

Temporal-order Iconicity Bias in Narrative Event Understanding and Memory

Xinming Xu ^a and Sze Chai Kwok ^{a, b, c*}

^a *Shanghai Key Laboratory of Brain Functional Genomics, Key Laboratory of Brain Functional Genomics Ministry of Education, School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China;* ^b *Shanghai Key Laboratory of Magnetic Resonance, East China Normal University, Shanghai 200062, China;* ^c *NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai 200062, China*

*Corresponding Author: Sze Chai Kwok, East China Normal University, Room 269, Geography Building, 3663 Zhongshan Road North, Shanghai 200062, China; Email: sze-chai.kwok@st-hughs.oxon.org

This work was supported by Ministry of Education of PRC Humanities and Social Sciences Research grant 16YJC190006; STCSM Shanghai Pujiang Program 16PJ1402800; STCSM Natural Science Foundation of Shanghai 16ZR1410200; Fundamental Research Funds for the Central Universities 2018ECNU-HWFW007; JoRISS Incubating Project 2017/4; NYU Shanghai and the NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai (S.C.K.).

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Temporal-order Iconicity Bias in Narrative Event Understanding and Memory

Incongruence between the narrated (encoded) order and the actual chronological order of events is ubiquitous in various kinds of narratives and information modalities. The iconicity assumption in text comprehension proposes that readers will by default assume the chronological order to match the narrated order. However, it is not clear whether this iconicity assumption would directly bias the inferred chronology of events and the memory of the narrated order. In the current study, using non-linearly narrated video narratives as encoding materials, we dissociated the narrated order and the underlying chronological order of events. In Experiment 1, we found that participants' judgments of the chronological order of events were biased by the narrated order, but not vice versa. In Experiment 2, when the chronological positions of events were provided during encoding, participants' judgments of the chronological order were not biased by the narrated order, rather, their memory of the narrated order of events was biased by the chronological order. Interpreting the bias under a descriptive Bayesian framework, we offer a new perspective on the role of the iconicity assumption as prior belief, apart from prior knowledge about event sequences, in event understanding as well as memory.

Keywords: event cognition, temporal order, narratives, iconicity, Bayesian inference

I used to think this was the beginning of your story. Memory is a strange thing. It doesn't work like I thought it did. We are so bound by time, by its order.

— *Arrival*, 2016

As time flows in one direction, events are thereby bound to occur in a fixed order and memories of events are theorised to be temporarily organised (Davachi & DuBrow, 2015; Eichenbaum, 2013; Ranganath & Hsieh, 2016). However, when these events are reported as and in narratives they could become “unbound by time” such that the narrated order and thus the encoded order of events might not be congruent with the actual chronological order of events in which they occurred. This dissonance frequently appears in news exposition, social media, daily conversation, or more broadly

in any information modality that depicts events and when people vicariously experience them. Researchers have made clear distinction between the structure of events happening in the real world and the structure in which they are narrated (Bordwell, 2013; Chatman, 1980; Genette, 1983; see also Cameron, 2008; Cutting, 2016). For example, the Russian Formalists referred to the chronological order of the events in the story as *fabula*, and the order they are narrated as *syuzhet*, while Chatman (1980) referred to this distinction as *story* and *discourse*. In this article, we refer to the order of events in the story world as chronological order, and the order of events in the narration (i.e., the order predetermined by the story makers) as narrated order; narrated order is essentially the order experienced and encoded by the viewers (e.g., Claus & Kelter, 2006; Rinck, Hähnel, & Becker, 2001; Zwaan & Radvansky, 1998).

Having correct representation of the chronological order of narrated events (i.e., the *fabula*) is crucial for comprehending events structure, as causes must precede their effects (Lagnado & Sloman, 2006) and real life events often unfold in a meaningful and structured way (Clewett & Davachi, 2017; Ekstrom, Copara, Isham, Wang, & Yonelinas, 2011). Thus, misordering the chronology of events would distort event perception and understanding. In text narratives, with the help of temporal markers (e.g., temporal connectives such as *before* and *after*), readers will construct a chronologically organised representation of a sequence of events containing flashbacks (Claus & Kelter, 2006; Foos, 1992). However, temporal markers may not be available in narratives, such as in films, as films do not have a built-in tense system as language does (Henderson, 1983). In such cases, the viewers therefore need to infer the underlying chronological order of the narrated events.

In text comprehension research, it is proposed that readers will assume by default that events in a text are presented in correspondence to the chronological order (Claus & Kelter, 2006; Zwaan, 1996; Zwaan & Radvansky, 1998), that is, the two orders are “iconic”, known as the iconicity

assumption. Evidence for iconicity assumption is that events narrated out of chronological order will lead to longer reading time (Mandler, 1986), worse comprehension (Ohtsuka & Brewer, 1992) slower causal relation processing (Briner, Virtue, & Kurby, 2012), and less accurate in reporting the chronological order (Baker, 1978) of misordered events, compared to linearly-ordered events.

However, these aforementioned studies have used relatively short text narratives in which temporal markers (e.g., *before* and *after*) were used in depicting events, thereby making chronological information directly available. The worse performance in misordered sequences compared to chronological sequences may simply reflect worse memory of the misordered sequences. Thus, it remains unclear whether the iconicity assumption would directly result in mistaking the narrated order for the chronological order when the two orders are incongruent.

