through all three stickers was more than 30 deg away from any of the four cardinal directions. In other words, a line 0-30 deg. from horizontal would be coded as either LR or RL (depending on sticker order), between 30-60 deg. would be coded as diagonal, and between 60-90 deg. would be coded as either TB or BT.

Table 3.1. Sticker arrangements in the spatial priming tasks.

Task	Group	N	Lines	Ord. lines	LR	RL	TB	BT	Diag.
Horizontal blocks	PS	40	38	37	32	3	3	0	2
	Kinder	23	23	23	23	0	0	0	0
	Adult	20	20	20	20	0	0	0	0
Vertical blocks	PS	46	37	35	2	3	26	3	1
	Kinder	31	31	28	1	0	27	0	0
	Adult	21	21	20	0	0	20	0	0

Table 3.2. Sticker arrangements in the Horizontal Prime condition.

Task	Group	N	Lines	Ord. lines	LR	RL	TB	BT	Diag.
Days	PS	31	24	24	17	6	0	0	1
	Kinder	24	21	20	15	2	2	1	0
	Adult	20	20	20	19	1	0	0	0
Meals	PS	31	22	21	12	9	0	0	0
	Kinder	24	23	22	12	8	1	1	0
	Adult	20	20	20	17	0	2	0	1
Both	PS	62*	46	45	29	15	0	0	1
	Kinder	48*	44	42	27	10	3	2	0
	Adult	40*	40	40	36	1	2	0	1

* denotes total number of arrangements created, rather than participants tested. Each participant created one arrangement in the Meals task and one in the Days task.

Table 3.3. Frequency of sticker arrangement types created in the No Prime condition.

Task	Group	N	Lines	Ord. lines	LR	RL	TB	BT	Diag.
Days	PS	67	33	24	9	7	3	2	3
	Kinder	34	31	21	16	1	1	0	3
	Adult	21	21	21	18	0	2	0	1
Meals	PS	67	30	21	10	5	2	1	3
	Kinder	34	27	21	11	4	1	4	1
	Adult	21	20	20	11	0	8	0	1
Both	PS	134*	64	45	19	12	5	3	6
	Kinder	68*	58	42	27	5	2	4	4
	Adult	42*	41	41	29	0	10	0	2

* denotes total number of arrangements created, rather than participants tested. Each participant created one arrangement in the Meals task and one in the Days task.

Table 3.4. Sticker arrangements in the Vertical Prime condition.

Task	Group	N	Lines	Ordered lines	LR	RL	TB	BT	Diag.
Days	PS	27	21	20	2	1	6	11	0
	Kinder	27	25	21	3	0	8	10	0
	Adult	20	20	20	13	0	2	3	2
Meals	PS	27	21	20	1	1	6	11	1
	Kinder	27	26	24	2	2	12	8	0
	Adult	20	20	20	10	0	5	4	1
Both	PS	54*	42	40	3	2	12	22	1
	Kinder	54*	51	45	5	2	20	18	0
	Adult	40*	40	40	23	0	7	7	3

* denotes total number of arrangements created, rather than participants tested. Each participant created one arrangement in the Meals task and one in the Days task.

2. Mixed-effects modeling

Using mixed effects logistic regression, we modeled the likelihood of a participant producing a linear sticker arrangement. As main effects, we included Age Group (preschool, kindergarten, adult), Priming Condition (horizontal, no prime, vertical), Trial Type (meals, days), and Trial Order (meals first, days first), as well as a random effect of subjects. Model parameters are reported in Table S5. Relative to reduced models which did not include these factors, both Age Group ($X^2(2) = 26.0$, $p < 0.001$) and Condition ($X^2(2) = 21.1$, $p < 0.001$) significantly improved the fit of the model. However, neither Trial Order (X2(1)=0.21, p=0.6) nor Trial Type ($X^2(1) = 0.02$, $p=0.9$) improved the model.

