

Manuscript in press. Please, cite as
Beracci, A., Santiago, J., & Fabbri, M. (in press). The categorical use of a
continuous time representation. *Psychological Research*.

The Categorical Use of a Continuous Time Representation

Authors: Alessia Beracci*¹, Julio Santiago², & Marco Fabbri¹

¹ Department of Psychology, University of Campania Luigi Vanvitelli, Caserta, Italy

² Department of Experimental Psychology, University of Granada, Spain

***corresponding author:**

Alessia Beracci, MS

Department of Psychology

University of Campania Luigi Vanvitelli

Viale Ellittico 31, 81100 Caserta (CE),

Italy

Tel: +39-0823-275333

Fax: +39-0823-274759

e-mail: alessia.beracci@unicampania.it

ABSTRACT

The abstract concept of time is mentally represented as a spatially oriented line, with the past associated with the left space and the future associated with the right. Although the line is supposed to be continuous, most available evidence is also consistent with a categorical representation that only discriminates between past and future. The aim of the present study was to test the continuous or categorical nature of the mental timeline. Italian participants judged the temporal reference of 20 temporal expressions by pressing keys on either the left or the right. In Experiment 1 ($N=32$), all words were presented at the center of the screen. In Experiment 2 ($N=32$), each word was presented on the screen in a central, left, or right position. In Experiment 3 ($N=32$), all text was mirror-reversed. In all experiments, participants were asked to place the 20 temporal expressions on a 10-cm line. The results showed a clear *Spatial-TEmporal Association of Response Codes (STEARC)* effect which did not vary in strength depending on the location of the temporal expressions on the line. However, there was also a clear Distance effect: latencies were slower for words that were closer to the present than further away. We conclude that the mental timeline is a continuous representation that can be used in a categorical way when an explicit past vs. future discrimination is required by the task.

Keywords: Space; Time; Mental Time Line; STEARC effect; Categorical Time.

1. Introduction

There is a growing consensus that abstract concepts such as numbers and time have a spatial mental representation. According to the literature, both numbers and time are represented along a laterally oriented line (Bonato, Zorzi & Umiltà, 2012). Thus, these lines are often referred to as the mental number line (MNL) and mental time line (MTL). One of the most important challenges in embodied and grounded cognition theories is to understand how abstract concepts (such as time, numbers, emotions, etc.) are acquired, represented, and used, and it is recognized that language plays an important role (Borghi, 2020). Language can work as both an inner and a social tool that influences the mental representation of abstract concepts through spatial mappings (Borghi & Setti, 2017). In addition, an investigation into the nature of the time line could shed light on how this representation indicates how we interact with the outside world, how (and why) we associate temporal elements with space and, finally, how we store memories in long-term memory, by rebuilding the mapping of life events.

The main supporting evidence regarding the MNL is the *Spatial-Numerical Association of Response Codes*, or SNARC effect (Dehaene et al., 1993; Hubbard, Piazza, Pinel & Dehaene, 2005; Nuerk, Bauer, Krummenacher, Heller & Willmes, 2005a; Nuerk, Wood & Willmes, 2005b; Schwarz & Keus, 2004), in which smaller numbers are responded to faster with left-sided responses, and relatively larger numbers are responded to faster with right-sided responses.

Likewise, in the temporal domain, an analogous effect to the SNARC effect has been found, for which Ishihara, Keller, Rossetti and Prinz (2008) coined the term *Spatial-TEmporal Association of Response Codes*, or STEARC effect. This effect has been found with three different strategies that are used to convey temporal information. The first strategy uses words that refer to either the past or the future (Santiago, Lupiáñez, Pérez & Funes, 2007; Torralbo, Santiago & Lupiáñez, 2006). The second strategy uses stimuli that vary in actual duration (Fabbri, Cancellieri & Natale, 2012; Fabbri, Cellini, Martoni, Tonetti, & Natale, 2013a,b; Ishihara et al., 2008; Vallesi, Binns, and Shallice, 2008). A third strategy uses events (presented by means of pictures) that can be construed as having a particular temporal location within a wider sequence (Fuhrman and Boroditsky, 2010; Santiago, Román, Ouellet, Rodríguez and Pérez-Azor, 2010). All three strategies have provided converging results that support a left-to-right spatial mapping along the MTL : shorter durations, expressions referring to the past, and prior events were responded to faster with left-sided responses, while longer durations, expressions referring to the future, and subsequent events were responded to faster with right-sided responses. Although Ishihara et al. (2008) coined the term

"STEARC effect" specifically in the context of the second strategy, we contend that the effect revealed by the three strategies is actually one and the same, and we will use the term generally hereafter.

The similarities between the SNARC and the STEARC effects go further than the intuitively obvious analogy between ordered numerical and temporal series that run from left to right (see Bonato, Zorzi, & Umiltà, 2012, for a detailed comparison). An important argument supporting the underlying representations having a great deal in common is the relationship of these effects with reading and writing direction. In their seminal report regarding the SNARC effect, Dehaene et al. (1993) reported that the effect is reduced in participants who read from right to left. Later, Zebian (2005) reported a reversed SNARC in Arabic monoliterates and Arabic-English biliterates; and Shaki, Fischer and Petrusic (2009) found the same result in Palestinians. Evidence for a causal link was provided by Fischer, Mills, and Shaki (2010), who changed the SNARC by manipulating the location of smaller and larger numbers on the page.

Similarly, the STEARC effect has also been shown to depend on reading and writing direction. Ouellet, Santiago, Israeli, and Gabay (2010) found a reversed STEARC effect in Hebrew readers using words with a temporal reference, and Fuhrman and Boroditsky (2010) reported the same result using pictures of sequential events. Vallesi, Weisblatt, Semenza, and Shaki (2014) found that the STEARC effect disappeared (although it was not reverted) in Hebrew readers. Casasanto and Bottini (2014) were able to revert the STEARC effect by asking their participants to judge the temporal reference of mirror-reversed expressions in Dutch, thus proving a causal link between reading direction and the directionality of the STEARC. Thus, both the MNL and the MTL seem to arise from directional experiences that correlate with the direction of the habitual script, suggesting a cultural origin. .

A non-cultural account of congruency effects that might be applicable to the SNARC and the STEARC effects was proposed by Proctor and Cho (2006), and was based on the idea of polarity correspondence. For a variety of binary classification tasks, participants code the stimulus alternatives and response alternatives as (+) polarity and (-) polarity, and response selection is faster when the polarities correspond than when they do not. Indeed, the SNARC and the STEARC effect could be due to correspondence in magnitude polarity [large number, long durations, “after”, future (+); small number, short durations, “before”, past (-)] and response polarity [right (+); left (-)]. However, the polarity correspondence account of the SNARC effect was explicitly put to the test by Santiago and Lakens (2015), who found it unsupported by the data.

Additional evidence supporting the similarity between the SNARC and the STEARC effects was provided by Fabbri, Cancellieri, and Natale (2012), who used a procedure in which time, numbers, and space were all task-relevant, and observed a triple interaction between all the elements. Additionally, Anelli et al. (2018), when examining the directionality of the MTL with a bimodal response setting, showed that the STEARC effect emerged at the stage of response selection (Anelli et al., 2018; for additional findings see Vallesi et al., 2008; Vallesi, McIntosh & Stuss, 2011), in agreement with previous findings concerning the SNARC effect (Gevers, Lammertyn, Notebaert, Verguts & Fias, 2006; Keus & Schwarz, 2005). All the previous results showing how the SNARC and STEARC effects share similar features suggest that the mental representations of numbers and time have a lot in common, being constituted in both cases by a mental line that runs along the lateral axis.

It is intrinsic that a mental line has an analogical and continuous nature. However, this assumption has not been systematically verified. For example, Gevers, Lammertyn, Notebaert, Verguts and Fias (2006) demonstrated that the SNARC effect (and consequently the spatial characteristics of the MNL) was categorical or continuous depending on task instructions. The authors observed that the size of the SNARC effect (the difference between the latency of right and left hand responses) varied continuously with number magnitude in a parity judgment task, but was oblivious to number size in a magnitude comparison task, suggesting an underlying categorical representation in the latter (see Figures 2A vs. 2B in Gevers et al., 2006, p. 34; Figure 3A vs. 3B, in Gevers et al., 2006, p. 35). The authors provided a general framework with which to interpret the SNARC effect, postulating the existence of two routes of activation during the processing of numbers: a relatively fast route that automatically codes for the location of the numeric stimulus on the line and a relatively slow route that is dependent on the task instruction and provides the mapping of the relevant attribute to the required response. When both routes converge on the same spatial response code (a congruent condition), the response can be relatively fast and accurate. When both routes converge on opposing response codes (an incongruent condition), reaction times are slower, and errors are more frequent. The evidence from the number domain is thus maximally consistent with an underlying continuous representation that can be used categorically when the task imposes explicit magnitude comparisons with a criterion.