Alternatively, it is possible that the iconicity assumption would in turn bias the memory of the narrated order of events, that is, participants might mistake what chronologically earlier for narrated earlier when the two orders are incongruent. Bower and colleagues (Bower, Black, & Turner, 1979) let participants remember the presentation order of ordered and unordered action scripts and they found that actions presented away from their canonical location were positioned less accurately during rearrangement, and were biased to their canonical location. However, it is not known whether such bias would also be implicated in the memory for narrative events, of which the chronological structure is determined in the absence of any canonical scripts.

In the current study, using long video narratives as materials, we tested the role of the iconicity assumption in event understanding and memory for narrative events, when the narrated order and the chronological order of events were dissociated. We made the participants encode a non-linear narrated video narrative. In Experiment 1, no chronological information was explicitly provided. In Experiment 2, chronological information was provided at the beginning of each event (see Methods).

At test, participants were instructed to judge the chronological order of events (Experiment 1a and Experiment 2a) and the narrated order of events (Experiment 1b and Experiment 2b). The narrated order and the chronological order of a pair of events were either congruent (the event narrated earlier indeed happened earlier in the narrative) or incongruent. If the iconicity assumption plays a role in the representation of chronological order of events, judgements on the chronological order would be biased by the narrated order, leading to lower performance in the incongruent condition compared to the congruent condition. Alternatively, if the iconicity assumption plays a role in the memory for the narrated order of events, then we would expect the narrated order judgments of events to be biased by the chronological order.

Experiment 1

In Experiment 1, the encoding material is a non-linearly narrated video narrative in which the chronological information of each event is not explicitly provided. In Experiment 1a, we tested participants' inference of the chronological order of events. If iconicity plays a role in judging the chronological order, then the performance in the incongruent condition should be lower than that in the congruent condition. If the chronological order and the narrated order of events are independently represented, then performance would not differ between the congruent and incongruent condition. In Experiment 1b, we tested participants' memory of the narrated order of events. If iconicity plays a role in judging the narrated order, then the performance in the incongruent condition should also be lower than that in the congruent condition.

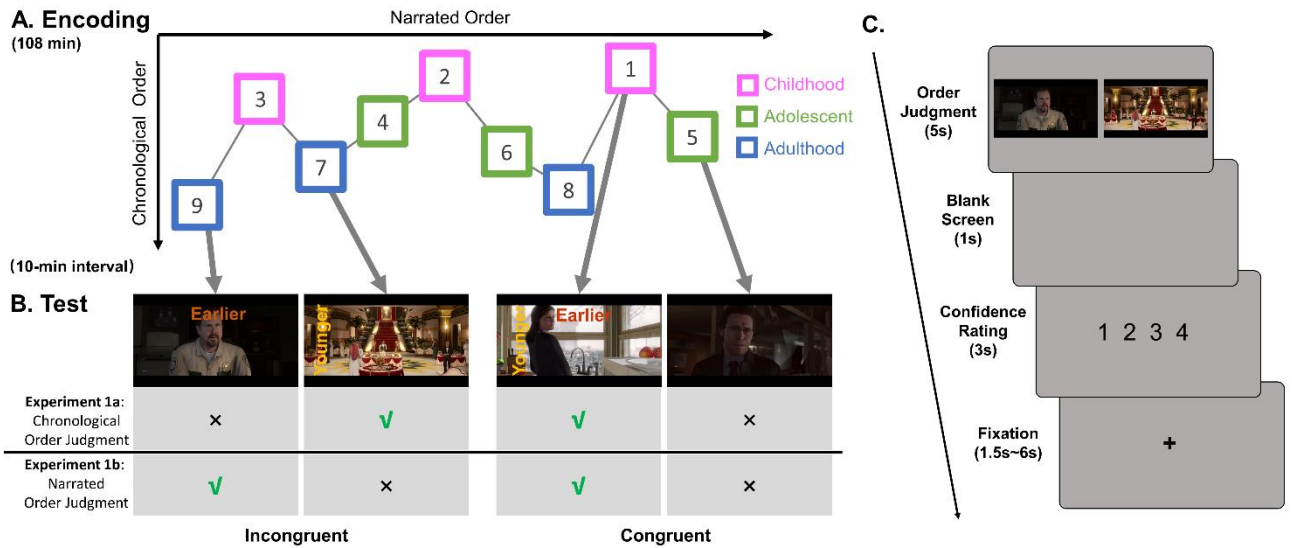


Fig. 1. Experiment 1 overview. (A) At encoding, participants watched a non-linearly narrated game video containing nine chapters (shown in squares). The number in the squares indicates the sequence order of each chapter in chronology. The colour of the squares indicates in which era in the life of the main character the chapter is narrated. (B) At test, in each trial, two probe images were presented side by side. The probe images did not contain the main character or subtitles to discourage participants from using additional cues to judge the chronological order. Participants were required to choose either the scene that was chronologically earlier in the story (Experiment 1a) or the scene that was narrated earlier in the video (Experiment 1b). Trials are either in a congruent condition (the event narrated earlier indeed happened earlier) or an incongruent condition. Ticks and crosses denote correct and incorrect choices. (C) In each trial, participants performed an order judgement and provided a confidence rating of their response on a scale of 1-4.