Separately, considering only linear arrangements, we modeled the likelihood of producing an LR line, using the same effects structure as above (see Table S6). Again, both Age Group ($X^2(2) = 16.6$, $p < 0.001$) and Condition ($X^2(2) = 19.3$, $p < 0.001$) significantly improved the model. Again, there was no effect of Trial Order ($X^2(1) = 1.4$, $p = 0.2$). However, unlike above, we also found that Trial Type (meals vs. days) was a significant predictor ($X^2(1) = 9.8$, $p = 0.002$).

Figure S1 shows the proportions of participants who made LR arrangements of days vs. meals in each condition. Interestingly, though we lacked sufficient statistical power to test for interaction effects in the models above, the effect of trial type appears is particularly strong in the adult controls.Through 86% (95% CI [53%-89%]) of adults spontaneously place stickers in a line in the No Prime condition, only 55% (95% CI [35%-73%]) did so with meals.

Table 3.5. Logistic regression model parameters predicting ordinal lines.

Fixed effect:	Estimate	Std. error	z	*p*
Intercept	1.7	0.61	2.8	< 0.01
Group (kinder)	1.6	0.52	3.4	< 0.001
Group (adult)	6.6	1.4	4.7	< 0.001
Prime (none)	-2.8	0.69	-4.2	< 0.001
Prime (vertical)	-0.14	0.64	-0.21	0.83
Trial type (meals)	-0.08	0.29	-0.29	0.77
Trial order (meals first)	0.18	0.45	0.40	0.69

Table 3.6. Logistic regression model parameters predicting LR lines.

Fixed effect:	Estimate	Std. error	z	*p*
Intercept	-0.4	0.56	-0.86	0.39
Group (kinder)	1.5	0.56	2.7	0.006
Group (adult)	5.2	0.97	5.3	< 0.001
Prime (none)	-2.5	0.67	-3.7	< 0.001
Prime (vertical)	-4.4	0.88	-5.0	< 0.001
Trial type (meals)	-1.1	0.32	-3.3	< 0.001
Trial order (meals first)	0.58	0.46	1.3	0.69

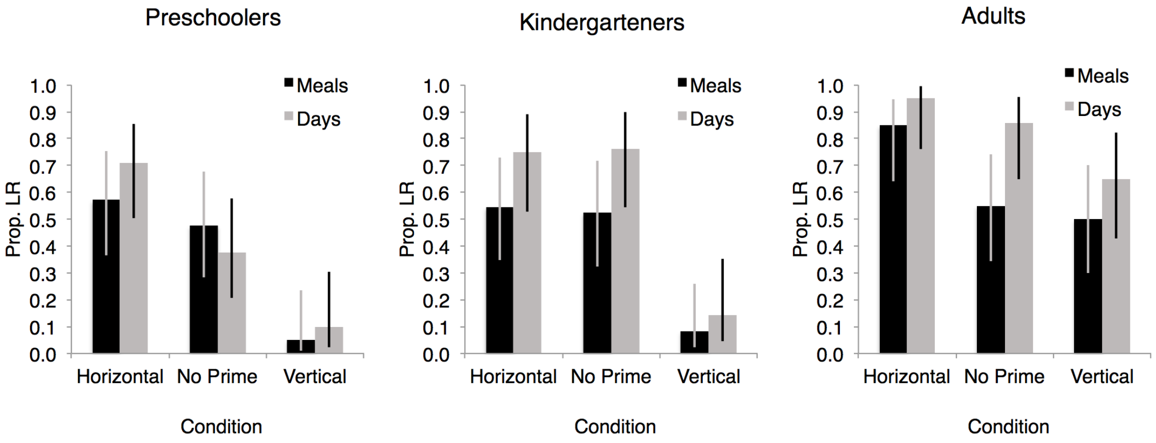

Figure 3.4. *Left-to-right arrangements of meals and days.* Proportion of linear representations of meals and days that were conventionally organized from left-to-right, in each priming condition.

Chapter 3, in full, is a reprint of the material currently under review at *Developmental Science.* Tillman, K. A., Tulagan, N., Fukuda, E., & Barner, D. The dissertation author was the primary investigator and author of this paper.

