COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 35 (2011) 1305–1328

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2011.01193.x

How Linguistic and Cultural Forces Shape Conceptions of Time: English and Mandarin Time in 3D

Orly Fuhrman,[a] Kelly McCormick,[a] Eva Chen,[b] Heidi Jiang,[c] Dingfang Shu,[d] Shuaimei Mao,[d] Lera Boroditsky[a]

[a]*Department of Psychology, Stanford University*
[b]*Graduate School of Education, Harvard University*
[c]*Department of Cognitive and Linguistic Sciences, Brown University*
[d]*Shanghai International Studies University*

Received 5 October 2010; received in revised form 18 February 2011; accepted 3 March 2011

Abstract

In this paper we examine how English and Mandarin speakers think about time, and we test how the patterns of thinking in the two groups relate to patterns in linguistic and cultural experience. In Mandarin, vertical spatial metaphors are used more frequently to talk about time than they are in English; English relies primarily on horizontal terms. We present results from two tasks comparing English and Mandarin speakers' temporal reasoning. The tasks measure how people spatialize time in three-dimensional space, including the sagittal (front/back), transverse (left/right), and vertical (up/down) axes. Results of Experiment 1 show that people automatically create spatial representations in the course of temporal reasoning, and these implicit spatializations differ in accordance with patterns in language, even in a non-linguistic task. Both groups showed evidence of a left-to-right representation of time, in accordance with writing direction, but only Mandarin speakers showed a vertical top-to-bottom pattern for time (congruent with vertical spatiotemporal metaphors in Mandarin). Results of Experiment 2 confirm and extend these findings, showing that bilinguals' representations of time depend on both long-term and proximal aspects of language experience. Participants who were more proficient in Mandarin were more likely to arrange time vertically (an effect of previous language experience). Further, bilinguals were more likely to arrange time vertically when they were tested in Mandarin than when they were tested in English (an effect of immediate linguistic context).

Keywords: Space; Time; Implicit association; Metaphor; Mandarin; English; Language; Culture

Correspondence should be sent to Lera Boroditsky, Stanford University, 450 Serra Mall, Jordan Hall, Bldg 420, Stanford, CA 94305-2130. E-mail: lera@stanford.edu

To represent time, people around the world rely on space. We spatialize time in cultural artifacts like graphs, timelines, orthography, clocks, sundials, hourglasses, and calendars, we gesture timelines, and rely heavily on spatial words (e.g., forward, back, long, short) to talk about the order and duration of events (e.g., Clark, 1973; Lakoff & Johnson, 1980; Traugott, 1978). People's private mental representations of time also appear to be based in space; irrelevant spatial information readily affects people's judgments of temporal order and duration (Boroditsky, 2000; Boroditsky & Gaby, 2010; Boroditsky & Ramscar, 2002; Casasanto & Boroditsky, 2008; Matlock, Ramscar, & Boroditsky, 2005; Núñez, Motz, & Teuscher, 2006), and people seem to implicitly and automatically generate spatial representations when thinking about time (Fuhrman & Boroditsky, 2010; Gevers, Reynvoet, & Fias, 2003; Ishihara, Keller, Rossetti, & Prinz, 2008; Miles, Nind, & Macrae, 2010; Santiago, Lupiáñez, Pérez, & Funes, 2007; Torralbo, Santiago, & Lupianez, 2006; Weger & Pratt, 2008).

However, the particular ways that time is spatialized differ across languages and cultures. Several aspects of linguistic, cultural, and personal experience have been shown to shape people's temporal reasoning.

1. Spatiotemporal metaphors: The spatial metaphors that are used to talk about time affect people's representations of event order and duration (Boroditsky, 2000, 2001; Boroditsky, Fuhrman, & McCormick, 2010; Casasanto et al., 2004; McGlone & Harding, 1998; Núñez & Sweetser, 2006). For example, languages differ in whether they talk about the past as being behind or in front of the observer and these patterns in metaphor affect how their speakers arrange time in gesture (Núñez & Sweetser, 2006).
2. Availability of spatial representations: Because people tend to recruit spatial representations to think about time, representations of time also differ depending on what spatial representations are most cognitively available to co-opt for time (either in the immediate environment or in the culture more generally) (Boroditsky, 2000; Boroditsky & Gaby, 2010; Boroditsky & Ramscar, 2002; Matlock et al., 2005; Núñez et al., 2006).
3. Writing direction: Cultural artifacts such as systems of writing also shape people's representations of time (Chan & Bergen, 2005; Fuhrman & Boroditsky, 2010; Ouellet, Santiago, Israeli, & Gabay, 2010; Tversky, Kugelmass, & Winter, 1991). For example, Hebrew and Arabic speakers (who read from right to left) are more likely to arrange time from right to left than are English speakers (who read from left to right).
4. Finally, a variety of other aspects of cultural or individual disposition, age, and experience importantly shape people's representations of time (e.g., Carstensen, 2006; Gonzalez & Zimbardo, 1985; Ji, Guo, Zhang, & Messervey, 2009).

Across these studies, people in different cultures or groups have been shown to differ in whether they think of time as stationary or moving, limited or open-ended, horizontal or vertical, oriented from left to right, right to left, front to back, back to front, east to west, and so on. In this paper we will focus on the representations of time in English and Mandarin speakers, and we examine the role of spatiotemporal metaphors and writing direction in shaping people's representations of time.

The question of whether English and Mandarin speakers differ in their representations of time has attracted much attention and controversy. Studies relying on one experimental paradigm have produced inconsistent results (e.g., Boroditsky, 2001; Chen, 2007; January & Kako, 2007; Liu & Zhang, 2009; Tse & Altarriba, 2008). Limitations of this paradigm are discussed in Boroditsky et al. (2010).[1] However, studies drawing on other empirical paradigms have revealed consistent differences between how English and Mandarin speakers tend to spatialize time (e.g., Boroditsky et al., 2010; Chan & Bergen, 2005; Miles, Tan, Noble, Lumsden, & Macrae, 2011).

Here, we aim to confirm and extend these previous findings with two new studies designed to assess how time is spatialized in three-dimensional space by speakers of English and Mandarin. Further, we examine what aspects of linguistic experience predict differences in temporal thinking.

Experiment 1 measures implicit space-time associations in English and Mandarin speakers along three axes: transverse (left/right), vertical (up/down), and sagittal (front/back). Experiment 2 uses an explicit spatial pointing task to extend and confirm these results, and to more closely examine the contributions of linguistic context and experience in influencing bilinguals' representations of time.

One conceptual innovation in previous work on this question was to test for effects of language on thought in bilinguals (Boroditsky, 2001). Boroditsky (2001) compared native English speakers and Mandarin–English (ME) bilinguals on the same task, with both groups tested in English. This kind of comparison allows for an exact match in experimental method used to compare two groups, without worry about whether differences in behavior may arise because of subtle differences in instructions or stimuli presented in different languages. In Experiment 2, we test ME bilinguals in both Mandarin and English, allowing us to separate out the effects of language of test and of previous linguistic experience, and to examine whether one or the other or both affect thinking. Further, Boroditsky (2001) compared data from ME bilinguals who had had different relative amounts of English and Mandarin experience. In Experiment 2, we replicate this logic to test whether the degree of proficiency in Mandarin and the amount of experience with different writing systems predicts how people organize time.

1. Space-time metaphors

Both English and Mandarin use horizontal front/back spatial terms to talk about time. In English, we can look forward to the good times ahead, or think back to travails past and be glad they are behind us. In Mandarin, front/back spatial metaphors for time are also common (Chun, 1997a; Chun, 1997b; Liu & Zhang, 2009; Scott, 1989; Zhang & Ding, 2003; Zhu, 2006). For example, Mandarin speakers use the spatial morphemes qián (''front'') and hòu (''back'') to talk about time.