Method

Participants

32 participants took part in this study (Experiment 1a: 16 participants, 13 female, mean = 22.1 years, $sd = 2.4$; Experiment 1b: 16 participants, 13 female, mean = 20.4 years, $sd = 1.4$). All participants were naïve to the material and had normal or corrected-to-normal eyesight. The study

was approved by University Committee on Human Research Protection of East China Normal University. All participants provided informed consent prior to participation.

Material

The material for encoding was the first nine chapters of a gameplay video of an interactive drama action-adventure video game *Beyond: Two Souls* (Dream, Q., 2013). The game begins with the main character (already in her adulthood) trying to put things in order right up to this moment, starting from her childhood, which is all mixed up in her head. We kept the original narrated order (play-order) of the chapters (Fig. 1A). Each chapter depicts an event of the main character, of which three chapters depicting her childhood, three depicting her adolescence, and the remaining three depicting her adulthood. The chapters were separated clearly in the video, with their titles briefly shown at the beginning of each chapter. Subtitles were provided for the participants to better understand the narrative. We determined the correct chronological order of chapters using the built-in chronological timeline provided by the gameplay. The duration of the video was 108 min, with each of the nine chapters ranging between 4 min and 24 min.

The material for the test was a set of 560 still frames extracted from the video, composing 280 image-pairs. The selected probe images did not contain the main character or subtitles to discourage participants from using additional cues to judge the chronological order. The same set of image-pairs was used in both Experiment 1a and 1b. On each trial, the two images always come from different chapters in the video. Half of the trials are in congruent condition (the event narrated earlier indeed happened earlier) and the other half in incongruent condition (Fig. 1B). In addition to congruence, the trial selection was further counterbalanced on whether two images come from the same era (within-era condition) of the main character (e.g., both from childhood) or two different eras (between-era condition) and the narration distance between the images in the video (7 evenly

distributed bins, ranging from “11m 02s - 11m 37s” to “81m 50s – 82m 25s”), with the three factors (Congruence, Era and Narration Distance) fully crossed.

Tasks and procedure

For encoding, participants watched the 108-min video with subtitles on a 21.5-inch, 1920 × 1080 resolution monitor (Dell E2214Hv), with the audio for the video delivered via over-ear headphones (Sennheiser HD 280 PRO). They were informed that their memory of the video would be tested later.

There was an unfilled 10-min interval between encoding and test, during which participants were instructed not to look up relevant content regarding the video. During test, participants performed either one of two experiment tasks which differed only in their instructional demands. In each trial, a pair of images were presented side-by-side. In Experiment 1a, participants were instructed to choose the scene that happened chronologically earlier in the story. In Experiment 1b, another group of participants were instructed to choose the scene that was narrated earlier in the video (Fig. 1B-C). On each trial, participants were allowed a maximum of 6 s to make a response with the right index or middle fingers by pressing either the “left arrow” or “right arrow” key on a keyboard. Following each order judgement, participants reported their confidence level of that order judgement within 3s, by pressing either one of the four keys “1” (very low) to “4” (very high) with the left hand (thumb excluded). Trials were separated by a black fixation cross, with jittered ITIs sampled from a distribution of 1.5 – 6.0s.

The test session contained 280 trials, divided into four blocks of 70 trials. The presentation order of all trials was randomised for each participant, with their left-right response contingency for the target image counterbalanced within blocks. The presentation was delivered by Psychtoolbox (Pelli, 1997) implemented on MATLAB (MathWorks, Natick, MA).