GENERAL DISCUSSION

Because time perception is highly subjective, the necessity of coordinating activities across large groups of people has led diverse cultures to create external symbolic systems to precisely encode the passage of time. Although humans have basic capacities for time perception from birth, these formal systems for time-keeping — including time words, clocks, and calendars — take children many years to learn. How do children acquire these symbolic systems, and how do these abstract constructs relate to the inner experience of time, early in development? In this dissertation, I explored how 3- to 8-year-old children first acquire duration words (Chapter 1), deictic time words (Chapter 2), and linear representations of time (Chapter 3). The results of each study indicated that children begin to construct abstract systems to represent and communicate about time as early as age 4. However, in each case, I argued that primitives available from infancy — including perceptual representations of duration, events, and spatial magnitudes — play a limited role in this process. In early acquisition, the meanings of abstract time words reflect only their relationships to other words within a linguistic domain. Further, the implicit spatial structures adults use when reasoning about time (i.e., the "mental timeline") arise in response to early experience with cultural tools and practices, not before.

In Chapter 1, I presented evidence that children's early meanings for duration words like "minute" do not reflect mappings to approximate, perceptual representations of duration. Instead, these early meanings encode the relations between time words within an ordered semantic domain – *prior* to being individually mapped to perception.

Similarly, in Chapter 2, I argued that early meanings for deictic time words like “yesterday” do not reflect their relations with experienced events, but rather their ordering in a common lexical domain. The findings of both Chapters 1 and 2 suggest that structured semantic domains for time words are formed primarily using cues found in natural language, and that mappings between time words and properties of experienced events do not emerge until after children are taught their formal definitions. Next, in Chapter 3, I presented evidence that the left-to-right “mental timeline” is not the result of innate spatial representations of time. I found that both the automatic use of the mental timeline and its conventional directional structure were absent in preschoolers. Together, these studies provide evidence that the abstract concept of time shared by most Western adults is not built from innate primitives or perceptual experiences of events. These relational structures are separately constructed as children acquire language, literacy, and experience with cultural tools. I review these findings in greater detail below, and further discuss their implications and possible directions for future work.

Children's acquisition of time words.

In most cases of early word learning, comprehension precedes production. However, in the case of time words like “minute” and “yesterday”, while some children begin using them as toddlers, up to five years may pass before they acquire adult-like meanings. This gap between production and adult-like comprehension can provide a window into the process of abstract word learning and conceptual development. Prior work showed that, although children use time words incorrectly for years, they make errors that seem to imply partial understanding. For example, one study showed that 4-year-old preschoolers respond to questions about duration with duration words, even if they do so highly inaccurately (Shatz et al., 2010). Importantly, this result suggests that preschoolers have already formed a lexical category for duration words. Further probing children's early comprehension of time words, I presented evidence that children's early lexical categories of time words are not simple groupings, but instead resemble ordered lists. For example, in Chapter 1, I found that, despite often using the words incorrectly, 4-year-olds could choose which of two duration words (e.g., "a minute" vs. "an hour") indicates the longer duration, and 5-year-olds could rank-order several duration words on a spatial scale, from shortest to longest. Additionally, in Chapter 2, I showed that 4-year-olds could spatially represent both the deictic (past/future) status of time words like "yesterday" relative to the present, and their ordering relative to other terms (e.g., that "last week" was before "yesterday") on a bidirectional timeline extending from the past to the future. In fact, children could do this just as well with time words as with familiar events (e.g., "breakfast"), and by age 6, most children could do this as well as adults.

However, beyond knowledge of their common semantic categories and relative ordering within them, my findings revealed that children's understanding of time words is limited. In the case of duration words, young children lacked knowledge of the absolute duration denoted by each word. Preschoolers were no better at comparisons of more disparate terms (e.g., "second" vs. "year") than of nearer ones. Even 6-year-olds often failed on comparisons like "3 minutes" vs. "2 hours," indicating profound ignorance of the approximate minute-to-hour ratio. Although 5-year-olds could order duration words along a timeline, they did not "space out" these words to indicate relative differences, as adult controls did, until age 7 or later. Similarly, I showed in Chapter 2 that although preschoolers could correctly assign deictic time words to the past or future, and put them in order, the majority of children did not demonstrate adult-like knowledge of their temporal *remoteness* from the present until after 7 years of age. Finding that these abilities do not emerge until children have entered school suggests that formal definitions (e.g., *hour* = *minute* × 60; *year* = *day* × 365) may be required.