Unlike English speakers, Mandarin speakers also systematically and frequently use vertical metaphors to talk about time (Chun, 1997a; Chun, 1997b; Liu & Zhang, 2009; Scott, 1989; Zhang & Ding, 2003; Zhu, 2006). The spatial morphemes shàng (上 ''up'') and xià (下''down'') are used to talk about the order of events, weeks, months, semesters, and more. Earlier events are said to be shàng or ''up,'' and later events are said to be xià or ''down.'' For example, ''shàng ge yuè'' (traditional: 上個月]/simplified: 上个月) is last (or previous) month, and ''xià ge yuè'' (traditional: 下個月/simplified: 下个月) is next (or following) month. Chen (2007) finds that a full 36% of spatial metaphors for time are vertical.

For our purposes, the key linguistic observation is that vertical metaphors are more frequent in Mandarin than they are in English. Although in English vertical spatial terms can also be used to talk about time (e.g., ''hand down knowledge from generation to generation''), these uses are not nearly as common or systematic as is the use of shàng and xià in Mandarin. The difference between the productivity and frequency of vertical metaphors between the two languages has been noted by a large number of scholars, including many in China, Taiwan, and Hong Kong (e.g., Chun, 1997a; Chun, 1997b; Liu & Zhang, 2009; Scott, 1989; Zhang & Ding, 2003; Zhu, 2006). This linguistic difference offers the prediction that Mandarin speakers would be more likely to conceive of time vertically than would English speakers.

2. Writing direction

In addition to differences in metaphors, Mandarin and English also differ in orthography. Patterns in orthography have been found to influence people's representations of time (Fuhrman & Boroditsky, 2010; Ouellet et al., 2010; Tversky et al., 1991). For example, English speakers (who read and write text written from left to right) tend to arrange time from left to right and associate earlier times with the left side of space, whereas Hebrew speakers (who read and write text arranged from right to left) also arrange time from right to left and associate earlier times with the right side of space (Fuhrman & Boroditsky, 2010; Tversky et al., 1991).

Traditionally, Chinese text was written in vertical columns arranged from right to left. Within the last century there has been a switch to writing in horizontal rows from left to right (same as in English). In the People's Republic of China the official switch occurred in 1956, and text in newspapers, books, and online is now nearly always arranged horizontally from left to right. In Taiwan, vertical writing has remained common, though official documents have been required to be written horizontally from left to right since 2004.

The fact that writing direction seems to influence representations of time in English and in other languages proposes an important further question for comparing English and Mandarin speakers' representations of time. If Mandarin speakers do think about time vertically more than English speakers do, is this difference due to experience speaking Mandarin (and using vertical time metaphors), or due to experience with reading and writing text arranged in vertical columns?

Because of the switch between vertical and horizontal writing in Chinese (and the geographical differences in the timing of the switch), different Mandarin speakers have different amounts of exposure to text arranged vertically; some have a lot of experience reading text in vertical columns and some never read this way. This variability in the population allows us to equate groups of English and Mandarin speakers on their experience with reading and writing vertical text, and test whether differences in time representations between English and Mandarin speakers might persist even when experience with vertical text is accounted for. To achieve such a comparison in Experiment 1, we only included Mandarin-speaking participants who reported reading and writing text arranged in horizontal rows from left to right (as in English).

3. Experiment 1

In Experiment 1, we use a non-linguistic implicit association task to measure English and Mandarin speakers' space-time associations. The task is modeled on the design used in Fuhrman and Boroditsky (2010) and Boroditsky et al. (2010). This task offers several advantages. First, the task is non-linguistic (the stimuli are photographs and the responses are button-presses; the task does not require participants to produce or process any language). Second, the task relies on reaction time (an implicit measure of processing that participants are unlikely to manipulate to please the experimenter). Third, the task tests temporal reasoning across a wide range of temporal progressions and durations. Finally, the study was designed to measure and distinguish space-time associations along each of the three major axes (transverse, sagittal, and vertical), allowing us to capture how time is spatialized in three-dimensional space.

3.1. Methods

3.1.1. Participants

Fifty-nine Stanford students and 75 students at Shanghai International Studies University and Shanghai University of Finance and Economics in Shanghai, China, participated in this study in exchange for course credit or payment. All participants completed a language background questionnaire. On the questionnaire participants report all of the languages they have been exposed to (with biographical information re: age of acquisition, countries lived in, etc.) and assess their own fluency in each of these languages on a scale from 1 to 5 (1 = not at all fluent, 5 = completely fluent). People also report their experience with reading and writing, in particular whether they read and write text arranged in each of the following four ways: rows from left to right, rows from right to left, columns from right to left, and columns from left to right. To be certain that people are thorough in reporting their reading and writing experience, the questionnaire asks separately about experience in each of the following six categories: reading books, reading magazines, reading newspapers, reading Internet-based text, typing text, and writing text by hand.

The students tested at Stanford were native English speakers, and all reported their proficiency in English to be 5 out of 5. None reported having any exposure to Mandarin. The students tested in Shanghai were all native Mandarin speakers, and all reported their proficiency in Mandarin to be 5 out of 5. In addition, none of them reported a proficiency level in English of above 4 out of 5, and none of them had lived in an English-speaking country. All of the participants tested in Shanghai reported reading and writing from left to right in all six categories (books, magazines, newspapers, Internet-based text, typing, and writing text by hand). This reflects the fact that the writing system in the People's Republic of China has switched to the left-to-right system over the last 60 years as discussed earlier.

3.1.2. Materials

Materials used in this study consisted of 168 images. The images belonged to 56 temporal progression themes (e.g., a banana being eaten, Julia Roberts at different ages, buildings from different eras). Within each theme there was an ''early,'' ''middle,'' and ''late'' time point. For example, in the ''banana being eaten'' theme, the ''early'' picture showed a whole banana, the ''middle'' picture a half-peeled banana, and the ''late'' picture—just the peel. The themes included a range of temporal intervals, from events that last only a few seconds (e.g., filling a cup of coffee), to intervals that spanned decades (e.g., a person at different stages in life).

3.1.3. Procedure

Participants were tested individually in a quiet testing room, all using the same MacBook laptop computer. Both English and Mandarin speakers completed the same experimental task, but read instructions in their native language at the beginning of each testing block. Instructions were verified and fine-tuned with the help of a native Mandarin-speaking research assistant in Shanghai.

On each trial, a fixation cross was presented in the middle of the screen, and participants were instructed to press the middle (blue) key to start the trial. Then, the picture showing the ''middle'' time point from one of the 56 sequences (e.g., the half-peeled banana) appeared in the center of the screen for 1500 ms, followed by either a picture of the whole banana or the empty banana peel. Participants were instructed to decide whether the second picture presented showed a conceptually earlier or later time point than the first picture.

Participants responded by pressing one of three adjacent keys on a USB keypad (see Appendix A). The middle key was masked with a blue sticker and participants were instructed to press it as soon as they saw the fixation cross on the screen, to start the trial. The key on one side of the blue key was designated ''earlier'' and masked with a black sticker, and the key on the other side of the blue key was masked with a white sticker and designated ''later.'' The keys were not labeled linguistically. The keypad was mounted on a rotating ball head (from a tripod mount), which allowed it to be rotated to align with the three different axes (such that the keys could be oriented left to right, top to bottom, or near to far, etc.). For both horizontal axes, the keypad was oriented so that it was parallel to the tabletop. For the vertical axis, the keypad was oriented vertically, with the keys facing the participants.