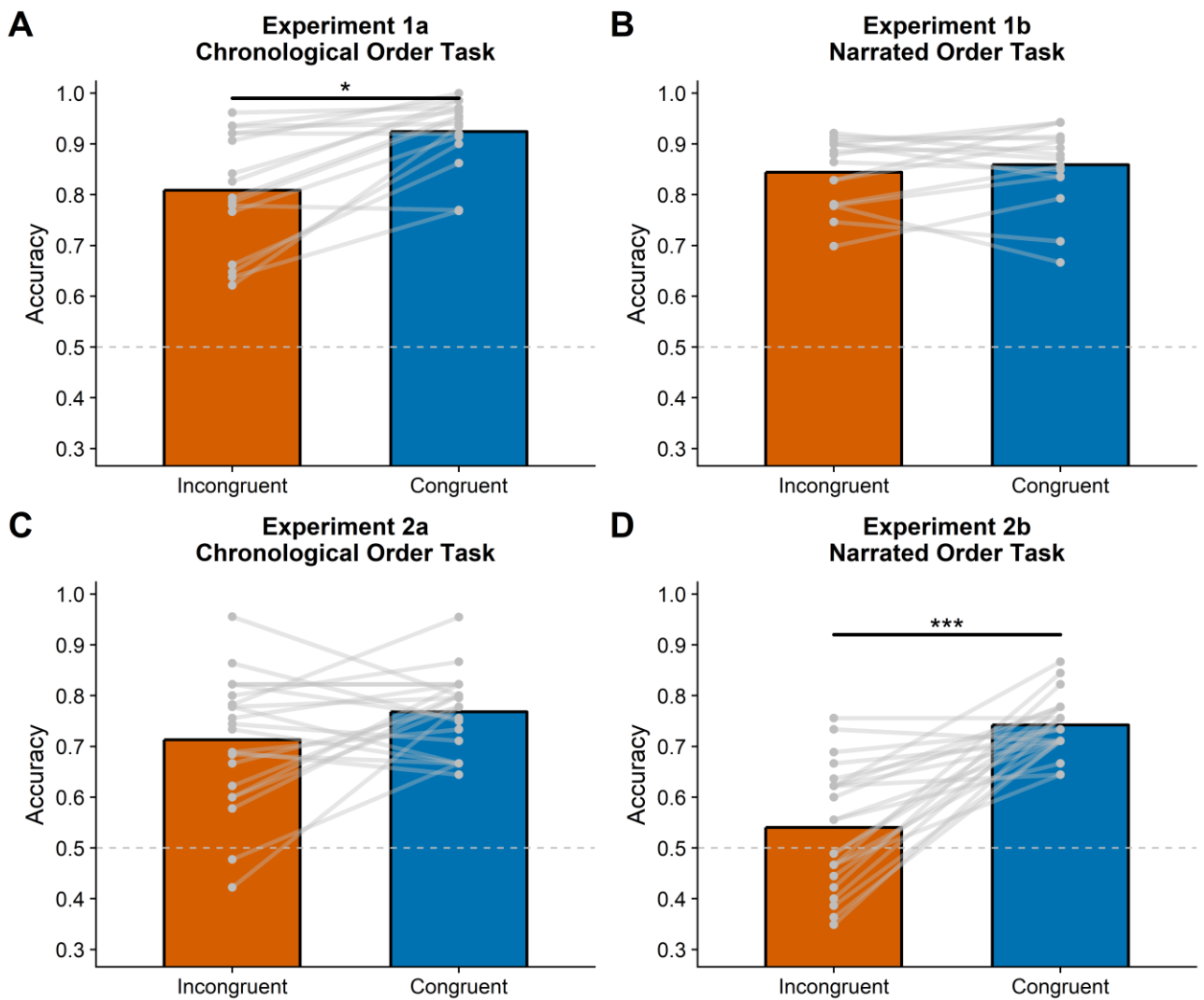


Fig. 2. Order judgement accuracy for the congruent and the incongruent condition in Experiment 1a (Chronological order task), 1b (Narrated order task), 2a (Chronological order task) and 2b (Narrated order task). Each grey line represents performance of one participant. Grey dashed lines denote chance level performance. * $p < 0.05$; *** $p < 0.001$.

Results

Experiment 1a

We analysed the judgement accuracy using generalized linear mixed models (GLMMs) with a logit

link, and RTs and confidence ratings using linear mixed models (LMMs) to conjointly account for participant- and item-related variance (Baayen, Davidson, & Bates, 2008; Jaeger, 2008; Judd, Westfall, & Kenny, 2017). Analyses were performed using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Development Core Team, 2018). Random effects included by-participant random intercepts and random slopes for each fixed factor, as well as random intercepts for each chapter pair and each trial, as trials were nested within chapter pairs. *P* values were extracted using the afex package in R (Singmann, Bolker, Westfall, & Aust, 2018). As we did not specifically predict a linear or quadratic trend of narration distance, we treated narration distance as a categorical variable. RT data from one participant was discarded for failing to respond within the response window.

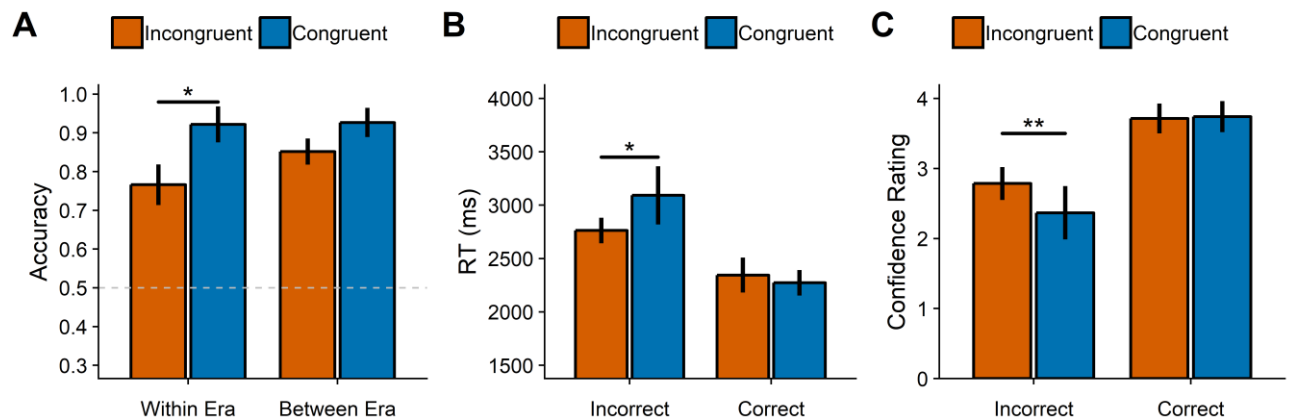


Fig. 3. Experiment 1a results. (A) Accuracy as a function of Congruence and Era. Grey dashed line denotes chance level performance. (B) (C) Response time and confidence rating as a function of response accuracy and Congruence condition. Error bars denote within-subject 95% confidence intervals (Morey, 2008). * $p < 0.05$; ** $p < 0.01$.