Importantly, the limits of children's understanding of time words are inconsistent with the hypothesis that children learn these words by associating them with innate, nonverbal representations of time, such as duration or event perception. Such mappings between verbal labels and nonverbal representations of time are possible in principle. Humans have access to perceptual representations of duration from infancy. In the first months of life, infants habituate to temporal patterns, displaying anticipatory behaviors time-locked to learned temporal delays (e.g., Brackbrill et al., 1967), and they can discriminate brief stimuli by their durations (e.g., Provasi et al., 2011). Furthermore, in Chapter 1, preschoolers were capable of using timelines to estimate the relative durations of familiar events, like washing hands vs. eating lunch, suggesting that the approximate durations of such events are encoded in memory. However, despite these capacities, reliable associations between words and durations did not appear until years later. Thus, the link between language and duration perception is apparently unintuitive for children

Our findings in Chapters 1 and 2 are consistent with the hypothesis that children's first meanings for time words are relational, reflecting word-to-word mappings, and may be inferred primarily from other language input. The facets of time-word meaning English-speaking children learn early (e.g., order, deictic status) are more readily encoded in English than those children acquire later (e.g., duration, remoteness). For instance, deictic status could potentially be inferred from verb tense, as predicted by a syntactic bootstrapping account (Brown, 1957; Gleitman, 1990; Gleitman et al., 2005; Landau & Gleitman, 1985; Naigles, 1996). The sentence "she *will go* to the beach *tomorrow"* indicates that "tomorrow" is in the future, but not its approximate remoteness from now. The order of terms within common domains could be learned from common sentence frames, contrastive usages, or order-of-mention in discourse. For example, a sentence like "it won't be done in a *minute*, because it takes a whole *hour*" indicates which word indicates the longer duration, and one like "she bought the apples *last week*, but didn't use them until *yesterday"* indicates which time word denotes the earlier time. However, such comparisons do not reveal absolute durations or specific locations in time, and young children are much less to hear sentences that do, e.g., "we went to the party *last year*, which was about *365 days ago*."

Our findings in the cases of duration words and deictic time words have implications for general theories of abstract word learning. For instance, consistent with the present findings, Susan Carey and others have proposed that, from early in development, abstract words are defined by their *inferential roles* within semantic structures (see Carey, 2009; Wagner, Tillman, & Barner, 2016). This "Quinian bootstrapping" idea is most frequently discussed in the context of the conceptual domain of number, which provides a relevant and useful comparison case for time. Similar to the time-word case, children produce number words long before they have adult-like meanings, and they acquire exact meanings for them gradually over an extended period (e.g., Le Corre & Carey, 2011; Wynn, 1990; 1992). As we discuss further elsewhere (Wagner, Tillman, & Barner, 2016), there are several similarities between the cases of number and duration words. For example, children appear to first acquire language-based “placeholder structures” relating sets of terms, but with little (if any) perceptual content. Adult-like meanings for most terms emerge late in development and in relative synchrony, around the time that children are taught exact meanings, indicating that the meanings of each word are mutually constrained by the others. And, perhaps most critically, the system as a whole is not mapped to perception until quite late in development, if at all.

There are also key differences between the cases of number and time words, which are more pronounced for the many time words that do not refer to event durations. Though their details vary substantially, almost all theories of the development of numerical concepts posit some key role for innate perceptual systems: a core system of individual object representation in visual memory, and/or an evolutionarily-ancient system for representing quantities approximately (e.g., Carey, 2009; Gelman & Gallistel, 1978; Leslie, Gelman & Gallistel, 2008; Spelke & Tsivkin, 2001). In the case of time, direct mappings from small duration words, like “second”, to nonverbal representations of duration, or associations between terms like “today” and the experience of daily routines or sun cycles are plausible. However, it is less clear how innate perceptual systems or word-to-world mappings could *ever* provide content for future-oriented terms like “last year”, “next week”; for calendar terms like “January” and “Thursday”; or even words that refer to very large spans of time, like “century” or “forever.”