All participants completed six key-mapping blocks, each consisting of 56 trials: two blocks on the transverse axis (one with the left response key designated as ''earlier'' and the right key as ''later'' and one with the reverse key mapping), two blocks on the vertical axis (top key as ''earlier''/bottom key as ''earlier''), and two on the sagittal axis (near key as ''earlier''/far key as ''earlier'').

The order of the blocks was counterbalanced across participants in six possible order conditions, such that two blocks of the same axis (e.g., ''left is earlier'' and ''left is later'') never followed each other, and never appeared as the first and last blocks in the same condition. Each of the ''earlier'' and ''later'' pictures of every temporal theme appeared once in every axis, such that participants saw the ''earlier'' picture in one block (e.g., ''top is earlier'') and the ''later'' picture of the same sequence in the remaining block of the same axis (e.g., ''top is later''). Assignment of pictures to blocks was counterbalanced across participants.

The first block started with 10 randomized practice trials, and each of the following five blocks started with only five randomized practice trials. Participants received feedback about their performance during the practice block, but not during the rest of the experiment. The items used in the practice trials were not used subsequently in the testing blocks.

3.1.4. Inclusion criteria

Responses of four participants (three English speakers and one Mandarin speaker) were discarded from analysis due to exceptionally high overall response times (more than 2 *SD* away from the language group mean). Error responses were not included in the analysis. Accuracy rate in the responses of the included participants was 94.6% for the English speakers and 89.4% for the Mandarin speakers. There were main effects of both accuracy and reaction time by testing group, with the Stanford testing group responding faster and more accurately overall. It is likely that this overall difference in performance is due simply to differences in familiarity with participating in psychological studies. Participants in the Stanford testing group are part of the Stanford Psychology testing pool and so have had much more experience completing these kinds of tasks. Finally, we excluded responses that were egregious reaction-time outliers (more than 10 *SD* away from the overall mean). These made up only 0.27% of total correct responses. All remaining responses were submitted for analysis.

3.2. Results

Results of interest are plotted in Fig. 1, and descriptive statistics are shown in Tables 1 and 2. To summarize, English speakers were fastest to answer questions about temporal order when responses were arranged horizontally from left to right. Mandarin speakers, on the other hand, were fastest to answer the same questions when responses were arranged vertically from top to bottom.

To examine interactions between native language and performance in the task broadly, we fit a multilevel logit model using Laplace Approximation implemented in the statistical software R in the lmer() function within the lme4 analysis package (Bates, Maechler, & Dai, 2008; R Development Core Team, 2008). This method allows us to model both participants and items as random effects in the analysis and to compare the explanatory strength of models that include or do not include interactions by native language. The full analysis modeled reaction time with key mapping (six levels) and native language (English or Mandarin) as fully crossed fixed effects and participants and items as random effects. The

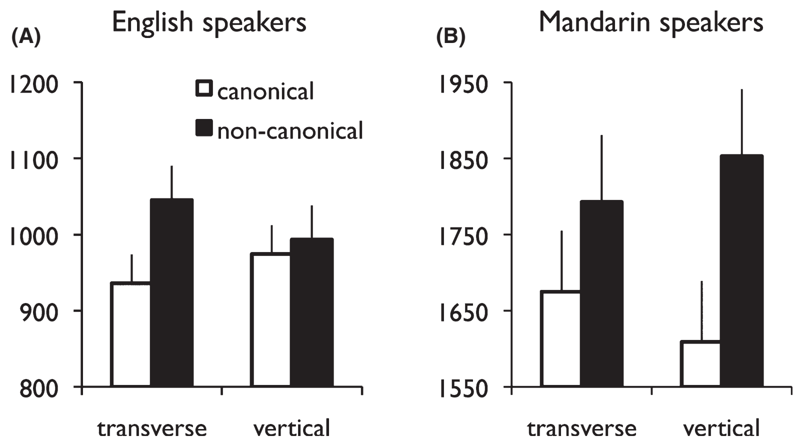

Fig. 1. Results of Experiment 1: Native English and native Mandarin speakers' response times. The figure plots by participants' mean response times (in ms). The error bars represent standard error. Responses along the transverse (left/right) and vertical axes are plotted (results on the sagittal axis yielded no reliable patterns and so are not included in this figure). For the transverse axis, the canonical bars show response times in the left-is-earlier key mapping, and the non-canonical bars show response times in the right-is-earlier key mapping. For the vertical axis, the canonical bars show response times in the top-is-earlier key mapping and the non-canonical bars show response times in the bottom-is-earlier key mapping.

Table 1
Results of Experiment 1: English and Mandarin speakers' average response times (in ms) for each of the six key-mapping blocks. Asterisks mark the fastest key mapping for each language group

	English	Mandarin
Left is earlier	* 936*	1675
Right is earlier	1045	1793
Near is earlier	1015	1745
Far is earlier	983	1666
Top is earlier	974	* 1609*
Bottom is earlier	993	1853

reduced model included all the same factors but did not model interactions between key mapping and native language. A comparison of the two models confirmed that including interactions between key mapping and native language significantly improved our ability to predict participants' reaction times, $\chi^2(5) = 67.19$, $p < .00000001$. The chi-square test serves as a measure of how much the model is improved by including the interaction terms.

To examine interactions between language, axis, and the canonicality of key mappings more specifically, we constrained our analysis to the transverse and vertical axes. On the transverse and vertical axes, there are clear predictions from writing direction (left is earlier) and from patterns in metaphor (in Mandarin, up is earlier) about which key mapping should

Table 2
Results of Experiment 1: English and Mandarin speakers' average accuracy for each of the six key-mapping blocks (shown as proportion correct)

	English	Mandarin
Left is earlier	0.95	0.89
Right is earlier	0.93	0.87
Near is earlier	0.95	0.89
Far is earlier	0.95	0.90
Top is earlier	0.95	0.91
Bottom is earlier	0.95	0.91

be easier. However, for the sagittal axis there is not a clear prediction. On the one hand ''the future is ahead of us,'' so near events might be earlier and far events later (provided all are in the future). On the other hand, in reading and writing, earlier elements are further away from us on the page, and later elements are nearer. For this reason, we conducted a further analysis, restricting the data set to only the transverse and vertical axes.

The prediction was that there would be a three-way interaction between axis, canonicality, and native language such that on the transverse axis both language groups would show effects of canonicality, but on the vertical axis only Mandarin speakers would. As before we fit and compared two multilevel logit models. The full analysis modeled reaction time with axis (transverse or vertical), canonicality (up or left is earlier vs. down or right is earlier) and native language (English or Mandarin) as fully crossed fixed effects and participants and items as random effects. The reduced model included all the same factors except the three-way interaction between axis, canonicality, and native language. A comparison of the two models confirmed that including the three-way interaction between axis, canonicality, and native language significantly improved our ability to predict participants' reaction times, $\chi^2(1) = 31.64$, $p < .10^{-7}$. This analysis again confirms that participants' native language affected their ability to map temporal responses onto different spatial arrangements.

To examine the results in more detail, we conducted 2 × 2 repeated measures ANOVAS (2 language × 2 key mapping) for each axis separately.