A GLMM with Congruence (congruent vs. incongruent) and Era (within-era vs. between-era) as fixed factors revealed a main effect of Congruence, $\chi^2(1) = 5.63$, $p = .02$, odds ratio (OR) = 5.15, indicating that the accuracy in the congruent condition (estimated marginal mean on the response

scale [EMM] = .98) was higher compared to that in the incongruent condition (EMM = .90) (Fig. 2A). The extra judgment errors made in the incongruent condition should be caused by bias induced by the narrated order, thus supporting the iconicity hypothesis that participants would assume the chronological order to match the narrated order. The Era main effect was not significant, between-era: EMM = .96, within-era: EMM = .94, $\chi^2(1) = 0.4$, $p = .53$, OR = 1.52. The effect of congruence was not modulated by Era, Congruency \times Era interaction: $\chi^2(1) = 1.36$, $p = .24$ (Fig. 3A).

We further ascertained that the participants were not fully aware of the bias to the narrated order of their judgments, as reflected by their markedly reduced meta- d' in the incongruent condition. Meta- d' reveals one's ability in discriminating between correct and incorrect responses, which is the predicted value of type 1 performance (d') derived from the observed confidence rating data assuming observers make optimal use of the type 1 information available to them when performing the type 2 confidence judgment task (Fleming & Lau, 2014; Maniscalco & Lau, 2012). We calculated meta- d' at individual level in each Congruence condition. Compared to the congruent trials, incongruent trials were associated with lower meta- d' , $t(15) = 3.37$, $p = .004$, Cohen's $d = 0.84$, implying that participants had a diminished ability to discriminate between correct and incorrect responses when the chronological order and the narrated order conflicted. However, their metacognitive efficiency, in terms of meta- $d' - d'$, was not affected by Congruence, $t(15) = -0.56$, $p = .58$, Cohen's $d = -0.14$.

Moreover, we looked into a specific case in which errors were made in line with the correct narrated order: the incongruent errors. In LMMs with Congruence and Accuracy as fixed factors, we found a significant Congruence \times Accuracy interaction in both RT ($p < .001$) and self-reported confidence ($p = .01$) (Fig. 3B-C). Compared to the error responses in the congruent condition, the error responses in the incongruent condition – that is responses made in line with the correct narrated order – were associated with faster RT ($p = .013$) and higher self-reported confidence ($p = .003$).

Experiment 1b

In Experiment 1b, in the GLMMs, the factor Era was replaced by the factor Narration Distance, which represented the temporal distance between two frames in the video during encoding. Confidence data from two participants were discarded, with one misunderstanding the instruction with no confidence responses elicited, and the other one having a negative meta-d' value.

In contrast to Experiment 1a, the main effect of Congruence was not significant, $\chi^2(1) = 0.15$, $p = .69$, OR = 1.12, indicating that accuracy in the congruent condition (EMM = .93) was not significantly higher than that in the incongruent condition (EMM = .92) (Fig. 2B). The main effect of Narration Distance was significant, $\chi^2(6) = 35.33$, $p < .0001$. There was no interaction between Congruence and Narration Distance, $\chi^2(6) = 4.80$, $p = .57$. Follow-up tests did not show any significant congruence effects in all the seven narration distance levels (all $ps > .249$) (Fig. 4A).

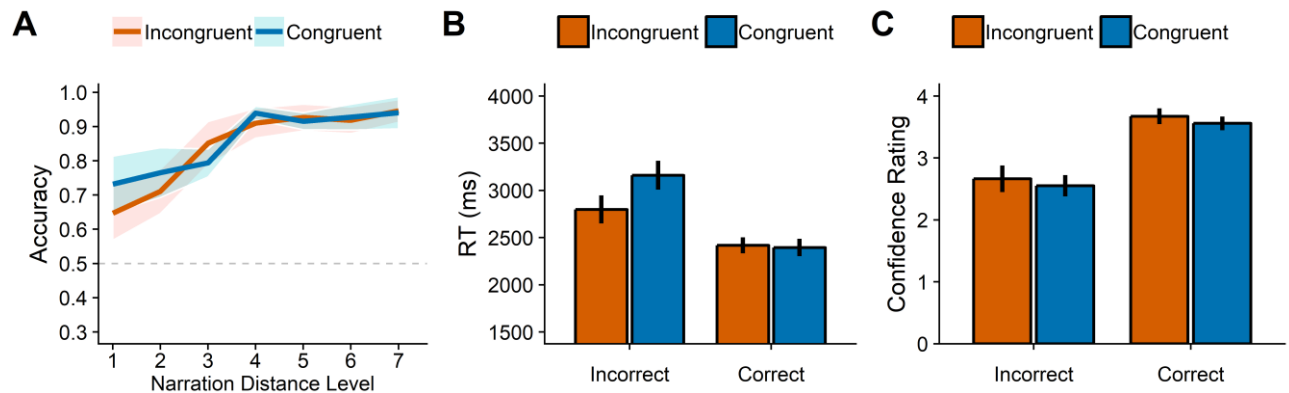


Fig. 4. Experiment 1b results. (A) Accuracy as a function of Congruence and Narration Distance. Grey dashed line denotes chance level performance. (B) (C) Response time and confidence rating as a function of response accuracy and Congruence condition. Error bars denote within-subject 95% confidence intervals (Morey, 2008).