Recently, Barner (2017) proposed an alternative to prior models of number word learning, in which perception does not provide the content for number concepts at any stage of their acquisition. Instead, the limitations of perception are taken as the impetus for human cultures (and individual children) use *other* resources for exact number representation, such as natural language or visual counting systems. This general framework may also provide a more comfortable fit to the current data on time words. Symbolic systems for time are useful precisely because they provide a means of precisely describing and quantifying the passage of time that innate perceptual systems cannot. Indeed, this early separation of the meanings of abstract temporal words from the temporal events they are used to reference may be critical for one of the central insights of a linear, Newtonian conception of time: the dimension of time is not a *feature* of particular events, but a *framework* for organizing and describing them.

Children's acquisition of the mental timeline.

Space can also provide framework for representing time, both externally and in the mind. Indeed, human cultures link space and time in many ways, including use of space-time metaphors in language, cultural practices like reading/writing in a particular orthography, and tools like clocks and calendars. Although space-time associations were not their focus, both the studies in Chapters 1 and 2 took advantage of children's early ability to use a space to represent time. Specifically, in both cases, we assessed children's knowledge of the meanings of time words in part by measuring their ability to place these items on left-to-right timelines. Preschoolers' ability to use and comprehend complex timelines in the present studies, and elsewhere in the literature (Busby Grant & Suddendorf, 2009; Friedman & Kemp, 1998; Friedman, 2000, 2002; Hudson & Mayhew, 2011), is striking, and suggests that mappings between time and space are, at a minimum, relatively intuitive for children.

A large body of research indicates that adults both explicitly and implicitly associate increasing time with progress along a particular linear path, known as the "mental timeline" (MTL). There is much debate over whether this phenomenon is a learned cultural convention, or whether it (also) reflects an underlying predisposition to reason about time using space (e.g., Boroditsky, 2011; Bottini & Casasanto, 2013; Casasanto & Boroditski, 2008; Chatterjee et al., 1999; Walsh, 2003; Winter, Marghetis, & Matlock, 2015) On the one hand, even infants readily associate spatial and temporal magnitudes, suggesting that some mappings between time and space may be innate (de Hevia et al., 2014; Brannon et al., 2007; Laurenco & Longo, 2010; Srinivasan & Carey, 2010). On the other hand, cross-cultural differences in the direction and orientation of the MTL in adults suggest that it is learned (Bergen & Lau, 2012; Fuhrman & Boroditsky, 2010; Ouellet et al., 2015; Nachson, 1983; Tversky, Kugelmass, & Winter, 1991).

Because timeline tasks like those we used in Chapters 1 and 2 provide children with a particular type of space-time mapping to adopt, these studies do not address whether children would invoke linear representations of time spontaneously. In Chapter 3, we used a more open-ended task, in which children placed stickers representing temporal items on a blank card, to assess their usage of linear representations of time (see also Tversky, Kugelmass, & Winter, 1991). We showed that preschoolers could be easily primed to create linear mappings between time and space. Critically, however, *despite* having the ability to do so, only a third of preschoolers spontaneously created ordinal sequences of time-denoting stickers. Further, among the subset of preschoolers who did create linear representations of time without a prime, there was no significant preference for LR over right-to-left lines. This pattern of results is inconsistent with the proposal that the mental timeline in general, or the LR timeline in particular, is an innate neurological constraint (e.g., Chatterjee et al., 1999). Instead, our results suggest that the mental timeline is gradually adopted in childhood, becoming more automatic and conventional with increasing cultural exposure. Relatedly, even the kindergarteners in our study remained much more flexible in their use of space to represent time than adult controls, suggesting that the process of internalizing the LR convention extends further into the elementary school years.