3.2.1. Transverse axis

Overall, participants responded faster when the ''earlier'' response was on the left than when it was on the right, as confirmed by a main effect of key mapping ($F(1, 55) = 34.0$, $p < .0001$ by items; $F(1, 128) = 4.95$, $p < .05$ by participants). Planned paired t tests showed that English speakers were indeed faster to respond when the earlier response was on the left than when it was on the right ($t(56) = 7.96$, $p < .0001$ by items; $t(56) = 2.34$, $p < .05$ by participants). For Mandarin speakers, the effect went in the same direction, but it was only reliable in the by-items contrast ($t(56) = 3.21$, $p < .01$ by items; $t(74) = 1.15$, $p = .15$ by participants). The effects of key mapping in the two groups did not significantly differ from one another: There was no interaction between language and key mapping ($F(1, 55) = 0.25$, $p = .62$ by items; $F(1, 128) = 0.007$, $p = .93$ by participants).

To confirm the results another way, we fit a multilevel logit model as described before, modeling both participants and items as random effects in the analysis, and comparing the explanatory strength of models that include or do not include interactions by native language. The full analysis modeled reaction time with key mapping (two levels) and native language (English or Mandarin) as fully crossed fixed effects and participants and items as random effects. The reduced model included all the same factors but did not model interactions between key mapping and native language. A comparison of the two models confirmed that including interactions between key mapping and native language did not significantly improve our ability to predict participants' reaction times, $\chi^2(1) = 0.33$, $p = .565$.

3.2.2. Vertical axis

English and Mandarin speakers showed different patterns of response on the vertical axis. Overall, participants responded faster when the ''earlier'' response was on top than when it was on the bottom, as confirmed by a main effect of key mapping ($F(1, 55) = 84.4$, $p < .0001$ by items; $F(1, 128) = 5.49$, $p < .05$ by participants). However, this effect was driven by the responses of the Mandarin speakers, as confirmed by a significant interaction between language and key mapping ($F(1, 55) = 64.8$, $p < .0001$ by items; $F(1, 128) = 4.01$, $p < .05$ by participants). Planned paired t tests showed that Mandarin speakers were indeed faster to respond when the earlier response was on top than when it was on the bottom ($t(56) = 9.84$, $p < .0001$ by items; $t(74) = 2.70$, $p < .01$ by participants), but English speakers showed no such difference ($t(56) = 0.62$, $p = .53$ by items; $t(56) = 0.03$, $p = .98$ by participants).

To confirm the main results of interest another way, we fit a multilevel logit model as described before, modeling both participants and items as random effects in the analysis, and comparing the explanatory strength of models that include or do not include interactions by native language. The full analysis modeled reaction time with key mapping (two levels) and native language (English or Mandarin) as fully crossed fixed effects and participants and items as random effects. The reduced model included all the same factors but did not model interactions between key mapping and native language. A comparison of the two models confirmed that including interactions between key mapping and native language significantly improved our ability to predict participants' reaction times, $\chi^2(1) = 56.09$, $p < .10^{-13}$.

3.2.3. Sagittal axis

Effects on the sagittal axis revealed a less clear pattern. Overall, participants responded less quickly when the ''earlier'' response was nearer to the body than when it was further away, but this contrast was not reliable in the by-participants analysis (main effect of key mapping $F(1, 55) = 33.7$, $p < .001$ by items; $F(1, 128) = 0.94$, $p = .33$ by participants). The effect was stronger in the Mandarin-speaking group, but this contrast again was not reliable in the by-participants analysis (interaction between language and key mapping $F(1, 55) = 5.4$, $p < .05$ by items; $F(1, 128) = 0.17$, $p = .68$ by participants).

To examine the main results of interest another way, we fit a multilevel logit model as described before, modeling both participants and items as random effects in the analysis, and comparing the explanatory strength of models that include or do not include interactions by native language. The full analysis modeled reaction time with key mapping (two levels) and native language (English or Mandarin) as fully crossed fixed effects and participants and items as random effects. The reduced model included all the same factors but did not model interactions between key mapping and native language. A comparison of the two models showed that including interactions between key mapping and native language did slightly improve our ability to predict participants' reaction times as seen in the other analyses (but not overwhelmingly so), $\chi^2(1) = 4.19$, $p < .05$.

Analyses of accuracy patterns confirmed that none of the effects found in reaction time resulted from speed-accuracy trade-offs. Accuracy did not vary significantly by key mapping along any of the axes in either language group (all $p > .34$). Accuracy results are shown in Table 2.

3.3. Discussion

Results show a difference in implicit space-time associations that is consistent with patterns in language and culture. On the transverse axis, both English and Mandarin speakers favored a left-to-right representation of time, consistent with writing direction for both groups. However, Mandarin speakers also showed evidence for a vertical time mapping, favoring a top-to-bottom representation of time, consistent with the set of vertical spatial metaphors for time in Mandarin. English speakers did not show a preference for time orientation along the vertical axis.

We also tested both English and Mandarin speakers along the sagittal axis, but neither group showed a robust time orientation preference in this axis. It is possible that patterns in language and culture provide conflicting cues to laying out time on this axis (see discussion earlier). Another possibility is that the sagittal axis is more likely to be used for thinking about time in the case of deictic representations (points in time with respect to now) or when thinking about biographical time. The photographs we used depicted time points that in many cases were not meaningfully temporally bound to the current moment. It is possible that this feature of the stimuli precluded people from using their existing representations of time along the sagittal axis. Yet another possibility is that the response task we used is not well tuned to test time representations on the sagittal axis: contrary to the common front-back or back-front linguistic metaphor, our paradigm asked participants to map both *earlier* and *later* onto points that were in front of them (near and far relative to their bodies). Of course a final possibility is that English and Mandarin speakers simply do not use the sagittal axis for representing time. This final possibility appears unlikely, as much previous evidence in other tasks does find evidence for front/back time representations (e.g., Miles et al., 2010).

4. Experiment 2

In Experiment 2 we explored what aspects of language experience or context are important in shaping how people organize time. We compared the responses of ME bilinguals with different degrees of proficiency in Mandarin. Further, we compared ME bilinguals tested in English or in Mandarin. Finally, we compared participants tested in California and in Taiwan. Does the propensity to organize time vertically depend on long-term aspects of language experience (e.g., proficiency in Mandarin), or the situational availability of a language (e.g., whether one is tested in an English or Mandarin-speaking situation), or both? In Experiment 2 we adapted the design used in Experiment 1 of Fuhrman and Boroditsky (2010) to answer these questions.

4.1. Methods

4.1.1. Participants

In total, 377 people participated in this study. Participants reported their level of proficiency in English and Mandarin (0 = do not speak at all, 5 = fluent), and whether they ever read text arranged in vertical columns (as in traditional Chinese writing). Of the total, 330 participants were tested in English at Stanford University or Foothill Community College. Of these, 134 were native English speakers who spoke no Mandarin ($M_{age} = 19.1$, $SD = 1.79$; mean Mandarin proficiency = 0, $SD = 0$; mean English proficiency = 5, $SD = 0$). The other 196 were ME bilinguals representing a range of proficiency in both English and Mandarin ($M_{age} = 23.7$, $SD = 5.7$; mean Mandarin proficiency = 4.52, $SD = 0.90$; mean English proficiency = 4.24, $SD = 0.95$). An additional 32 ME bilinguals were tested in Mandarin outside of the Cupertino Public Library or outside of 99 Ranch (a supermarket specializing in Asian foods) in Cupertino ($M_{age} = 36.3$, $SD = 15.9$; mean Mandarin proficiency = 4.63, $SD = 0.74$; mean English proficiency = 3.53, $SD = 1.09$). Finally, 15 ME bilingual Taiwan residents were tested in Mandarin in Taiwan ($M_{age} = 22.9$, $SD = 15.1$; mean Mandarin proficiency = 5, $SD = 0$; mean English proficiency = 3.53, $SD = 1.16$). The experimenters conducting the study in Mandarin were native speakers of Mandarin.