Can we extrapolate this account to the temporal domain? In order to do so, the STEARC effect should be categorical in explicit tasks and continuous in implicit tasks. The available findings are inconclusive. Is the STEARC effect continuous in implicit time tasks? The main issue in

answering this question is that it is unclear whether, and how, it is possible to find a STEARC effect in implicit tasks at all. Ulrich and collaborators (Ulrich and Maienborn, 2010; Ulrich, Eikmeier, de la Vega, Ruiz Fernández & Alex-Ruf, 2012; Maienborn, Alex-Ruf, Eikmeier & Ulrich, 2015) asked participants to decide whether past and future-related sentences were sensible or nonsensical by moving a lever. In this implicit task they failed to observe the STEARC effect at all, but the effect was clearly present when the task was to judge the temporal reference of the sentence. The authors reasoned that, when time is task-irrelevant, either the MTL is not automatically activated by the linguistic information or its activation is too weak to be detected.

However, there are studies which have succeeded in observing a STEARC effect in an implicit task. Lakens, Semin, and Garrido (2011) asked participants to place eight past and eight future-related expressions on a horizontal line. The expressions varied in their temporal distance from the present (from "past" through "the day before yesterday", "a moment ago", "immediately", "the day after tomorrow", to "future"). The results showed that the words were positioned at different locations on the line following a left-to-right progression that agreed with their temporal reference, suggesting that their meanings are represented along a spatial continuum. In a subsequent experiment, the expressions were presented auditorily and were intermixed with words that were unrelated to time (e.g., "table" or "glass"). Participants were asked to judge which ear was presented with the louder stimulus. Crucially, time-related expressions were all presented with the same volume. The results showed that future-related words were judged to be louder in the right ear more often than past-related words. Furthermore, the visual-spatial ordering of the expressions obtained in Experiment 1 was linearly related to the probability of judging an expression as louder in the right ear. The authors also tested the relationship using a categorical predictor (past vs. future), but a significant fit failed to emerge, suggesting that the pattern was linear. However, this study has one important caveat: the temporal range of the expressions was small. Besides the words "past" and "future", which do not refer to any specific temporal distance, the remaining expressions ranged from the day before yesterday to the day after tomorrow. Moreover, the agreement between participants in locating some of the expressions may have been low (e.g., "been", "coming"). These factors may underlie the fact that the observed linear relationship, though significant, was small ($r^2=.30$), and may elicit doubts about its replicability.

Because of the strong relationship between sequential order and temporal order, studies using sequences of ordered elements are also relevant to the present discussion. However, they also suffer from problems that make them inconclusive. Gevers, Reynvoet, and Fias (2003, 2004)

presented months, letters of the alphabet, and days of the week to be compared with a central reference (explicit task), as well as in implicit, order-irrelevant tasks (e.g., say whether the month name contains the letter "R"). They observed STEARC effects in both the explicit and implicit tasks. Inspection of the figures suggests a categorical pattern in the explicit order judgement tasks and a continuous pattern in the implicit tasks. However, they did not contrast a categorical versus a continuous predictor, and therefore, it is uncertain whether the observed pattern should be considered continuous or categorical.

Is the STEARC effect categorical in explicit time judgement tasks? Available evidence does not provide a clear answer to this question either. Vallesi et al. (2008) asked their participants to judge the duration of a visual stimulus varying from 0.5 to 3.5 sec as smaller or larger than a criterium (2 sec), and compared the degree of fit obtained by a continuous and a categorical predictor of the size of the STEARC effect. The categorical predictor provided a better fit than the continuous predictor. However, Ding, Feng, Cheng, Liu, and Fan (2015) found some evidence of a continuous pattern while exploring the symmetry of the STEARC effect toward the past and the future. They compared the size of the effect when expressions that referred to a close (a day) versus a far (a year) distance in the past or the future were judged. They showed that the STEARC effect was greater for temporally distant than for closer events, suggesting a continuous underlying representation. However, this study also has many caveats: first, only past, but not future distant events showed a greater STEARC effect; and second, there was no explicit test for a continuous vs. categorical STEARC effect. Also inconclusive is the study by Santiago et al. (2010) which used an explicit order judgement task. The authors asked participants to watch movie clips or picture sequences depicting everyday events (e.g., getting up and making breakfast). Participants then had to classify clip frames or pictures as belonging to the first or second half of the story by means of left or right key presses (there was no implicit task). In this case, visual inspection of the figures suggests a categorical pattern, but the authors did not expressly test for this by comparing a categorical vs. a continuous predictor.

It is important to note that in all the studies that tested for a Distance effect, it was in the explicit tasks that the effect was found: responses were slower for items that were closer to the reference (Gevers et al., 2003, 2004; Santiago et al., 2010). Since Moyer and Landauer (1967) originally reported that comparing closer numbers is more difficult than comparing distant numbers, the Distance effect has been considered the hallmark of a continuous representation of numbers (see Dehaene, 1992, for a review; Dehaene et al., 1993; Fias, 2001; Fias, Brysbaert, Geypens, &

d'Ydewalle, 1996). Thus, this effect would suggest a continuous underlying representation of time in explicit tasks. However, the data are again unclear: while Santiago et al. (2010) found a symmetrical Distance effect, Gevers et al. (2004) found it only on one side (Gevers et al., 2004), and it was also affected by unexplained interactions (Gevers et al., 2003).

In summary, in terms of the existence of a continuous or categorical representation underlying the STEARC effect in explicit and implicit temporal tasks, the evidence is scarce, mixed, and unclear. Because of the difficulty of finding the STEARC effect in implicit tasks, we focused on an explicit task. The main aim of our study was to provide a conclusive test of 1) whether the STEARC effect is continuous or categorical, and 2) whether there is a Distance effect when an explicit temporal judgement is required. Thus, in three experiments we requested participants to perform a STEARC explicit temporal task in which they had to categorize, as past or future, 20 Italian temporal expressions that referred to time points varying from the distant past to the distant future. In order to make sure that the meanings of the temporal terms were distributed more or less continuously over space and to estimate their spatial location with precision, after the STEARC task we asked the same participants to place them on a line. This also allowed us to use the position of each term on the line as reported by each individual to predict both the Distance effect (the effect of spatial position on average reaction time) and the STEARC effect (the effect of spatial position on the difference in reaction time between the right and left responses) of that individual. This allowed us to reach general conclusions regarding the relationship between space and time at the same time as we controlled for idiosyncratic patterns. More importantly, we followed the logic used by Gevers et al. (2006), Vallesi et al. (2008), and Lakens et al. (2011) to directly test for a continuous versus a categorical pattern in the STEARC effect by contrasting the goodness of fit of a continuous predictor with a categorical predictor. The three experiments were close replications that varied only in aspects of stimulus presentation: central presentation in Experiment 1, central and lateral presentation in Experiment 2, and mirror-reversed central and lateral presentation in Experiment 3. With reference to the study by Santiago, Román and Ouellet (2011), we reasoned that the degree of saliency of the lateral axis might affect how this axis is used to perform the task. In all experiments, the lateral axis is task-relevant because it is used to guide manual responses. In Experiment 2, we expected its saliency to increase due to the presence of a left-right contrast of stimulus presentation on the screen. In Experiment 3 we used mirror-reversed stimuli following the work by Casasanto and Bottini (2014). We expected this manipulation to reverse the STEARC effect, but our main goal was to push the saliency of the left-right axis to the extreme, in the hope of capturing a continuous STEARC effect. Thus, the decision to perform three

different experiments with different degrees of saliency of the spatial left-right axis was made in order to provide a strong test of continuous vs. categorical nature of the MTL in explicit tasks.

2. Methods

2.1 Participants

The studies reported here were approved by the Ethical Committee of the University of Caserta. All participants were Italian university students, who took part in exchange for course credit, had normal or corrected-to-normal vision, provided written informed consent, and filled in a demographic questionnaire as well as the Edinburgh Handedness Inventory (EHI; Oldfield, 1971). Experiment 1 included 32 students (29 females; mean age = 24.44 years; SD = 2.00 years). According to the EHI scores, 28 were right-handed (M = 92.44; SD = 13.42) and 4 were left-handed (M = -75.00; SD = 30.62). Experiment 2 included a separate group of 32 students (30 females; mean age = 23.91 years; SD = 1.65 years; 30 were right-handers, M = 89.73; SD = 17.72; and 2 left-handers, M = -94.74; SD = 7.44). In Experiment 3 there was a new sample of 32 students (27 females and 5 males; mean age = 24.13 years; SD = 1.48 years; 26 were right-handed, M = 85.37; SD = 14.82; and 6 were left-handed, M = -82.48; SD = 15.95).