Implications for education.

Several findings in the present dissertation suggest that some aspects of the way Western adults think about time cannot be easily learned outside the context of formal education. For example, Chapters 1 and 2 indicate that the ability to estimate durations and temporal distances using conventional units of time develops late in development, likely relying on formal training on time-word definitions in school. Furthermore, Chapter 3 revealed that children do not spontaneously create linear representations of time until they reach kindergarten, which suggests that formal instruction on reading, writing, and calendar use could be critical in shaping the way children and adults associate time with space.

In the United States, children receive many years of explicit instruction about time and timekeeping. Telling time becomes a focus of standard mathematics curricula in the U.S. starting in Grade 1 and continuing through Grade 4. Students are expected to use analog and digital clocks to tell time to the hour in Grade 1, to the nearest 5 minutes in Grade 2, and to the minute in Grade 3 (Common Core State Standards Initiative, 2012). Children are also asked to perform mathematical operations involving temporal quantities in Grades 3 and 4, and to represent time using spatial timelines (Common Core State Standards Initiative, 2012). Instruction on time is (necessarily) closely intertwined with that of other core elements in early mathematics, and one recent study indicates that time knowledge in 6- to 11-year olds is correlated with both their knowledge of number facts and their number-line estimation skills (Labrell et al., 2016).

Although teaching time is widely considered to be one of the hardest challenges primary school teachers face, research on this topic in the educational literature is surprisingly scarce (Burny et al., 2009; Kamii & Russell, 2012). Educational psychologists have stressed the need for new, evidence-based instructional methods in the domain of time (e.g., Burny et al., 2009). Importantly, knowledge of the meanings of time words is a prerequisite for successful clock use (Burny et al., 2009; Cohen et al., 2000; Friedman & Laycock, 1989), and it has been argued that difficulty in calculating elapsed time may stem primarily from inability to coordinate units such as minutes and hours (Kamii & Russell, 2012). In order for educators to deploy developmentally-appropriate scaffolding for the acquisition of temporal terms like these, it is vital to take into account children's understanding of these terms *before* formal instruction begins. Therefore characterizing preschoolers' knowledge of time-related words, as the current studies have done, may aid the quest for more effective teaching strategies in the elementary classroom.

The current studies also contribute to our understanding of children's ability to use spatial tools to represent time. Suggesting that the specific spatial tools used in the classroom matter for children's learning outcomes, one recent study showed that 3rd and 4th graders' performance on temporal problem solving differed according to what type of manipulative clock they were given as an aid (Earnest, 2017). Spatial timelines are prevalent in elementary school textbooks and worksheets. However, their effectiveness as learning aids in the classroom is understudied, and educators are provided no guidance on how to best explain them or how to interpret children's work (Burny et al., 2009). The current studies indicate that children are capable of representing time linearly by the age of 4. Timelines and other spatial tools may therefore be more effective as learning aids if they are introduced to students earlier, perhaps in pre-kindergarten classrooms. Moreover, in Chapter 2, children's ability to use spatial timelines provided a more sensitive measure of their knowledge of time words than did verbal questioning. The timeline scoring systems developed in Chapters 1 and 2, which characterize performance in terms of separable facets of time knowledge (i.e., deictic status, order, and remoteness) could therefore have practical value as assessment tools in the classroom.

Future directions.

The present dissertation leaves many open questions and avenues for future work. For example, while the current results suggest that linguistic cues play an important role in constraining children's early hypotheses about the meanings of time words, the question of exactly *which* cues those are remains open. This is important both because it could further inform theories about abstract word learning and because it could help inform teaching strategies. As mentioned above, children's early acquisition of the deictic status of words like "yesterday" consistent with a *syntactic bootstrapping* account, and their early acquisition of relative ordering suggests that other cues from discourse (e.g., order-of-mention, lexical contrast) also play a role. However, the present studies do not directly test these hypotheses.