4.1.2. Procedure and materials

All participants were tested on the same three sets of questions about time. An experimenter stood next to the participant (such that experimenter and participant were facing in the same direction), used his or her hand to select a spot in space directly in front of the participant (about a foot in front of the chest, with the palm facing up and the fingers brought together into a cone resembling the Italian ''che vuoi?'' gesture), and said (for example), ''If this here is TODAY, where would you put YESTERDAY?'' The experimenter waited for the participant to point. After the participant pointed, the experimenter asked, ''And where would you put TOMORROW?'' and again waited for the participant to point.

After the participant answered both questions, the experimenter discretely marked down the responses on a coding sheet, marking whether each point was above, below, to the left, to the right, in front of, or behind the reference point. In mixed direction cases, for example, if participant pointed to a spot that was equally above and to the left of the reference, this was noted as .5 above and .5 left.[2]

Three sets of time points were tested with each participant (the same order was used for all participants). In addition to Yesterday/Today/Tomorrow, participants were also asked to locate breakfast and dinner with respect to lunch, and the last and next month with respect to the current month. The question about months of the year was asked using month names. For example, for a participant tested in September, the experimenter would ask, ''If this here is September, where would you put August? And where would you put October?'' The central time point was always mentioned first (e.g., today), followed by the other two time points either with the earlier (e.g., yesterday) mentioned first or the reverse. No spatial language was used to describe the time points in either English or Mandarin (e.g., participants were not asked which is the ''up month'' or the ''down month'' in Mandarin). The stimuli used in Mandarin were as follows: 午餐 (wǔ cān) lunch, 早餐 (zǎo cān) breakfast, 晚餐 (wǎn cān) dinner; 今天 (jīn tiān) today, 昨天 (zuó tiān) yesterday, 明天 (míng tiān) tomorrow; 九月 (jiǔ yuè) September, 八月 (bā yuè) August, 十月 (shí yuè) October (and other month names as appropriate).

4.2. Results

Results and analyses are shown in Tables 3–5. Figs. 2 and 3 illustrate the main findings of interest. In brief, English speakers predominantly arranged time horizontally from left to right. Mandarin speakers likewise often arranged time from left to right. Unlike English speakers, Mandarin speakers also often arranged time vertically from top to bottom (more so than English speakers, and more if they were more proficient in Mandarin or were tested in Mandarin).

To quantify the data, we coded the linear trajectory for each arrangement in terms of six main directions (rightward, leftward, upward, downward, forward, and backward). The codings were quantified by assigning each of the six possible main directions a value from 0 to 1, with the sum of the six directions adding up to 1. Some example codings: If the participant put yesterday to the left of today and tomorrow to the right of today (a rightward arrangement), the directionality coding for that trial would be 1 for the rightward direction, and 0 for all the others. If the participant put yesterday about equally below and to the right

Table 3
Results of Experiment 2 shown separately for different groups of participants

	English Speakers	ME Bilinguals	ME Bilinguals	ME Bilinguals	ME Bilinguals
Mandarin skill	None	Low (3 or less)	High (4 or 5)	High (4 or 5)	High (4 or 5)
Language of test	English	English	English	Mandarin	Mandarin
Country of test	USA	USA	USA	USA	Taiwan
Rightward	0.88	0.78	0.64	0.38	0.47
Leftward	0.06	0.06	0.09	0.08	0.02
Downward	0.02	0.02	0.13	0.39	0.33
Upward	0.01	0.02	0.03	0.04	0.11
Forward	0.03	0.12	0.08	0.06	0.04
Backward	0.01	0	0.04	0.05	0.02
N	134	26	170	32	15

Table 4
Results of Experiment 2 shown separately for the subset of Mandarin–English bilinguals who reported never reading text arranged in vertical columns and the subset of Mandarin–English bilinguals who reported sometimes reading text arranged in vertical columns

	ME Bilinguals	ME Bilinguals
Read vertically?	Never	Sometimes
Rightward	0.66	0.56
Leftward	0.07	0.09
Downward	0.15	0.19
Upward	0.02	0.04
Forward	0.06	0.10
Backward	0.04	0.03
N	125	118

Table 5
Linear regression model fits for Experiment 2. The table shows the role of five predictors (rows) in six models (columns). The cells report standardized beta coefficients, *t* values, and *p* values. Overall ANOVA results and adjusted *R*-squared are reported in the bottom rows

		Rightward	Leftward	Downward	Upward	Backward	Forward
Mandarin proficiency	beta	−0.21	0.04	0.18	−0.02	0.16	0.05
	t	−3.33	0.61	2.82	−0.25	2.33	0.73
	p	0.001	0.542	0.005	0.800	0.020	0.466
English proficiency	beta	0.08	−0.05	−0.04	−0.05	−0.05	0.00
	t	1.46	−0.78	−0.69	−0.87	−0.83	0.03
	p	0.145	0.437	0.488	0.387	0.410	0.975
Language of test (Eng or Man)	beta	−0.25	−0.02	0.39	0.07	0.03	−0.03
	t	−4.25	−0.36	6.68	1.16	0.47	−0.44
	p	0.001	0.720	0.001	0.247	0.641	0.663
Country of test (USA or Taiwan)	beta	0.10	−0.07	−0.07	0.11	−0.04	−0.06
	t	1.67	−1.04	−1.21	1.75	−0.61	−0.88
	p	0.095	0.300	0.225	0.081	0.544	0.382
Experience with vertical text	beta	−0.12	0.04	0.02	0.11	−0.08	0.15
	t	−2.20	0.69	0.43	1.73	−1.29	2.37
	p	0.029	0.489	0.669	0.085	0.196	0.018
ANOVA	*F*	19.10	0.88	21.00	4.32	2.41	2.13
	p	0.001	0.494	0.001	0.001	0.036	0.061
Adjusted R^2		0.196	−0.002	0.211	0.043	0.019	0.015

of today and put tomorrow about equally above and to the left of today, the coding would be .5 leftward, .5 upward, and 0 for all the others. If an arrangement could not be interpreted as having a linear order, it was coded as not having a direction, and not included in analyses (this was true for only 0.3% of all arrangements).

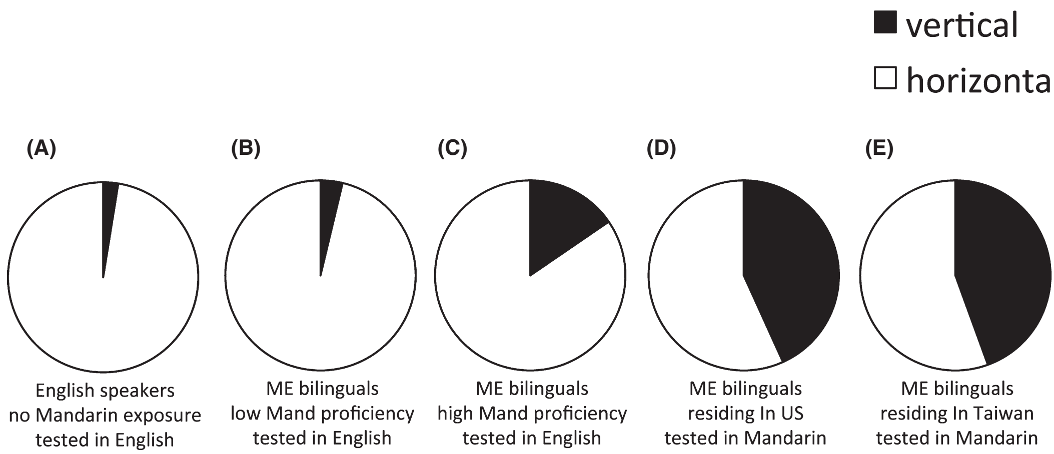

Fig. 2. Results of Experiment 2 shown separately for different groups of participants. (A) English speakers tested in English (participants with no exposure to Mandarin) (N = 134). (B) Mandarin–English bilinguals with low Mandarin proficiency (3 or less) tested in English (N = 26). (C) Mandarin–English bilinguals with high Mandarin proficiency (4 or more) tested in English (N = 170). (D) Mandarin–English bilinguals tested in Mandarin in the United States (N = 32). (E) Mandarin–English bilinguals tested in Mandarin in Taiwan (N = 15). The graphs plot all horizontal (leftward, rightward, forward, and backward) responses and all vertical (upward and downward) responses made by participants in each group.