2.2 Materials

All experiments reported hereafter used the same set of materials. The stimuli consisted of 20 Italian temporal expressions, 10 referring to the past ("anticamente", formerly; "tempo fa", long ago; "in passato", in the past; "una volta", once; "precedentemente", previously; "l'altro ieri", the day before yesterday, "l'altro giorno", the other day; "prima", before; "ieri", yesterday; "recentemente", recently) and 10 referring to the future ("dopo", after; "tra poco", in a short while; "presto", soon "conseguentemente", subsequently; "domani", tomorrow; "in seguito", later; "dopodomani", the day after tomorrow; "successivamente", thereafter; "prossimamente", next; "in futuro", in the future). Note that the temporal reference of the English translations may not match exactly the reference in Italian.

2.3 Procedure

All stimuli were presented in white font on a black laptop computer screen. The display had a resolution of 1,072 × 960 pixels and a refresh rate of 72 Hz. Participants sat at a viewing distance

of 55 cm and were individually tested in a quiet room. Stimulus presentation and data collection were controlled using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002).

The experimental session had two parts, for a total duration of approximately 20 minutes. In the first part, participants were required to categorize the target temporal expressions as referring to the past or to the future by means of a yellow key on the left and a blue key on the right of a response box (Cedrus RB-x40 Response Pads) that was connected to the laptop. In each trial, a fixation cross (+) symbol was presented first at the center of the screen for 1000 ms, followed by a temporal expression for 5000 ms or until a response was recorded. All target stimuli were presented in two separate blocks with different time–key mappings. In one block, the left key was pressed in order to categorize the target as past, and the right key to categorize the target as future (see Figure 1). The other block had the reverse mapping. The order of blocks was counterbalanced over participants. Each block comprised a total of 180 trials (nine repetitions of each target expression) in a random order. Thus, altogether the participants judged 360 trials. Before the test, a 10-trial training block was run. The training phase could be repeated if requested by the participant. After each block, the participant was allowed to take a short break.

After the time judgement task, participants were requested to perform a paper-and-pencil Time-to-Position task (see Figure 2). They were presented with each of the 20 temporal expressions written above a 10-cm long line that was flanked by the labels “*passato lontano*” (“distant past”) on the left and “*futuro lontano*” (“far future”) on the right. Only one expression was presented in each trial, in a randomized order, and participants could not see their responses to prior expressions. They then had to decide where the temporal meaning of the expression should be positioned on the line by making a vertical mark with a pencil (see Fabbri & Guarini, 2016, and Fabbri & Natale, 2016, for a similar procedure with numbers).

The only difference between experiments regarded the presentation of the target stimuli (see Figure 3). In Experiment 1, all target expressions were presented centrally on the screen. In Experiment 2, expressions were presented either at the center, on the left or on the right of the screen. In Experiment 3, expressions were mirror-reversed and presented at the same three locations as in Experiment 2.

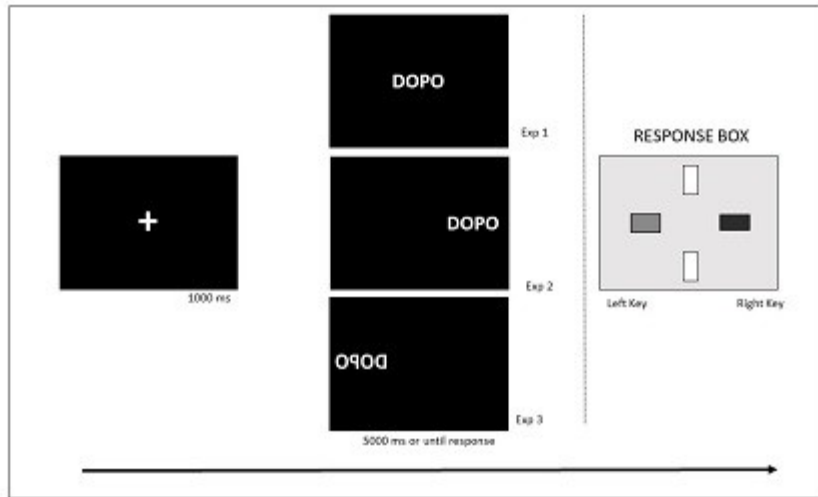


Figure 1. Example of trial sequences of the temporal judgement task.

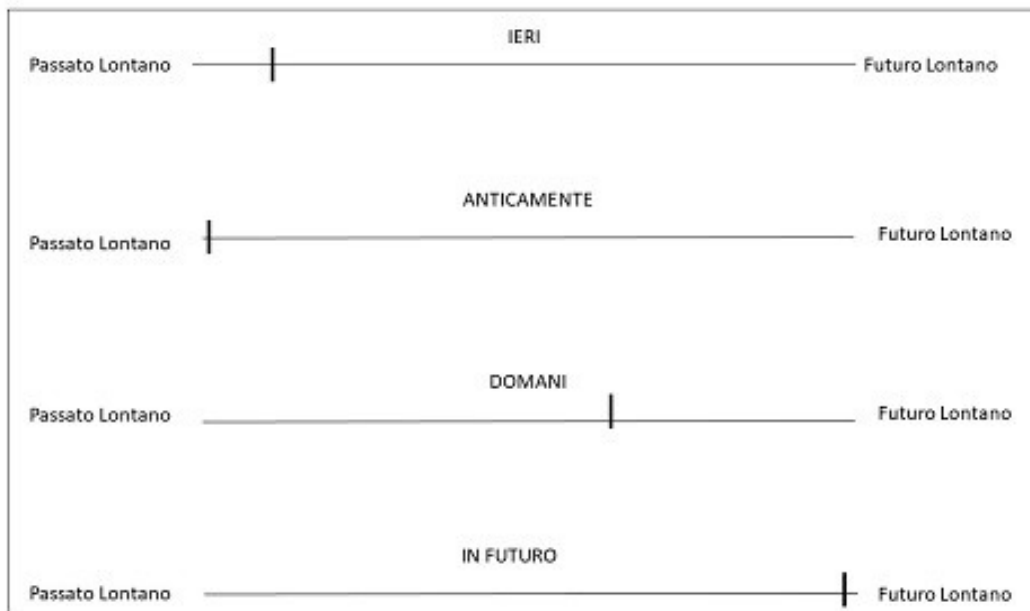


Figure 2. Example of several trials in the Time-To-Position Task. Only one stimulus and line was presented in each trial.

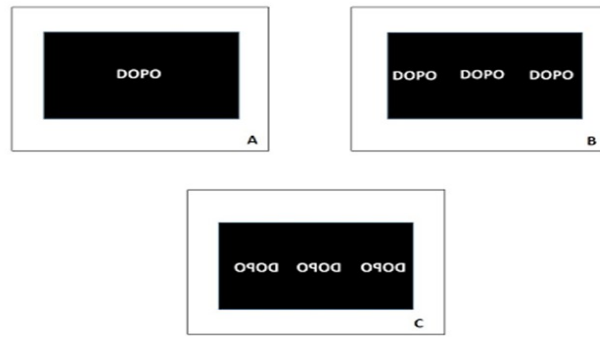


Figure 3. a) Example of stimulus presentation in Experiment 1; b) Locations of stimuli presentation in Experiment 2; c) Locations of stimuli presentation in Experiment 3. Note that in experiments 2 and 3 only one expression was presented on the screen in each trial.

3. Data Analysis

In order to verify the presence of a STEARC and a Distance effect and to assess whether the former takes a continuous or categorical form, data were analyzed as follows.

First, we analysed the Time-to-Position Task, in order to make sure that the reference of the temporal expressions spreads more or less homogeneously over the space from the distant past to the distant future. As in previous studies (Ebersbach, Luwel, Frich, Onghena & Verschaffel, 2008; Fabbri & Guarini, 2016; Fabbri & Natale, 2016; Siegler & Opfer, 2003; White & Bull, 2008), we measured (in millimeters), for each participant and each temporal expression, the distance from the left side of the line to the subjective mark positioned on the line. In order to put these measures on a common metric we computed z scores for each participant. This standardized score indicated how far the subjective position of each expression on the line was from the subjective center (average position of all expressions) in standard deviation units.