Congruence also did not affect participants' ability in discriminating between their correct and incorrect responses, meta- d' : $t(13) = 0.056, p = .96$, Cohen's $d = 0.01$, or metacognitive efficiency, meta- $d' - d'$: $t(13) = -0.626, p = .54$, Cohen's $d = -0.16$. Congruence did not affect RT and confidence in the correct trials (RT: $p = .91$; confidence ratings: $p = .06$) or in the incorrect trials (RT: $p = .57$; confidence ratings: $p = .73$) (Fig. 4B-C).

Discussion

In Experiment 1, when chronological information was not provided, we found that participants' judgments of the chronological order were biased by the narrated order (Experiment 1a), while judgments of the narrated order were not biased by the chronological order (Experiment 1b). However, the relatively small number of events and the ceiling effect caused by the high narrated order performance may obscure the potential bias caused by the chronological order on the memory of the narrated order. Experiment 2 was designed to address these issues.

Experiment 2

In Experiment 2, we chose another non-linear narrated video narrative as encoding material which contains more discrete events. Unlike Experiment 1, in this narrative, chronological information is provided such that participants will be shown the chronological position of each forthcoming event (see Methods). We anticipated that the bias from the narrated order on the chronological order judgments should be diminished and we will have more power to detect the potential bias, if any, from the chronological order on the judgment of the narrated order of events.

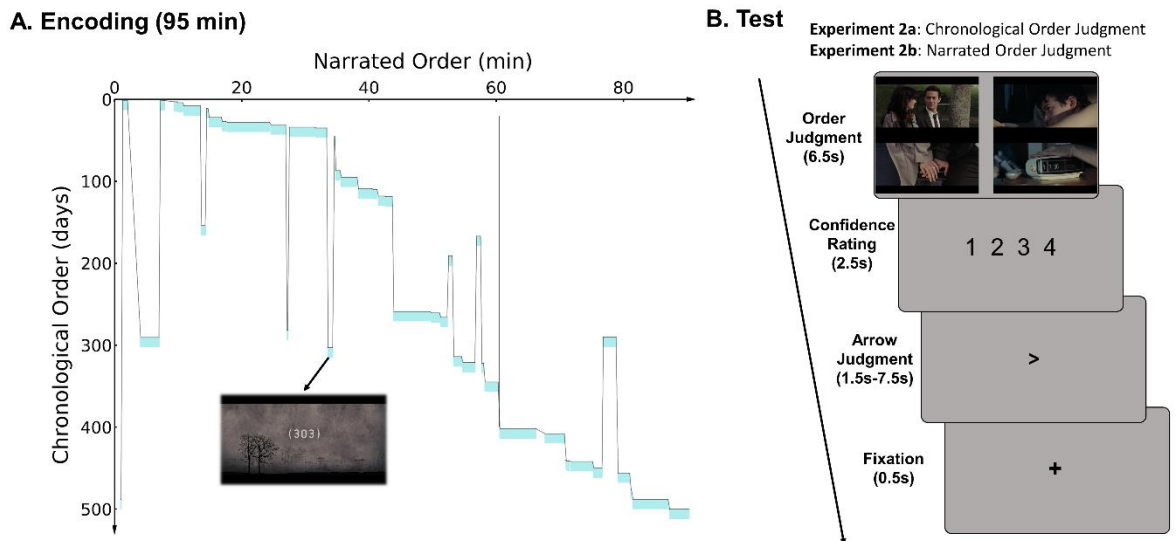


Fig. 5. Experiment 2 overview. (A) At encoding, participants watched a non-linearly narrated film, with an onscreen timer displaying the exact day in chronology at the beginning of each event. (B) At test, participants performed the chronological order task (Experiment 2a) and the narrated order task (Experiment 2b) in succession, with the order of task counterbalanced across participants. (C) In each trial, two pairs of images were presented side by side, depicting two scenes in the film. participants either performed a chronological order judgement or a narrated order judgment and provided a confidence rating of their response on a scale of 1-4.

Methods

Participants

25 participants took part in this study (16 female, mean = 21.4 years, sd = 2.1). Data of two participants were not registered due to machine fault, leaving with 23 participants for further analysis. One further participant's data were discarded in Experiment 2a for chance-level performance in both the congruent and incongruent conditions. All participants were naïve to the material and had normal or corrected-to-normal eyesight. The study was approved by University Committee on Human Research Protection of East China Normal University. All participants provided informed consent

prior to participation.

Material

The encoding material was a 95-min-long feature length film (*500 Days of Summer* (Webb, 2009)). The film was narrated in a nonlinear fashion, jumping between days within the 500 days of Tom and Summer's relationship (Fig. 5A). The whole film is segmented by the days (1-500) in which each event took place. Events from 41 days were shown in the film. There is an onscreen timer showing the day at the beginning of each event. Subtitles were provided for the participants to better understand the narrative. We determined the correct chronological order of events based on the onscreen timer provided.