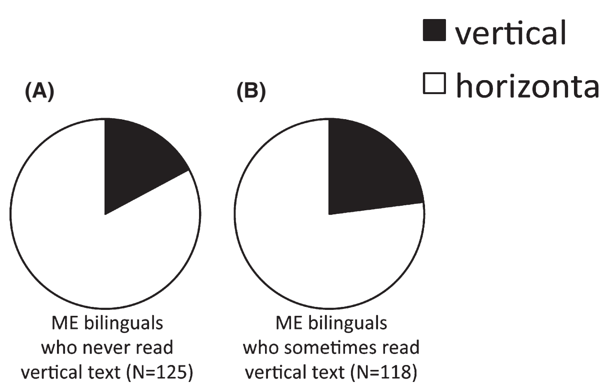

Fig. 3. Results of Experiment 2 shown separately for different groups of participants. (A) Data from the subset of Mandarin–English bilinguals who reported never reading text arranged in vertical columns (N = 125). (B) Data from the subset of Mandarin–English bilinguals who reported sometimes reading text arranged in vertical columns (N = 118). The graphs plot all horizontal (leftward, rightward, forward, and backward) responses and all vertical (upward and downward) responses made by participants in each group.

To analyze the data, we fit linear regression models for each of the six main directions (as coded above, by participants) with the following five factors as predictors: (a) proficiency in Mandarin (0–5); (b) proficiency in English (0–5); (c) language of test (English or Mandarin); (d) country of test (United States or Taiwan); and (e) experience with vertical text (whether participants reported ever reading text in vertical columns).[3] This set of five predictors captured a significant proportion of the variance in four of the six models (rightward, upward, downward, and backward arrangements). The full results of these regressions are reported in Table 5. Tests for multicollinearity between predictor variables revealed no violations (VIF values of 1.82, 1.51, 1.61, 1.50, 1.46 for the five predictors in Table 5 respectively).

To summarize, proficiency in Mandarin was a powerful independent predictor of time arrangements along all three axes (transverse, vertical, and sagittal), even when experience with vertical writing was controlled for. Participants who were more proficient in Mandarin were more likely to arrange time as proceeding downward, less likely to arrange time as proceeding rightward, and more likely to arrange time as proceeding backward than those who were less proficient in Mandarin.[4] Language of test was also an important independent predictor: Participants tested in Mandarin were more likely to arrange time as proceeding downward and less likely to arrange time as proceeding rightward than those tested in English. Finally, experience with vertical text had an effect on rightward arrangements but did not predict vertical arrangements for time. Participants who reported at least sometimes reading text arranged in vertical columns were less likely to make rightward arrangements for time (likely because vertical columns in Mandarin are typically arranged from right to left), but they did not reliably differ in vertical arrangements from participants who never read vertical text.

The four models that yielded significant results are discussed below.

Rightward time: Mandarin proficiency, language of test, and experience with vertical text all emerged as significant independent predictors of rightward time arrangements. Participants who were more proficient in Mandarin, who were tested in Mandarin, or who had experience with reading or writing vertical text were less likely to lay out time as proceeding from left to right. The effect of writing direction here is likely due to the fact that vertical columns in traditional Chinese writing are arranged from right to left, the opposite direction of English writing. The effects of Mandarin proficiency and language of test were stronger than that of experience with vertical writing (standardized betas of −.21, −.25, and −.12, respectively).

Downward time: Mandarin proficiency and language of test emerged as significant independent predictors of downward time arrangements. Participants who were more proficient in Mandarin or who were tested in Mandarin were more likely to lay out time as proceeding from top to bottom. Experience with vertical writing was not a significant predictor. The effects of Mandarin proficiency and language of test were stronger than that of experience with vertical writing (standardized betas of .18, .39, and .02, respectively).

Upward time: While the five predictors together explained a significant amount of variance in this model, the adjusted R-squared was rather modest (.043), and none of the predictors emerged as significant on their own.

Backward time: The adjusted R-squared for this model was very modest (.019), and only Mandarin proficiency emerged as a significant predictor. Participants who were more proficient in Mandarin were more likely to lay out time as proceeding from front to back (with the past in front)[4].

Proficiency in English and country of test did not emerge as significant predictors of the directionality of time arrangements in any of the models.

To verify that the language of test independently affected the bilinguals' responses, we repeated the same set of regressions, this time including only ME bilinguals (using the same five predictors as before). The five predictors together captured a significant proportion of the variance in three of the six models: rightward (adjusted $R^2 = .102$, $F = 6.43$, $p < .001$), upward (adjusted $R^2 = .030$, $F = 2.47$, $p < .05$), and downward (adjusted $R^2 = .142$, $F = 8.89$, $p < .001$). As before, bilinguals tested in Mandarin were significantly less likely to make rightward time arrangements (standardized beta = −.294, $t = -3.94$, $p < .001$), and significantly more likely to make downward arrangements (standardized beta = .407, $t = 5.57$, $p < .001$) than bilinguals tested in English. Also as before, experience with vertical writing predicted fewer rightward time arrangements (standardized beta = −.130, $t = -2.034$, $p < .05$) and did not predict downward time arrangements (standardized beta = .022, $t = 0.353$, $p = .724$). No other predictors were individually significant in any of the other models.

4.3. Discussion

Results of Experiment 2 confirm and extend the findings of Experiment 1. Further, they demonstrate that both long-term and immediate aspects of language experience play an important role in how people spatialize time.

Participants with high proficiency in Mandarin were more likely to arrange time vertically (15.4%–44.4% depending on testing conditions) than were English speakers who had had no exposure to Mandarin (2.5%). Further, of the bilinguals highly proficient in Mandarin, those tested in Mandarin were more likely to arrange time vertically (43.2–44.4% depending on testing conditions) than those tested in English (15.4%). It appears that experience with language has both long-term and online effects on people's representations of time. These results confirm the original conclusions that speaking Mandarin makes one more likely to conceive of time vertically than speaking English.

While English speakers were overwhelmingly likely to arrange time horizontally (97.4% horizontal), the dominant direction of these horizontal arrangements was left to right (88%), not back to front (3.2%). The left to right pattern found here (and in many other studies discussed earlier) appears to be tied to writing direction. Participants who reported at least sometimes reading text arranged in vertical columns from right to left were less likely to arrange time as proceeding from left to right.

Mandarin speakers' propensity to arrange time vertically did not seem to depend on reading and writing experience in our sample. ME bilinguals who reported never reading vertical text and those who reported at least sometimes doing so were about equally likely to arrange time vertically (17.2% and 23.0%, respectively).