Second, we focused on reaction times (RTs) in correct trials (94.98% of total trials) and analysed them by means of mixed models (LMM) using the package lme4 (Bates, Mächler, Bolker, & Walker, 2015) in R version 3.6.1 (R Core Team, 2018). The models used sum contrasts. Categorical predictors were dummy coded and centered. As in the studies by Barr, Levy, Scheepers, and Tily (2013) and Barr (2013), we started by specifying the maximally complex model and then simplified it until the convergence problems were resolved, and then until the simplest model was found that did not lose goodness of fit (Bates, Kliegl, Vasishth, & Baayen, 2015). To compare models we used likelihood tests as implemented in the anova() function of the package LmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). The final choices of parametric (linear) models

were interpreted by means of the ANOVA tables and p-values provided by the `anova()` function of `LmerTest` using the Satterthwaite method. Generalized models were interpreted using the `Anova()` function of the package `car` (Fox & Weisberg (2019)). Whenever we found convergence problems we re-ran the analysis using the `bobyqa` optimizer, which sometimes succeeded (using a different optimizer does not change the results of a model, but may sometimes alleviate convergence problems).

RTs smaller than 200 ms or larger than 2.5 SDs above the participant's mean were discarded as outliers (3.25% of correct trials). In order to test whether the STEARC effect takes a continuous or categorical form, we first averaged the RT for each hand, item, and participant. Then, we subtracted the RT of the left hand from the RT of the right hand. The resulting distribution of differential RTs showed no asymmetry, lending validity to the use of parametric approaches of analysis (i.e., assuming a Gaussian distribution). Differential RTs were predicted from the fixed factors Experiment (1 to 3), Time (past vs. future), and Distance in the Time-to-Position Task. We followed two parallel analytical approaches: one started with the continuous predictor, found the simplest model, and then added the categorical predictor; another started with the categorical predictor, found the simplest model, and then added the continuous predictor. In the final models we could thus compare which (the continuous or the categorical) predictor provided the best fit.

In order to assess the Distance effect, for each participant and item we computed the average RT of both left and right responses. We also computed the absolute values of the standardized distances from the subjective center, so that they served as a non-directional measure of distance from the center. The effect of Absolute Distance was then assessed together with the effects of categorical Time (past vs. future), Experiment (1 to 3), and their interactions. An interaction between Absolute Distance and Time would reveal asymmetrical Distance effects toward the past and the future. Mean RTs showed a skewed distribution, so we used the strategy proposed by Lo and Andrews (2015): using generalized mixed models, we compared a model that assumed a Gamma distribution with another assuming an Inverse Gaussian distribution, and retained the best model as the starting point of the analysis.

All data and analysis scripts are available for download via the following Open Science Framework repository: <https://osf.io/7jckp/quickfiles>

4. Results

4.1 Time-to-Position Task

As shown in Figure 4, the standardized distances to the subjective center in the Time-to-Position task revealed that the 20 temporal terms spanned a wide spatial interval, always keeping to the expected order. Thus, Italian speakers placed the 20 temporal expressions in an ordered pattern from left to right along a linear (time) line. Although the selected temporal expressions did not fall at regularly spaced intervals, their number and dispersion support their use as a continuous predictor of reaction times in the experimental tasks.

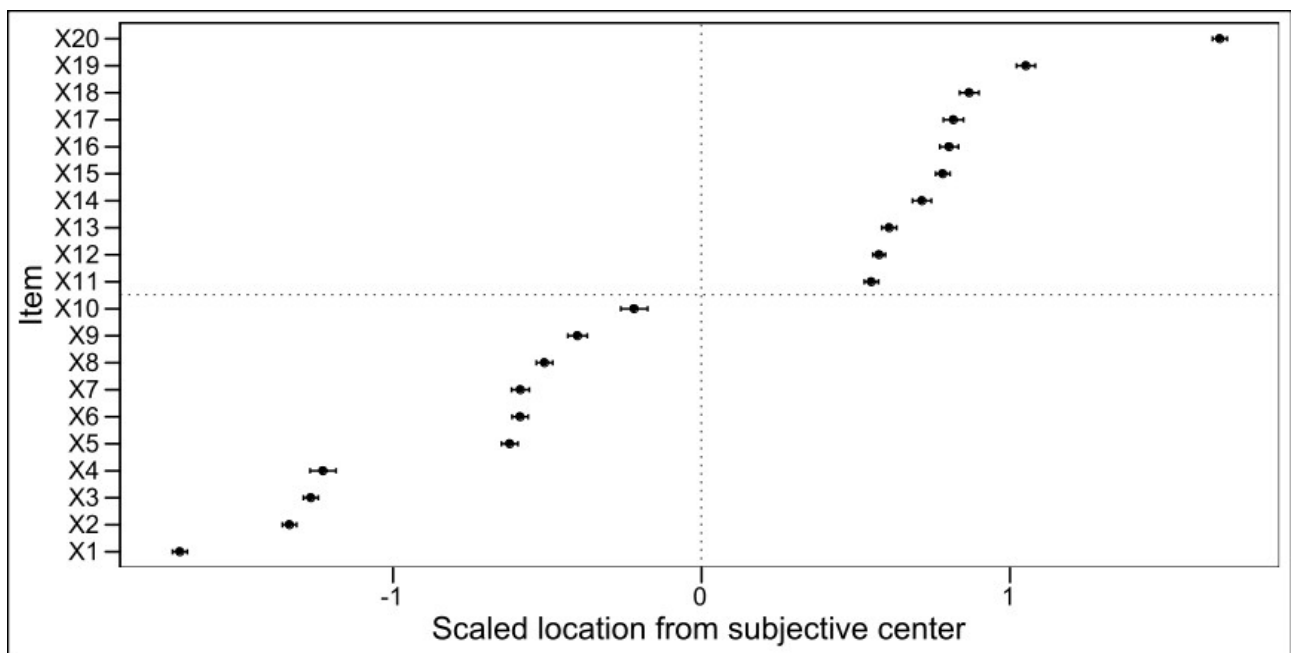


Figure 4. Distance from the subjective center in SD units for each temporal expression in the Time-to-Position Task. Error bars show SEM. Items: X1 = Anticamente, X2 = Tempo fa, X3 = In passato, X4 = Una volta, X5= Precedentemente, X6 = L'altro ieri, X7 = L'altro giorno, X8 = Prima, X9 = Ieri, X10 = Recentemente, X11 = Dopo, X12 = Tra poco, X13 = Presto, X14 = Conseguentemente, X15 = Domani, X16 = In seguito, X17 = Dopodomani, X18 = Successivamente, X19 = Prossimamente, X20 = In futuro.

4.2 Continuous vs. categorical STEARC Effect

In the first analytical approach we started using the continuous predictor (see Table 1 for a description of the sequence of analytical steps and their results). The random term in the maximal model included random intercepts per participant as well as random slopes of standardized Distance

over participants. The fixed effects were Experiment (1 to 3), Distance and their interaction. The dependent variable was the right hand RT minus left hand RT for each item and participant. The maximal model failed to converge. Inspection of the results suggested that the convergence problems arose because of variability in random intercepts that was too small (as the dependent variable was the difference in latencies between the two hands, it showed a size close to zero and a very small variability). Therefore, we removed participant intercepts from the random term, keeping the random slopes of Distance. This model did converge and revealed a main effect of Distance but no effect of Experiment or interaction with Distance. We then searched for the simplest model without loss of fit. A model keeping only the Distance factor did not have a significantly different goodness of fit ($\chi^2(4)=0.43$, $p=.98$), so we removed Experiment and the Experiment x Distance interaction. In this simplest model, Distance had a very clear effect ($F(1,94.94)=29.46$, $p<.001$). Therefore, in agreement with our expectations, there was a clear STEARC effect. In contrast to expectations, the STEARC effect did not change over experiments, even when all the materials were mirror-reversed in Experiment 3. Of central relevance to the question of interest, in the simplest model we then introduced the categorical predictor Time (past vs. future). This change produced an important increase in overall goodness of fit ($\chi^2(1)=77.57$, $p<.001$). Moreover, the categorical predictor Time became the only factor responsible for the fit, as it showed a significant main effect ($F(1,1819.6)=79.15$, $p<.001$) while the continuous predictor Distance became non-significant ($F(1,194.2)=0.15$, $p=.70$). To provide a final check, we removed the continuous predictor from the model. The goodness of fit of the resulting model, including only the categorical predictor Time, did not differ from the model that included both predictors ($\chi^2(1)=0.15$, $p=.70$).

In the second analytical approach we started using the categorical predictor Time (see Table 2). The random term of the maximal model included random intercepts for participants and random slopes over Time. The fixed effects included Time, Experiment, and their interaction. The maximal model converged and revealed only a main effect of Time. We then removed Experiment and the interaction Time x Experiment from the model and compared the resulting simplest model with the maximal model: the goodness of fit did not change ($\chi^2(4)=0.53$, $p=.97$). The effect of the categorical predictor Time was very clear ($F(1,94.96)=30.29$, $p<.001$). We then added the continuous Distance predictor to the simplest model, but it did not improve fit ($\chi^2(1)=0.83$, $p=.36$).