The material for the test was still frames extracted from the film. For each trial, two pairs of images were presented side by side, representing two scenes from two different days. In each pair, the two images were aligned vertically and were within 15 s in the film. This manipulation was for participants to better recall the event shown in the images. The probe images did not contain any subtitles. Two sets of material, each containing 90 trials, were created. For each task, half of the trials are in the congruent condition (the event narrated earlier indeed happened earlier) and the other half in the incongruent condition. The trials in each set were also counterbalanced on the chronological distance of the events depicted with three levels (short: 1-75 days; medium: 79-178 days; long: 182-439 days), analogues to the factor Era in Experiment 1, and the narration distance between the scenes in the video also with three levels (short: 0-6 min; medium: 6-20.1 min; long: 20.4-76 min), with the three factors (Congruence, Chronological Distance and Narration Distance) fully crossed.

Tasks and procedure

For encoding, participants watched the 95-min film with subtitles on a 21.5-inch, 1920 × 1080

resolution monitor (Dell E2214Hv), with the audio for the video delivered via over-ear headphones (Sony MDR-XB450AP).

The mean interval between encoding and test was 75 min (sd = 22 min). In this unfilled period, participants were instructed not to look up relevant content regarding the video. During test, participants performed the chronological order task (Experiment 2a) and the narrated order task (Experiment 2b) in succession, each containing 90 trials from one set of material, with the order of tasks and the task-set combination counterbalanced across participants. Behavioral data were acquired during a neuroimaging experiment.

Each task was divided into two blocks with 45 trials each. For each trial, two pairs of images were presented side by side, representing two scenes from two different days (Fig. 5B). Participants were asked to either judge the chronological order or the narrated order of the two scenes in each task. Half of the participants judged which scene was earlier (i.e., narrated earlier in narrated order task and chronologically earlier in chronological order task), and half the participants judged which scene was later. Another 15 trials were embedded in each block in which participants were required to judge which scene was more positive or negative in valence, data of which were not reported here. Participants were allowed a maximum of 6.5 s to make a response with the right index or middle fingers via an MRI compatible five-button response keyboard (Sinorad). Following each temporal order judgement, participants reported their confidence level of that order judgement within 2.5 s, from “1” (very low) to “4” (very high) with four fingers (thumb excluded) using the left hand. Trials were separated by a black fixation cross, embedded with an arrow judgment task, with jittered ITIs sampled from a distribution of 2 – 8 s. The visual stimuli were delivered by Psychtoolbox (Pelli, 1997) implemented on MATLAB (MathWorks, Natick, MA), as back-projected via a mirror system to the participant.

Results

Experiment 2a

An analogous analysis approach to Experiment 1 was adopted, except that in Experiment 2 the narration distance and chronological distance were both divided into three levels (short, medium and long), and that random intercepts for each trial were included but without any nested factor given that each trial was selected from a different event pair.

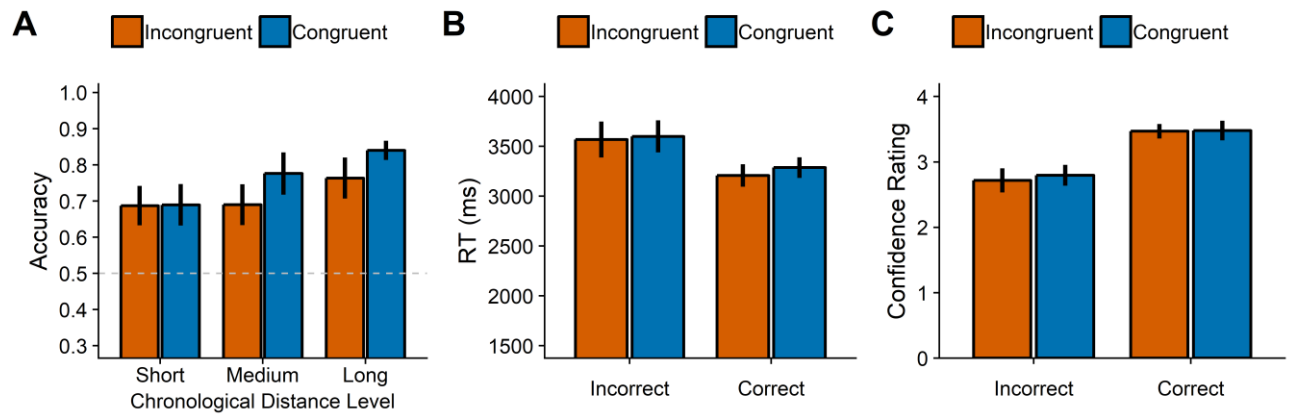


Fig. 6. Experiment 2a results. (A) Accuracy as a function of Congruence and Chronological Distance. Grey dashed line denotes chance level performance. (B) (C) Response time and confidence rating as a function of response accuracy and Congruence condition. Error bars denote within-subject 95% confidence intervals (Morey, 2008).

The main effect of Congruence was not significant, $\chi^2(1) = 2.10$, $p = .15$, odds ratio (OR) = 1.33, indicating that accuracy in the congruent condition (EMM = .80) was not significantly higher than that in the incongruent condition (EMM = .75) (Fig. 2C). The main effect of Chronological Distance was significant, $\chi^2(1) = 8.19$, $p = .02$. The accuracy in the long chronological distance condition (EMM = .85) was higher than that in the short condition (EMM = .71, OR = 2.30, $z = 3.41$,

$p < .002$), but not significantly higher than that in the medium condition (EMM = .77, OR = 1.69, $z = 2.28$, $p = .058$), and accuracy in the medium condition was not significantly higher than that in the short condition (OR = 1.35, $z = 1.48$, $p = .303$). The congruence effect was not observed in any chronological distance level (all $ps > .16$) (Fig. 6A). An alternative GLMM including Congruence and Task Order as fixed effects also revealed no significant interaction between Congruence and Task Order ($p = .89$). These results suggested that participants' chronological order judgments were not biased by the narrated order.