It is of course still possible that both experience with vertical time metaphors and experience with vertical writing make people more likely to think about time vertically, even though writing did not appear predictive in our sample. One possibility is that testing a wider range of age groups would include a wider range of experience with vertical writing, providing a more sensitive measure. It may also be that tasks that happen in table-space (rather than in unrestricted space around the body) or are otherwise more closely associated with the experience and practice of writing would be more likely to show effects of vertical writing experience on the arrangement of time (see e.g., Chan & Bergen, 2005).

5. General discussion

In two experiments we examined the representation of time in speakers of English and Mandarin. A number of interesting differences emerged. Results of Experiment 1 revealed similarities between English and Mandarin speakers along the transverse axis but differences along the vertical axis. Mandarin speakers' data revealed evidence for a vertical representation of time with earlier events represented above and later events below. This difference was observed in an implicit non-linguistic task, suggesting that culturally specific spatial representations for time are accessed automatically when people are making temporal judgments.

Results of Experiment 2 confirmed and extended these findings in a different paradigm. In Experiment 2, both English and Mandarin speakers arranged time horizontally, most often from left to right. In addition, participants who spoke Mandarin often produced vertical arrangements for time. The proportion of vertical representations of time depended on participants' proficiency in Mandarin and also on whether participants were tested in Mandarin or in English. Participants who were more proficient in Mandarin and participants who were tested in Mandarin were more likely to represent time vertically, suggesting that both long-term language experience and proximal linguistic context can shape people's representations of time. Differences in vertical time representations were not explained by differences in writing direction. Writing direction did emerge as an important predictor on the left/right axis; participants who always read text arranged horizontally from left to right were more likely to arrange time from left to right than participants who at least sometimes read text arranged in other ways.

Being able to observe effects from both long-term language experience and short-term linguistic manipulations (e.g., switching language of test) helps us understand more about the mechanisms through which language shapes thought. Temporary and chronic effects of language are not mutually exclusive. Indeed, chronic biases may themselves arise out of the sum of many temporary episodes; that is, many individual experiences may strengthen representations and make them available for a broader set of reasoning contexts. There is almost certainly continuity between short- and long-term effects of language on representations of time.

The existing evidence comparing English and Mandarin representations of time helps highlight this interplay between effects of long-term linguistic habits and effects of immediate linguistic context. To test whether experience in a particular culturo-linguistic environment creates long-term habits in thought, studies have tested ME bilinguals in English (e.g., Boroditsky, 2001; Boroditsky et al., 2010; Experiment 2 of this paper; Miles et al., 2011). In each of these studies, ME bilinguals were found to be more likely to think about time vertically than were native English speakers. These findings suggest habits of thought created while immersed in a Mandarin-speaking environment at least to some extent persist even when reasoning about time in an English-speaking context.

To test for effects of immediate context, researchers have compared the results from ME bilinguals when tested in Mandarin versus in English (Experiment 2 of this paper; Lai & Boroditsky, under review), when tested in a context with American versus Chinese cultural cues (Miles et al., in press), and when tested in Mandarin using vertical versus horizontal metaphors (Lai & Boroditsky, under review). In each of these cases, the data reveal that immediate context (whether it is language of test, cues to cultural context, or the particular metaphors being used in the moment) also plays an important role in guiding which representations of time people will be likely to activate for reasoning about time. Mandarin–English bilinguals are more likely to represent time vertically when tested in Mandarin (as opposed to English), when cued with photographs of a Chinese actor (as opposed to an American actor), or when processing vertical time metaphors in Mandarin (as opposed to horizontal metaphors in Mandarin).

Across these studies, there is a consistent pattern: Mandarin speakers are more likely to arrange time vertically than are English speakers, and the degree to which vertical representations are invoked depends on (a) how proficient people are in Mandarin; and (b) whether the current linguistic or cultural context primes more English-like or more Mandarin-like representations.

In this paper we have focused on the role of linguistic metaphor and cultural artifacts (writing direction) in guiding people's representations of time. Of course, time is spatialized in cultures not only in metaphor and writing direction but in a large number of other cultural artifacts: clocks, calendars, timelines, and so on. That is, experience with language is only one part of the differences in cultural experience between English and Mandarin speakers. In addition to differences in metaphor and writing considered here, English and Mandarin speakers may also have different experience with orientation of time on calendars and timelines. Patterns in linguistic and extra-linguistic aspects of culture often go together and mutually reinforce each other. Could it be that cross-group differences observed in all of these studies are indeed created entirely by differences in extra-linguistic cultural experience and not by patterns in language as suggested here? Results available to date offer a number of reasons to believe that experience with language per se is an important ingredient in shaping people's representations of time.

In this paper, we compared the relative influence of linguistic variables (proficiency in Mandarin, and language of test) with that of extra-linguistic variables (experience with vertical writing, country of test) in predicting vertical representations of time. The results were clear. The linguistic variables (Mandarin proficiency and language of test) strongly predicted vertical time orientation, whereas the extra-linguistic variables (experience with vertical writing, country of test) did not. At least in this case, linguistic information serves as a strong cue to culturally specific representations of time.

Another way to examine whether patterns in language per se can create some of the cross-group differences we have observed is to teach people a new pattern of metaphors and see whether this can induce a temporary shift in their representations of time. For example, Boroditsky (2001) trained English speakers to use vertical terms (above/below) to talk about time and found a corresponding shift in time representations. The effects of such quick in-lab training are similar to those of other immediate contextual cues (e.g., whether bilinguals are tested in one language or another). Of course, any such laboratory training will have only a temporary effect if the linguistic patterns learned are not reinforced and used continuously in a variety of contexts as they would be in natural language. Future work applying this logic to new non-linguistic experimental paradigms will help shed further light on the scope and durability of such linguistic training.

Another set of clues that linguistic metaphors per se play a role in guiding representations of time comes from studies that specifically manipulate which metaphors are used within a language and check for consequences in the resulting representations of time (e.g., Boroditsky, 2000; Gentner, Imai, & Boroditsky, 2002; Lai & Boroditsky, under review; McGlone & Harding, 1998). For example, McGlone and Harding (1998) primed people with either ego-moving (e.g., ''We are approaching the holidays'') or time-moving (e.g., ''The holidays are approaching'') metaphors in English and found that processing one or the other of these metaphors affected how people reasoned about a subsequent ambiguous temporal scenario. Lai and Boroditsky (under review) tested Mandarin speakers on the same time-pointing paradigm as described in Experiment 2 of this paper, but instead of using non-spatial terms in the temporal prompts, specifically used horizontal or vertical Mandarin metaphors (e.g., asking Mandarin speakers to arrange the up/down month or front/back month relative to now). The study asked whether Mandarin speakers flexibly re-organize time along the front-back or up-down axis depending on whether they are processing front-back or up-down metaphors for time. The fact that Mandarin uses both horizontal and vertical metaphors frequently allows an opportunity to ask this question. If spatial metaphors are simply fanciful ways of talking, then which metaphor is used should have no effect on how people arrange time. However, if such metaphors have psychological reality, then people should arrange time differently when processing different metaphors. The results revealed a strong effect of metaphor. Mandarin speakers were twice as likely to arrange time vertically when prompted with vertical metaphors as when prompted with horizontal metaphors. They were also more than twice as likely to arrange time sagitally (on the front-back axis) when prompted with front-back metaphors as when prompted with vertical metaphors. Interestingly, metaphors did not significantly affect arrangements along the transverse (left-right) axis (responses on this axis seem to be guided by experience with reading and writing text). These results strongly suggest that metaphors matter. In this case, metaphors have an in-the-moment effect on how people organize and imagine time.