In order to provide a final test of the categorical nature of the STEARC effect, we ran independent analyses of the effect of the continuous predictor Distance on each Time condition (past vs. future; see Table 3). In the Past condition, the maximal model included the effects of

Distance, Experiment, and their interaction, random intercepts for participants and random slopes of Distance over participants. The model converged and revealed no significant findings. We simplified the model, leaving only the fixed effect of Distance. The resulting simplest model did not differ from the maximal model in goodness of fit ($\chi^2(4)=0.15, p=.997$). In this model, the effect of Distance was also non-significant ($F(1,79.94)=0.46, p=.50$). In the Future condition, the maximal model converged and also showed no significant findings. The simplest model, including only the fixed effect of Distance, produced a singular fit, so we also simplified the random term by removing the participant intercepts, which solved the problem. To allow a fair comparison of the goodness of fit of the maximal and the simplest models we also removed the participant intercepts from the maximal model (which did not change the null findings). The simplest model showed the same goodness of fit as the latter maximal model ($\chi^2(4)=0.71, p=.95$). In the simplest model, Distance failed to make any significant contribution ($F(1,292.63)=0.21, p=.65$). Therefore, when analyzed only within the range of past or future expressions, the continuous predictor Distance failed to have any effect on differential RTs (see Figure 5).

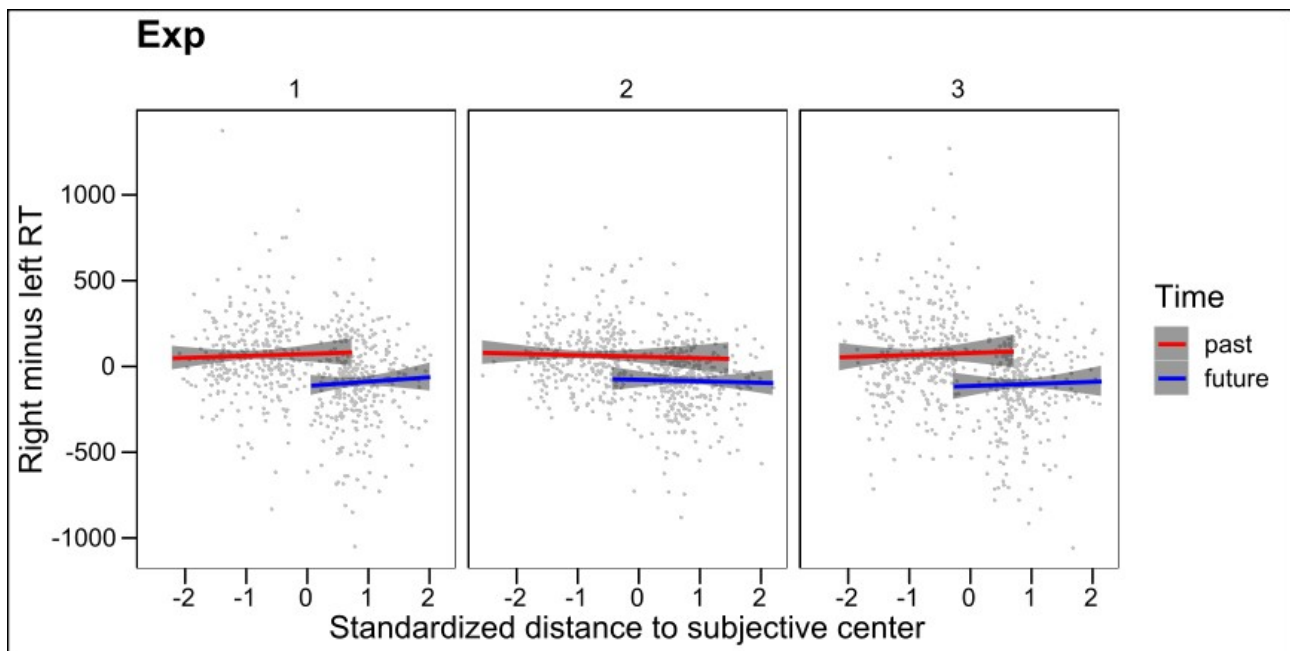


Figure 5. Differential RTs (average RT of the right hand for each participant and expression minus the average of those produced with the left hand) for the past and future expressions. Lines represent the best fitting linear model within each temporal category. The gray areas represent the 95% confidence interval.

4.3 Distance Effect

The analysis of the Distance effect used the factor Absolute Distance (see above) in order to produce a non-directional estimation of distance from the subjective center (see Table 4). The fixed effects in the maximal model included the factors Absolute Distance, the categorical factor Time (past vs. future), Experiment, and all their interactions. The random term included participant intercepts, and the random slopes over participants of Absolute Distance, Time, and their interaction. As in the study by Lo and Andrews (2015) we tried both a model that assumed an Inverse Gaussian distribution and one that showed a Gamma distribution. The former model failed to converge, so we continued the analysis using the latter. The maximal model revealed main effects of Experiment, Absolute Distance, and Time, as well as an Absolute Distance x Time interaction. We then simplified the model by removing all the non-significant interactions. The resulting simplest model did not change goodness of fit ($\chi^2(4)=0.71$, $p=.95$). Any further simplifications reduced the predictive ability of the model. The simplest model showed clear main effects of Experiment ($\chi^2(2)=32.15$, $p<.001$), Absolute Distance ($\chi(1)=85.34$, $p<.001$), Time ($\chi(1)=19.50$, $p<.001$), and the interaction between Absolute Distance and Time ($\chi(1)=14.17$, $p<.001$). As shown in Figure 6, RTs were slower in Experiment 3, and increased when the expression was closer to the subjective middle; RTs for future expressions were faster than for past expressions, and the slope of the Distance effect was steeper for past than for future expressions.

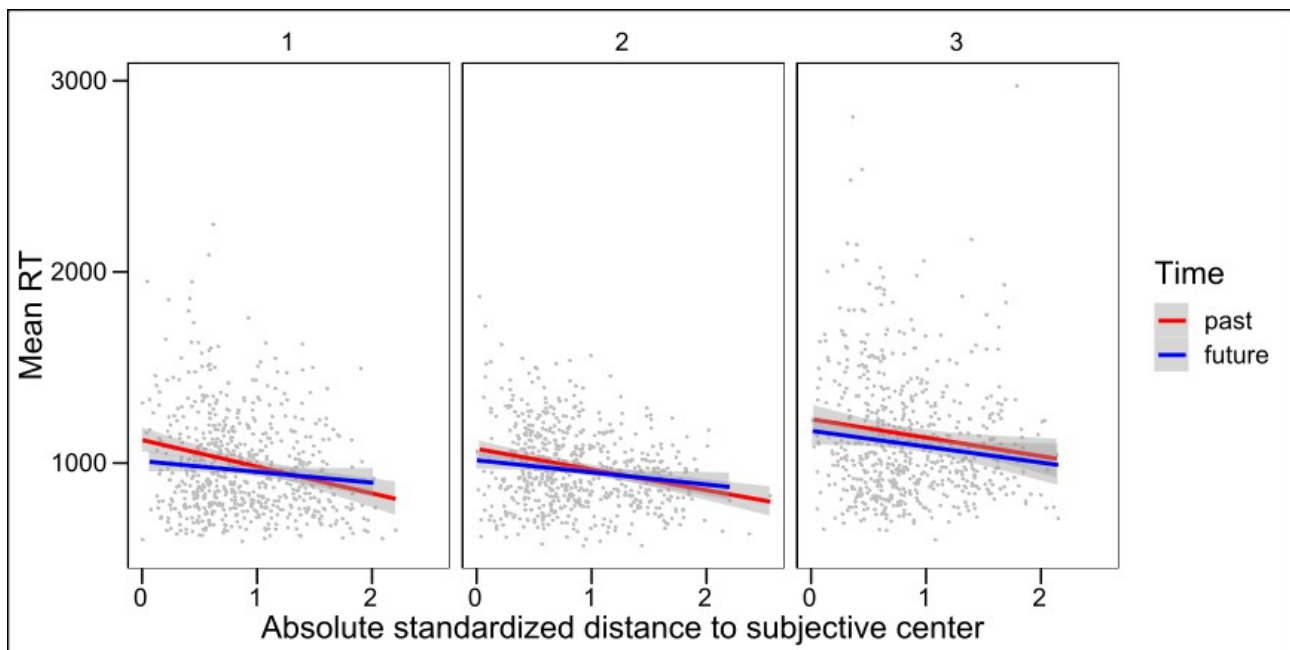


Figure 6. Average RTs for each participant and expression for the past and future expressions. Lines represent the best fitting linear model within each temporal category. The gray areas represent the 95% confidence interval.