In ongoing and future work, I will more precisely characterize the learning cues available to children in their speech input from adults. One means of doing this is by analyzing corpora of child-directed speech, and a second means is by leveraging cross-linguistic diversity. Languages vary widely in how they encode time. For instance, unlike English, some languages have no morphological tense system, while others have multiple tenses within the past and future, each encoding a different degree of remoteness from the present (e.g., Comrie, 1985). To the extent that children selectively rely on syntax to guide their inferences about time-word meaning, the expressivity of a language's tense system ought to impact children's acquisition of time words. To test this, in one ongoing study, I am comparing comprehension of "yesterday" and "tomorrow" in children who speak English and children who speak Cantonese, a language that lacks a morphological tense system. In addition to differing in temporal syntax, languages also differ in how many time words they have, and which facets of time those words encode. For instance, unlike English, many languages include single lexical items that indicate a time precisely two days in the future (like "the day after tomorrow") and/or two days in the past ("the day before yesterday"). Future cross-linguistic comparisons would also allow me to test how variation in the set of alternatives available in a class of time words impacts children's early meanings.

In the case of space–time interactions, it remains unclear how to best account for both the prior findings that infants appear to spontaneously associate time and space (e.g., de Hevia et al., 2014; Brannon et al., 2007; Laurenco & Longo, 2010; Srinivasan & Carey, 2010) and the current finding that preschoolers do not. One interesting possibility is that this difference is driven by the fact that the two lines of work test different facets of time: duration in the infants, and sequence in the preschoolers. Another key difference between the studies is that in the infant work, both the spatial (i.e., length) and temporal (i.e., duration) cues to be integrated are presented to the infant simultaneously, thus setting the stage for an alignment. Because the sticker-placement task requires children to *create* a spatial representation of time, this production task could have been less sensitive to early associations between time and space in preschoolers. To address this, I am currently developing new comprehension tasks with more limited response demands.

Understanding the nature of time in the mind is one of the biggest problems in cognitive science, and one of the most fascinating. Here, I have argued that little of the Western adult's concept of time is innate, and much of it is shaped by experience with language and cultural artifacts. Many questions about the implications of this are left to be answered. However, characterizing how temporal language and space-time mappings develop in children may be a necessary step toward understanding how these symbolic systems impact other aspects of human cognition.

References

Barner, D. (2017). Language, procedures, and the non-perceptual origin of number word meanings. *Journal of Child Language*, 1-38.

Bender, A., & Beller, S. (2014). Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition*, *132*(3), 342-382.

Bergen, B. K., & Lau, T. T. C. (2012). Writing direction affects how people map space onto time. *Frontiers in psychology*, *3*.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*(10), 2257-2273.

Boroditsky, L. (2011). How Languages Construct Time. In *Space, Time and Number in the Brain* (pp. 333–341).

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently?. *Cognition*, *118*(1), 123-129.

Boroditsky, L., & Gaby, A. (2010). Remembrances of times east: Absolute spatial representations of time in an Australian aboriginal community. *Psychological Science*, *21*(11), 1635-1639.

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMic?. *Frontiers in psychology*, *4*.

Brackbill, Y., Fitzgerald, H. E., & Lintz, L. M. (1967). A developmental study of classical conditioning. *Monographs of the society for research in child development*, 1-63.

Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental science*, *10*(6), 770-777.

Brown, R. W. (1957). Linguistic determinism and the part of speech. The Journal of Abnormal and Social Psychology, 55(1), 1.

Burny, E., Valcke, M., & Desoete, A. (2009). Towards an agenda for studying learning and instruction focusing on time-related competences in children. *Educational Studies*, *35*(5), 481-492.

Busby Grant, J., & Suddendorf, T. (2009). Preschoolers begin to differentiate the times of events from throughout the lifespan. *European Journal of Developmental Psychology*, *6*(6), 746-762.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, *37*(4), 395-402.

Cohen, M. J., Ricci, C. A., Kibby, M. Y., & Edmonds, J. E. (2000). Developmental progression of clock face drawing in children. *Child Neuropsychology*, *6*(1), 64-76.

Common Core State Standards Initiative (2012). *Common core state standards for mathematics*.