Taken all together, these findings demonstrate that the metaphors we use to talk about time have both immediate and long-term consequences for how we conceptualize and reason about this fundamental domain of experience. How people conceptualize time appears to depend on how the languages they speak tend to talk about time and also on the particular metaphors being used to talk about time in the moment. Results of the two experiments reported in this paper are consistent with this growing body of work on representations of time in Mandarin, as well as with a broad body of evidence regarding the role of spatial metaphors in shaping temporal thinking, and with research exploring diversity in temporal thinking more broadly.

6. Conclusions

In this paper we presented evidence for one cross-cultural difference in temporal thinking: whether Mandarin speakers think about time vertically more than English speakers do. This pattern in thought is predicted by patterns in spatiotemporal metaphors; Mandarin speakers use vertical terms to talk about time more so than English speakers do. Converging evidence from two paradigms strongly suggests that Mandarin speakers also think about time vertically more often than English speakers do. It appears that patterns in language and culture can induce differences in thought in even such fundamental conceptual domains as time.

Notes

1. The new studies improve on the original design in a number of key ways. The original task collapsed the left-right and the front-back axes into one ''horizontal'' dimension, whereas the new studies treat these different axes separately. Second, the new studies separate the directions within the axes (so that a leftward spatial arrangement is treated differently from a rightward arrangement, not simply both as ''horizontal''). Further, in many of the new studies the task is non-linguistic (the stimuli are photographs and the responses are button-presses). Finally, the new studies test temporal reasoning across a wide range of temporal progressions and durations (previous work looked primarily at the order of months of the year).
2. Of the testing sessions conducted at Stanford and Foothill College, 237 were video-taped (the other participants did not give their permission to be video-taped), and the videos were later cross-checked with the experimenter's in-the-moment codes. The sessions conducted in the field (Cupertino and Taiwan) were coded in the moment (not video-taped). Because the behavior being coded in this study is an explicit evoked communicative point, not a subtle subconscious movement, it was possible to code people's points in the moment. While in-the-moment coding can open up the possibility of unconscious experimenter bias, we note that all such worries are ruled out in the implicit reaction-time task in Experiment 1. If the cross-linguistic differences observed in Experiment 2 were due only to experimenter bias, then one would not expect to find converging patterns in the data of Experiment 1. Instead, the two studies provide converging results. Further, experimenters were not aware of which dimensions of variation in the linguistic and cultural background of the participants would turn out to be of interest in analyses, and they did not collect this background information until after conducting the task.
3. Several other options for coding experience with writing direction are possible. It can be coded as proportion vertical (vs. horizontal), or proportion left to right (vs. right to left), or categorically as reading only horizontally left to right versus any other pattern. All of these codings produce the same pattern of results (no other effects of writing direction emerge as significant with any of these codings, nor do any of the other predictors lose significance).

4. The pattern of laying out time with the past in front and the future behind in Mandarin speakers is consistent with some patterns in the Mandarin language. For example, in Mandarin the ''front year'' is 2 years ago. This extra dimension of variation between the two languages is interesting and has been previously discussed in linguistic analyses of Mandarin. For space reasons, we do not focus on this dimension in this paper beyond simply noting here that our data are consistent with these previous linguistic analyses.

Acknowledgments

We are grateful to Harry F. Bahlman II for his ingenuity in designing and fabricating our experimental apparatus, and to Herbert H. Clark and the citizens of Cognation for helpful feedback on this manuscript and insightful discussions regarding this work.

References

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using s4 classes* [R package version 0.999375-28]. Available at: http://lme4.r-forge.r-project.org/.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28.

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

Boroditsky, L., Fuhrman, O., & McCormick, K. (2010). Do English and Mandarin speakers think about time differently? *Cognition*, *118*, 123–129. doi:10.1016/j.cognition.2010.09.010

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.

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, *13*, 185–189.

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

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S., Santiago-Diaz, I., Fotokopoulu, O., Pita, R., & Gil, D. (2004). How deep are effects of language on thought? Evidence from English, Greek, and Spanish speakers. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Chicago, IL.

Chan, T. T., & Bergen, B. (2005). Writing direction influences spatial cognition. *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society*.

Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition*, *104*(2), 427–436.

Chun, L. (1997a). A cognitive approach to UP metaphors in English and Chinese: What do they reveal about the English mind and the Chinese mind? Research degree progress report for Hong Kong Polytechnic University, pp. 125–140.

Chun, L. (1997b). Conceptualizing the world through spatial metaphors: An analysis of UP/DOWN vs. SHANG/XIA metaphors. *Proceeding of the 19th Annual Meeting of the Cognitive Science Society*, Mahwa, NJ: Erlbaum.

Clark, H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit non-linguistic task. *Cognitive Science*, *34*, 1430–1451.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space time metaphors. *Language and Cognitive Processes.*, *17*(5), 537–565.

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

Gonzalez, A., & Zimbardo, P. G. (1985, March). Time in perspective. *Psychology Today*, 21–26.

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

January, D., & Kako, E. (2007). Re-evaluating evidence for the linguistic relativity hypothesis: Response to Boroditsky (2001). *Cognition*, *104*(2), 417–426.

Ji, L., Guo, T., Zhang, Z., & Messervey, D. (2009). Looking into the past: Cultural differences in perception and representation of past information. *Journal of Personality and Social Psychology*, *96*(4), 761–769.

Lai, V., & Boroditsky, L. (under review). Spatiotemporal metaphors exert both immediate and chronic influence on people's representations of time: Examples from English and Mandarin. *Cognitive Linguistics*.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Liu, L., & Zhang, J. (2009). The effects of spatial metaphorical representations of time on cognition. *Foreign Language Teaching and Research*, *41*(4), 266–271.

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). The experiential link between spatial and temporal language. *Cognitive Science*, *29*, 655–664.

McGlone, M. S., & Harding, J. L. (1998). Back (or forward?) to the future: The role of perspective in temporal language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1211–1223.

Miles, L. K., Nind, L. K., & Macrae, C. N. (2010). Moving through time. *Psychological Science*, *21*(2), 222–223. DOI: 10.1177/0956797609359333

Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., & Macrae, C. N. (in press). Can a mind have two timelines? Exploring space-time mapping in Mandarin and English speakers. *Psychonomic Bulletin and Review*, *18*, 598–604.

Núñez, R., Motz, B., & Teuscher, U. (2006). Time after time: The psychological reality of the ego- and time-reference-point distinction in metaphorical construals of time. *Metaphor and Symbol*, *21*, 133–146.

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. (2010). Is the future the right time? *Experimental Psychology*, *57*(4), 308–314. doi: 10.1027/1618-3169/a000036

R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Development Core Team [ISBN 3-900051-07-0].

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

Scott, A. (1989). The vertical dimension and time in Mandarin. *Australian Journal of Linguistics*, *9*, 295–314.

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

Traugott, E. (1978). On the expression of spatiotemporal relations in language. In J. H. Greenberg (Ed.), *Universals of human language. Vol. 3: Word structure* (pp. 369–400). Stanford, CA: Stanford University Press.

Tse, C. S., & Altarriba, J. (2008). Evidence against linguistic relativity in Chinese and English: A case study of spatial and temporal metaphors. *Journal of Cognition and Culture*, *8*, 335–357.

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

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

Zhang, L., & Ding, C. (2003). Comparative study of temporal metaphor in English and Chinese. *Foreign Languages and Their Teaching*, *174*(9), 31–34.

Zhu, H. (2006) Spatio-temporal metaphor in English and Chinese. *Sino-US English Teaching*, *3*(11) (Serial No.35), 75–78.

Appendix A

The input device used in Experiment 1 shown in the six key-mapping configurations.