Table 1. Comparisons of Fit Models. Key: AIC: Aikake Information Criterion; **BIC:** Bayesian Information Criterion; **LL:** LogLikelihood; **Df:** degrees of freedom
 Note that: **RminusL:** right hand (RTs) minus left hand (RTs); **Exp:** Experiment (1 to 3); **stmm:** Distance; **Time:** Time (past vs. future); **0+stmm|SubjectID:** no random intercept and random slopes of Distance over participants.

Model Equation	Fixed Effects	Model fit				
		Df	AIC	BIC	LL	Deviance
RminusL ~ Exp + stmm + Exp:stmm + (0+stmm SubjectID)	Experiment + Distance + Experiment x Distance	8	26024	26068	-13004	26608
RminusL ~ stmm + (0+stmm SubjectID)	Distance	4	26016	26039	-13004	26608
RminusL ~ stmm + Time + stmm:Time + (0+stmm SubjectID)	Distance + Time + Distance x Time	5	25941	25969	-12965	25931
RminusL ~ Time + (0+stmm SubjectID)	Time	4	25939	25961	-12966	25391

Table 2. Comparisons of Fit Models. Key: AIC: Aikake Information Criterion; **BIC:** Bayesian Information Criterion; **LL:** LogLikelihood; **Df:** degrees of freedom
 Note that: **RminusL:** right hand (RTs) minus left hand (RTs); **Exp:** Experiment (1 to 3); **stmm:** Distance; **Time:** Time (past vs. future); **1+Time|SubjectID:** random intercepts and random slopes of Time over participants.

Model Equation	Fixed Effects	Model fit				
		Df	AIC	BIC	LL	Deviance
RminusL ~ Exp + Time + Exp:Time + (1+Time SubjectID)	Experiment + Time + Experiment x Time	10	25723	25779	-12852	25703
RminusL ~ Time + (1+Time SubjectID)	Time	6	25716	25749	-12852	25704
RminusL ~ Time + stmm + (1+Time SubjectID)	Time + Distance	7	25717	25756	-12852	25703

Table 3. Comparisons of Fit Models. Key: **AIC:** Aikake Information Criterion; **BIC:** Bayesian Information Criterion; **LL:** LogLikelihood; **Df:** degrees of freedom.

Note that: **RminusL:** right hand (RTs) minus left hand (RTs); **Exp:** Experiment (1 to 3); **stmm:** Distance; **Time:** Time (past vs. future); **0+stmm|SubjectID:** no random intercept and random slopes of Distance over participants; **1+stmm|SubjectID:** random intercepts and random slopes of Distance over participants.

Model Equation	Fixed Effects	Model fit				
PAST						
		Df	AIC	BIC	LL	Deviance
RminusL ~ Exp + stmm + Exp:stmm + (1+stmm SubjectID)	Experiment + Distance + Experiment x Distance	10	12958	13007	-6469.0	12938
RminusL ~ stmm + (1+stmm SubjectID)	Distance	6	12950	12979	-6469.1	12938
FUTURE						
		Df	AIC	BIC	LL	Deviance
RminusL ~ Exp + stmm + Exp:stmm + (1+stmm SubjectID)	Experiment + Distance + Experiment x Distance	10	12874	12923	-6427.3	12844
RminusL ~ stmm + (1+stmm SubjectID)	Distance	6	12867	12896	-6427.7	12855
RminusL ~ stmm + (0+stmm SubjectID)	Distance	4	12932	12951	-6461.8	12924
RminusL ~ Exp + stmm + Exp:stmm + (0+stmm SubjectID)	Experiment + Distance + Experiment x Distance	8	12939	12923	-6461.4	12923

Table 4. Comparisons of Fit Models. Key: **AIC:** Aikake Information Criterion; **BIC:** Bayesian Information Criterion; **LL:** LogLikelihood; **Df:** degrees of freedom.

Note that: **RTmean:** reaction times average; **Exp:** Experiment (1 to 3); **stmmAbs:** Absolute Distance; **Time:** Time (past vs. future); **1+stmmAbs*Time|SubjectID:** random intercept and random slopes of Absolute Distance x Time over participants.

Model Equation	Fixed Effects	Model fit				
		Df	AIC	BIC	LL	Deviance
RTmean ~ Exp * stmmAbs *Time + (1+stmmAbs*Time SubjectID)	Experiment + Absolute Distance + Time + Experiment x Absolute Distance + Experiment x Time + Absolute Distance x Time + Experiment x Absolute Distance x Time	23	24514	24642	-12234	24468
RTmean ~ Exp + stmmAbs *Time + (1+stmmAbs*Time SubjectID)	Experiment + Absolute Distance + Time + Absolute Distance x Time	17	24507	24601	-12236	24473

5. Discussion

The purpose of our study was to provide a conclusive answer to the question of whether the mental representation of time is continuous or categorical in an explicit time judgement task. In order to do so, we asked participants to place 20 Italian temporal expressions on a line and then used the position of each term on the line to assess both the Distance and STEARC effects in three experiments with increasing saliency of the lateral spatial dimension. The results showed that Italian speakers placed the 20 temporal expressions in an ordered pattern from left to right along the line. Even if the temporal meanings were not perfectly spaced at regular intervals along the line, they covered a broad spectrum and could be used as a continuous predictor. Then, in order to verify our hypothesis, we contrasted the predictive ability of the continuous vs. the categorical predictor of the RT difference between the right and the left hands. The results were very clear. Responses with the right hand were faster for future than past terms compared to responses of the left hand, confirming the STEARC effect (see Bonato et al., 2012, for a review). Crucially, the model that only included the categorical predictor Time (past vs. future) showed a better fit than the model including only the continuous Distance predictor. Moreover, when both predictors were included simultaneously in the regression, the continuous predictor made no contribution to model fit over and above that made by the categorical predictor. Therefore, the STEARC effect is categorical, and not continuous, in explicit time judgement tasks. However, the results also showed a Distance effect, which is continuous by its own nature. The fact that the very same responses that reveal a categorical STEARC effect when the latencies of the right hand are subtracted from the left hand also reveal a continuous Distance effect when considered independently of the responding hand suggests that both effects reveal the characteristics of a single underlying representation. If this representation were categorical, it could not support a continuous effect such as the Distance effect. However, a continuous representation can be used categorically if a criterium is set and used to make binary decisions. Therefore, the present findings can be best interpreted as the categorical use of a continuous mental representation.

This result runs contrary to the findings by Ding et al. (2015), who observed a greater STEARC effect when people considered past events that were distant than when they were closer to the present. The present data show that the magnitude of the STEARC effect does not depend on the temporal distance from the present, neither toward the past nor toward the future. So far, the reasons for this discrepancy remain unclear. The temporal intervals used by Ding et al. (2015; a day and a year) were well within the range used in the present experiments. Other obvious differences are the

language (Chinese vs. Italian) and the cultures (Western vs. East Asian). However, we cannot think of any reason why these factors could produce a continuous SNARC effect in China, but a categorical one in Italy.

The present results are, on the contrary, perfectly consistent with the study by Gevers et al. (2006), who observed, in the numerical domain, that the size of the SNARC effect varied continuously with number magnitude in a parity judgment task (implicit task) but was categorical in a magnitude comparison task (explicit task), suggesting an underlying continuous representation that can be used categorically when the task imposes explicit magnitude comparisons with a criterion. For a perfect parallelism between the domains of number and time, the STEARC effect should take a continuous form in implicit temporal tasks, as reported by Lakens et al. (2011), and by Gevers et al. (2003, 2004) in implicit order tasks. In order to test this final prediction, future research should replicate these studies and expressly assess the categorical vs. continuous form of the effect. In any case, the present study demonstrated a further similarity between the SNARC and STEARC effects, suggesting that similar mapping mechanisms are used to mentally represent the abstract concepts of time and number.

The present data also reveal interesting aspects regarding the continuous underlying representation of time, as they confirm that the Distance effect is not symmetrical toward the past and the future: latencies decreased with distance from the present at a greater rate toward the past than toward the future. In other words, psychological distance toward the future grew more slowly than toward the past. This was so in spite of the fact that most future references were placed further away from the subjective center than past references in the Time-to-Position task. Moreover, latencies for future temporal references were overall faster than those for the past. It is tempting to interpret this pattern as the result of cultural effects on temporal cognition, as Western cultures pay more attention to the future than to the past (de la Fuente, Santiago, Román, Dumitrache, & Casasanto, 2014; Callizo-Romero et al., 2020); they feel that the future is closer than the past (Caruso, Van Boven, Chin, & Ward, 2013), give it a greater emotional and economic valuation (Caruso, Gilbert, & Wilson, 2008; Molouki, Hardisty, & Caruso, 2019), and feel more continuity with their future than past selves (Quoidbach, Gilbert, & Wilson, 2013). Some studies with Chinese participants (but not all) have found the opposite pattern (Guo et al., 2012; Guo & Spina, 2019; Ji, Guo, Zhang, & Messervey, 2009; see Gao, 2016, for a balanced review). Lacking a cross-cultural comparison, this possibility must remain an interesting speculation for now. Additionally, this asymmetrical pattern could be explained by the age of the participants. According to de la Fuente,

Santiago, Román, Dumitrache, and Casasanto (2014; see also Bylund, Gyga, Samuel, and Athanopoulos, 2020), older participants focus less on the future and more on the past than younger participants. Since all our participants were university students, their young age may have influenced their future temporal horizons. Further studies are needed to investigate this issue.