Comrie, B. (Ed.). (1985). *Tense* (Vol. 17). Cambridge University Press.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*, *111*(13), 4809-4813.

Dobel, C., Diesendruck, G., & Bölte, J. (2007). How writing system and age influence spatial representations of actions a developmental, cross-linguistic study. *Psychological Science*, *18*(6), 487-491.

Droit-Volet, S., & Coull, J. (2015). The Developmental Emergence of the Mental Time-Line: Spatial and Numerical Distortion of Time Judgement. *PloS one*,*10*(7), e0130465.

Earnest, D. (2017). Clock Work: How Tools for Time Mediate Problem Solving and Reveal Understanding. *Journal for Research in Mathematics Education*, *48*(2), 191-223.

Friedman, W. J. (1990). Children’s representations of the pattern of daily activities. *Child Development, 61*, 1399–1412.

Friedman, W. J. (2000). The development of children's knowledge of the times of future events. *Child development*, 913-932.

Friedman, W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Journal of Cognition and Development*, *3*(3), 333-356.

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development*, *13*(3), 335-367.

Friedman, W. J., & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development*, 357-371.

Fuhrman, O., Boroditsky, L. (2010). Cross-cultural differences in mental representationsof time: evidence from an implicit nonlinguistic task. *Cognitive Science* 34, 1430—1451.

Gelman, R., & Gallistel, C. (1978). The child’s understanding of number. Cambridge, MA: Harvard University Press.

Gleitman, L. (1990). The structural sources of verb meanings. Language Acquisition, 1(1), 3–55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. Language Learning and Development, 1, 23–64.

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. Developmental Psychology, 11(6), 864–865.

Hudson, J. A., & Mayhew, E. M. (2011). Children's temporal judgments for autobiographical past and future events. *Cognitive Development*, *26*(4), 331-342.

Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex*, *44*(4), 454-461.

Kamii, C., & Russell, K. A. (2012). Elapsed time: why is it so difficult to teach?. *Journal for Research in Mathematics Education*, *43*(3), 296-315.

Labrell, F., Mikaeloff, Y., Perdry, H., & Dellatolas, G. (2016). Time knowledge acquisition in children aged 6 to 11 years and its relationship with numerical skills. *Journal of experimental child psychology*, *143*, 1-13.

Lai, V. T., & Boroditsky, L. (2013). The immediate and chronic influence of spatio temporal metaphors on the mental representations of time in English, Mandarin, and Mandarin-English speakers. *Frontiers in psychology*, *4*.

Landau, B., & Gleitman, L. R. (1985). Language and experience: Evidence from the blind child. Cambridge, MA: Harvard University Press.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395-438.

Leslie, A. M., Gelman, R., & Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends in cognitive sciences*, *12*(6), 213-218.

Lourenco, S. F., & Longo, M. R. (2010). General Magnitude Representation in Human Infants. *Psychological Science*, *21*(6), 873-881.

Nachson, I. (1983). Directional preferences in bilingual children. *Perceptual and Motor Skills*, 56, 747—750.

Naigles, L. (1996). The use of multiple frames in verb leaning via syntactic bootstrapping. Cognition, 58, 222–251.

Núñez, R. E., & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive science*, *30*(3), 401-450.

Ouellet, M., Santiago, J., Israeli, Z., & Gabay, S. (2015). Is the future the right time?. *Experimental psychology*.

Provasi, J., Rattat, A. C., & Droit-Volet, S. (2011). Temporal bisection in 4-month-old infants. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(1), 108.

Santiago, J., Lupáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512-516.

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11,* 16-36.

Srinivasan, M., & Carey, S. (2010). The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition*, *116*(2), 217-241.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, *23*(4), 515-557.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, *7*(11), 483-488.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127* (3), 307-317.

Wagner, K., Tillman, K. A., & Barner, D. (2016). Inferring number, time, and color concepts from core knowledge and linguistic structure. *Core Knowledge and Conceptual Change.* Oxford University Press.

Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. Carroll (Ed.), Language, thought, and reality. MIT Press.

Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209-224.