An additional, and surprising finding of the present study is that Experiment 3, using mirror-reversed text, failed to replicate the findings by Casasanto and Bottini (2014). These authors found that a short exposure to mirror-reversed instructions and stimuli reversed the direction of participants' MTL in the second block of their study (after 48 trials), which they explained in terms of representational flexibility. In our data, the STEARC effect did not vary across experiments. This was so even when only the second half of all three experiments was analyzed by means of ANOVA ($F < 1$). What is the cause of this failure of replication? There were several procedural differences between Casasanto and Bottini (2014) and the present Experiment 3: first, they presented the instructions in mirror-reversed text whereas we did not; second, they used temporal sentences with a fixed structure (e.g., "a month before", "a year before", "a month after", "a year after") whereas we used adverbial temporal expressions; third, they presented each stimulus only once per block whereas we presented them nine times; fourth, they presented all materials centered on the screen whereas we presented them in left, center, and right positions. Even though we did not mirror-reverse the instructions, we provided much more practice with the experimental stimuli during the experiment (360 versus 86 trials), so we doubt that insufficient exposure to mirror text was an issue. It does not seem likely to us that the differences in materials and screen presentation could block the reversal of the effect. We think that the best candidate for the differential findings is stimulus repetition: by repeating the expressions nine times, participants may have adopted a holistic stimulus recognition strategy. By doing so, they would identify the expression and access its temporal meaning without actually proceeding from right to left over the letters of the stimulus. Thus, they were not having a right-to-left sensory-motor experience that could reverse the directionality of the MTL. If we are correct, future studies on the effect of mirror-reversing text on temporal cognition should be careful to avoid repeating the materials up to the point of making them recognizable as wholes. It is worth noting that ours is not the only study which has failed to replicate Casasanto and Bottini (2014). Yang, Feng, Jin, Liu, and Sun (2020) showed that in Japanese speakers, who are used to reading both horizontally and vertically, mirror reading reversed the vertical, but not the horizontal time line. These authors suggest that representational flexibility may depend on the relative practice using a script with a given orientation. . Although this

possibility does not apply to the present study (our participants only had practice with a left-to-right script), Yang et al.'s (2020) study suggests that we still do not know all the factors that allow for or prevent reversals of the MTL as a result of mirror reading.

In conclusion, the present study located the position of 20 temporal abstract concepts on a spatial line, from the far past to the far future, and used it to predict the size of the STEARC effect and the Distance effect in an explicit temporal judgement task. The STEARC effect showed a categorical pattern, but the Distance effect suggested that the underlying mental representation of time is continuous. This continuous MTL can be used categorically when the task requires a fixed criterium to be set in order to support temporal decision making.

Acknowledgement

Julio Santiago was supported by FEDER/Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación, Project PGC2018-096096-B-I00.

Declarations

Ethical Standards: The manuscript does not contain clinical studies or patient data.

Consent to participate: Informed consent was obtained from all individual participants included in the study.

REFERENCES

Anelli, F., Peters-Founshtein, G., Shreibman, Y., Moreh, E., Forlani, C., Frassinetti, F., & Arzy, S. (2018). Nature and nurture effects on the spatiality of the mental time line. *Scientific Reports*, 8:11710, doi: 10.1038/s41598-018-29584-3.

Arzy, S., Adi-Japha, E., & Blanke, O. (2009). The mental time line: An analogue of the mental number line in the mapping of life events. *Consciousness and Cognition*, 18, 781-785, doi: 10.1016/j.concog.2009.05.007

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, 4(328), doi: 10.3389/fpsyg.2013.00328.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278, doi: 10.1016/j.jml.2012.11.001.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv, arXiv:1506. doi: arXiv:1506.04967.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48, doi: 10.18637/jss.v067.i01.

Bergen, B.K., & Chan Lau, T.T. (2012). Writing direction affects how people map space onto time. *Frontiers in Psychology*, 3, 109, doi: 10.3389/fpsyg.2012.00109.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience and Behavioral Reviews*, 36, 2257-2273, doi: 10.1016/j.neubiorev.2012.08.007.

Borghi, A. M., & Setti, A. (2017). Abstract concepts and Aging: An Embodied and Grounded Perspective. *Frontiers in Psychology*, 8, 430, doi: 10.3389/fpsyg.2017.00430.

Borghi, A. M (2020). A Future of Words: Language and the Challenge of Abstract Concepts. *Journal of Cognition*, 3(1), 42, doi: 10.5334/joc.134

Bylund, E., Gygax, P., Samuel, S., & Athanasopoulos, P. (2020). Back to the future? The role of the temporal focus for mapping time onto space. *Quarterly Journal of Experimental Psychology*, 73(2), 174-182, doi: 10.1177/1747021819867624

Callizo-Romero, C., Tutnjevic, S., Pandza, M., Ouellet, M., Kranjec, A., Ilić, S., Gu, Y., Goksun, T., Chahboun, S., Casasanto, D., Santiago, J. (2020). Temporal focus and time specialization across cultures. *Psychonomic Bulletin & Review*, 27, 1247-1258, doi: 10.358/s13423-020-017602-5.

Caruso, E. M., Gilbert, D. T., & Wilson, T. D. (2008). A wrinkle in time: Asymmetric valuation of past and future events. *Psychological Science*, 19(8), 796-801, doi: 10.1111/j.1467-9280.2008.02159.x.

Caruso, E. M., Van Boven, L., Chin, M., & Ward, A. (2013). The temporal Doppler effect: When the future feels closer than the past. *Psychological Science*, 24(4), 530-536, doi: 10.1177/0956797612458804.

Casasanto, D., & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General*, 143, 473-479, doi: 10.1037/a0033297.

Casasanto, D. & Pitt, B. (2019). The Faulty Magnitude Detector: Why SNARC-Like Tasks Cannot Support A Generalized Magnitude System. *Cognitive Science*, 43, doi: 10.1111/cogs.12794.

de la Fuente, J., Santiago, J., Román, A., Dumitrache, C., & Casasanto, D. (2014). When you think about it, your past is in front of you: How culture shapes spatial conceptions of time. *Psychological Science*, 25(9), 1682-1690, doi: 10.1177/0956797614534695.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44, 1-42, doi: 10.1016/0010-0277(92)90049-N.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General*, 122, 371–396, doi: 10.1037/0096-3445.122.3.371.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or Linear? Distinct Intuitions of Number Scale in Western and Amazonian Indigene Cultures. *Science*, 320, 1217, doi: 10.1126/science.1156540

Di Bono, M.G., Casarotti, M., Gava, L., Umiltà, C. & Zorzi, M. (2012). Priming the Mental Time Line. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 838-842, doi: 10.1037/a0028346.

Ding, X., Feng, N., Cheng, X., Liu, H., & Fan, Z. (2015). Are past and future symmetric in mental time line? *Frontiers in Psychology*, 6, 208, doi: 10.3389/fpsyg.2015.00208.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year old children: Evidence of a segmented linear model. *Journal of Experimental Child Psychology*, 99, 1-17, doi: 10.1016/j.jecp.2007.08.006.

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, doi: 10.1016/j.actpsy.2011.09.006.

Fabbri, M., Cellini, N., Martoni, M., Tonetti, L., & Natale, V. (2013a). Perceptual and motor congruency effects in time-space association. *Attention, Perception, & Psychophysics*, 75, 1840-1851, doi: 10.3758/s13414-013-0519-9.

Fabbri, M., Cellini, N., Martoni, M., Tonetti, L., & Natale, V. (2013b). The mechanisms of space-time association: comparing motor and perceptual contributions in time reproduction. *Cognitive Science*, *37*, 1228-1250, doi: 10.1111/cogs.12038.

Fabbri, M., & Guarini, A. (2016). Finger counting habit and spatial-numerical association in children and adults. *Consciousness and Cognition*, *40*, 45-53, doi: 10.1016/j.concog.2015.12.012.

Fabbri, M., & Natale, V. (2016). Finger Counting and (2D:4D) Digit Ratio in Spatial-Numerical Association. *Perception*, *45*(1-2), 136-155, doi: 10.1177/0301006615594913.

Fias, W. (2001). Two routes for the processing of verbal numbers: Evidence from the SNARC effect. *Psychological Research*, *65*, 250-259, doi: 10.1007/s004260100065.

Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of the magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition*, *2*, 95-110, doi: 10.1080/135467996387552.

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial-numerical associations. *Brain and Cognition*, *72*(3), 333-336, doi: 10.1016/j.bandc.2009.10.010

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). SAGE Publications.

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural Differences in Mental Representation of Time: Evidence From an Implicit Nonlinguistic task. *Cognitive Science*, *34*, 1430-1451, doi: 10.1111/j.1551-6709.2010.01105.x

Gao, X. (2016). Cultural differences between East Asian and North American in temporal orientation. *Review of General Psychology*, *20*(1), 118–127, doi: [10.1037/gpr0000070](https://doi.org/10.1037/gpr0000070).

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition*, *87*, B87-B95, doi: 10.1016/S0010-0277(02)00234-2.

Gevers, W., Reynvoet, B., & Fias, W. (2004). The mental representation of ordinal sequences is spatially organised: evidence from days of the week. *Cortex*, *40*, 171-172, doi: 10.1016/s0010-9452(08)70938-9.

Gevers, W., Lammertyn, J., Notebaert, W., Verguts, T & Fias, W. (2006). Automatic response activation of implicit spatial information: Evidence from the SNARC effect. *Acta Psychologica*, *122*, 221-233, doi: 10.1016/j.actpsy.2005.11.004.

Guo, T., Ji, L.-J., Spina, R., & Zhang, Z. (2012). Culture, temporal focus, and values of the past and the future. *Personality & Social Psychology Bulletin*, *38*(8), 1030-1040, doi: 10.1177/0146167212443895.

Guo, T., & Spina, R. (2019). Cultural Asymmetry Between Perceptions of Past and Future Personal Change. *Frontiers in Psychology*, *10*, doi: 10.3389/fpsyg.2019.00885.

Hubbard, E.M., Piazza, M., Pinel, P. & Dehaene, S. (2005). Interactions Between Number and Space in Parietal Cortex. *Natural Reviews Neuroscience*, *6*(6), 435-448, doi: 10.1038/nrn1684.

Ishihara, M., Keller, P.E., Rossetti, Y., Prinz, W., 2008. Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex*, *44*, 454-461, doi: 10.1016/j.cortex.2007.08.010.

Ji, L.-J., Hong, E. K., Guo, T., Zhang, Z., Su, Y., & Li, Y. (2019). Culture, psychological proximity to the past and future, and self-continuity. *European Journal of Social Psychology*, *49*(4), 735-747, doi: 10.1002/ejsp.2544.

Keus, I.M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition*, *33*(4), 681-695, doi: 10.3758/BF03195335.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), doi: 10.18637/jss.v082.i13.

Lakens, D., Semin, G. R., & Garrido, M. V. (2011). The sound of time: cross-modal convergence in the spatial structuring of time. *Consciousness and Cognition*, *20*, 437-443, doi: 10.1016/j.concog.2010.09.020.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, doi: 10.3389/fpsyg.2015.01171.

- Maienborn, C., Alex-Ruf, S., Eikmeier, V., & Ulrich, R. (2015). Do we map remembrances to the left/back and expectations to the right/front of a mental timeline? Space–time congruency effects with retrospective and prospective verbs. *Acta Psychologica*, *156*, 168–178, doi: 10.1016/j.actpsy.2014.11.006.
- Molouki, S., Hardisty, D. J., & Caruso, E. M. (2019). The Sign Effect in Past and Future Discounting. *Psychological Science*, *30*(12), 1674–1695, doi: 10.1177/0956797619876982.
- Moyer, R.S., & Landauer, T.K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520, doi: 10.1038/2151519a0.
- Nuerk, H. C., Bauer, F., Krummenacher, J., Heller, D., & Willmes, K. (2005). The power of the mental number line: How the magnitude of unattended numbers affects performance in an Eriksen task. *Psychology Science*, *47*(1), 34–50.
- Nuerk, H.C., Wood, G., & Willmes, K. (2005). The Universal SNARC effect. The Association between Number Magnitude and Space is Amodal. *Experimental Psychology*, *52*(3), 187–194, doi: 10.1027/1618-3169.52.3.187.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113, doi: 10.1016/0028-3932(71)90067-4.
- Ouellet, M., Santiago J., Israeli, Z., & Gabay S. (2010a). Is the Future the Right Time? *Experimental psychology*, *57*(4), 308–314, doi: 10.1027/1618-3169/a000036.
- Proctor, R.W., & Cho, Y.S. (2006). Polarity Correspondence: A General Principle for Performance of Speeded Binary Classification Tasks. *Psychological Bulletin*, *132*(3), 416–442, doi: 10.1037/0033-2909.132.3.416
- Quoidbach, J., Gilbert, D. T., & Wilson, T. D. (2013). The End of the History Illusion. *Science*, *43*(6115), 96–98, doi: 10.1126/science. 1229294.
- R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- Santiago, J., Lupiáñez, J., Perez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*, 512–516, doi: 10.3758/BF03194099.

Santiago, J., Román, A., Ouellet, M., Rodríguez, N., & Pérez-Azor, P. (2010). In hindsight, life flows from left to right. *Psychological Research*, *74*, 59–70, doi: 10.1007/s00426-008-0220-0.

Santiago, J., Román, A., & Ouellet, M. (2011). Flexible foundations of abstract thought: A review and a theory. In A. Maas & T. Schubert (Eds.), *Spatial dimensions of social thought* (pp. 41-110). Berlin: Mouton de Gruyter.

Santiago, J., & Lakens, D. (2015). Can conceptual congruency effect between number, time, and space be accounted for by polarity correspondence? *Acta Psychologica*, *156*, 179-191, doi: 10.1016/j.actpsy.2014.09.016

Shaki, S., Fischer, M.H., & Petrusic, W.M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review*, *16*(2), 328-331, doi: 10.3758/PBR.16.2.328

Schwarz, V. & Keus, I.M. (2004). Moving the eyes along a mental number line: Comparing SNARC effects with saccadic and manual responses. *Perception & Psychophysics*, *66*(4), 651-664, doi: 10.3758/BF03194909.

Siegler R. S., & Opfer, J.E. (2003). The Development of Numerical Estimation: Evidence for Multiple Representations of Numerical Quantity. *Psychological Science*, *14*(3), 237-243, doi: 10.1111/1467-9280.02438.

Torralbo, A., Santiago, J., & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science*, *30*, 745-757, doi: 10.1207/s15516709cog0000_67.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, *23*, 515-557, doi: 10.1016/0010-0285(91)90005-9.

Ulrich, R., & Maienborn, C. (2010). Left-right coding of past and future in language: The mental timeline during sentence processing. *Cognition*, *117*, 126-138, doi: 10.1016/j.cognition.2010.08.001.

Ulrich, R., Eikmeier, V., Vega, I., Ruiz Fernández, S., Alex-Ruf, S., & Maienborn, C. (2012). With the past behind and the future ahead: Back-to-front representation of past and future sentences. *Memory & Cognition*, *40*(3), 483–495, doi: 10.3758/s13421-011-012-4.

Vallesi, A., Binns M.A., & Shallice, T. (2008). An effect of spatial-temporal association of response code: Understanding the cognitive representations of time. *Cognition*, *107*, 501-527, doi: 10.1016/j.cognition.2007.10.011.

Vallesi, A., McIntosh, A.R., & Stuss, D.T. (2011). How time modulates spatial responses. *Cortex*, *47*, 148-156, doi: 10.1016/j.cortex.2009.09.005.

Vallesi, A., Weisblatt, Y., Semenza, C., & Shaki, S. (2014). Cultural modulations of space-time compatibility effects. *Psychonomic Bulletin & Review*, *2*, 666-669, doi: 10.358/s13423-013-0540-y.

Wang, Q., Hou, Y., Tang, H., & Wiprovnick, A. (2011). Travelling backwards and forwards in time: Culture and gender in the episodic specificity of past and future events. *Memory*, *19*(1), 103-109, doi: 10.1080/09658211.2010.537279.

Yang, W., Feng, X., Jin, J., Liu, Y., & Sun, Y. (2020). Can mirror reading reverse the flow of time? Evidence from Japanese speakers. *Psicologia: Reflexão e Crítica*, *33*, 19, doi: 10.1186/s41155-020-001

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 Illiterates Arabic Speakers. *Journal of Cognition and Culture*, *51*(1-2), 165-190, doi: 10.1163/1568537054068660