Congruence also did not affect participants' meta- d' , $t(21) = 0.192$, $p = .84$, Cohen's $d = 0.04$, or meta-efficiency, meta- $d' - d'$: $t(21) = -0.928$, $p = .36$, Cohen's $d = -0.19$. Congruence did not affect RT and confidence in the correct trials (RT: $p = .17$; confidence ratings: $p = .92$) or in the incorrect trials (RT: $p = .65$; confidence ratings: $p = .90$) (Fig. 6B-C).

Experiment 2b

Accuracy in the congruent condition (EMM = .79) was higher than that in the incongruent condition (EMM = .57), as revealed by a main effect of Congruence, $\chi^2(1) = 18.69$, $p < .0001$, odds ratio (OR) = 2.88 (Fig. 2D), suggesting that participants' judgments of narrated order were biased by the chronological order. Moreover, as predicted, we found a main effect of Narrated Distance, $\chi^2(2) = 36.03$, $p < .0001$, indicating that accuracy in the long narration distance condition (EMM = .81) was higher than that in the medium condition (EMM = .67, OR = 2.14, $z = 3.11$, $p < .006$), which in turn was marginally higher than that in the short condition (EMM = .55, OR = 1.66, $z = 2.14$, $p = .082$). There was a significant interaction between Congruence and Narrated Distance, $\chi^2(2) = 25.12$, $p < .0001$ (Fig. 7A). The congruence effect was observed in the long narration distance condition (OR = 6.47, $z = 5.20$, $p < .0001$) and the medium condition (OR = 2.93, $z = 3.02$, $p < .003$), but not in the short condition (OR = 1.25, $z = 0.65$, $p = .52$). An alternative GLMM including Congruence and Task

Order as fixed effects revealed no significant interaction between Congruence and Task Order ($p = .97$), suggesting that the bias remained the same irrespective of whether the narrated order task was performed first or otherwise.

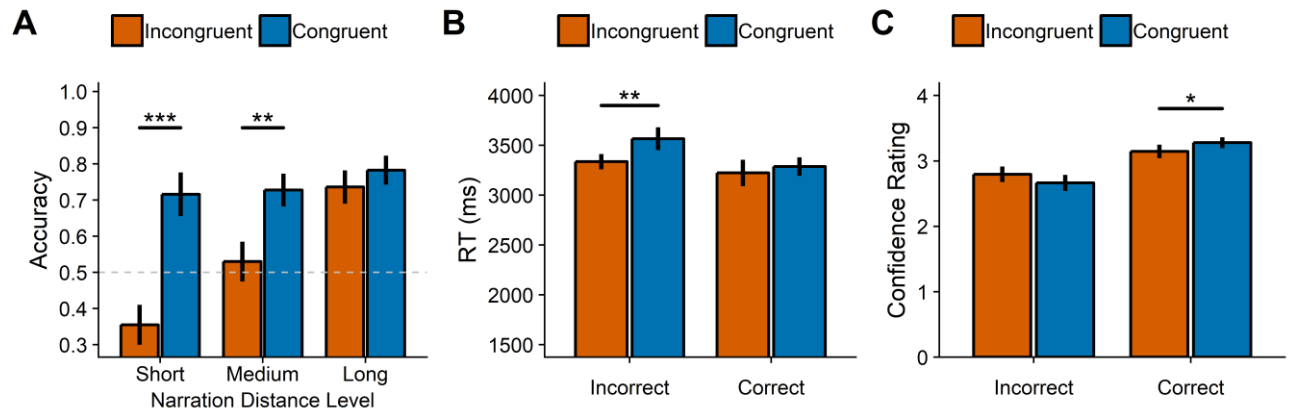


Fig. 7. Experiment 2b results. (A) Accuracy as a function of Congruence and Narration Distance. Grey dashed line denotes chance level performance. (B) (C) Response time and confidence rating as a function of response accuracy and Congruence condition. Error bars denote within-subject 95% confidence intervals (Morey, 2008). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Participants' meta- d' were also lower in the incongruent condition compared to the congruent condition, $t(22) = 4.421$, $p < .001$, Cohen's $d = 0.92$. Meta-efficiency, in terms of meta- $d' - d'$, were higher in the incongruent condition, $t(22) = -2.365$, $p = .03$, Cohen's $d = -0.493$. We observed significant interactions between Congruence and Accuracy in both RT ($p = .03$) and confidence ($p = .02$). Error responses in the incongruent condition were associated with faster response time ($p = .005$) but not with higher confidence ratings ($p = .22$) (Fig. 7B-C). The interaction between Congruence and Accuracy in confidence was driven by higher confidence with correct responses in the congruent condition compared to that in the incongruent condition ($p = .05$).

Discussion

In Experiment 2, when the chronological positions of each event were provided during encoding, participants' judgments of the chronological order were no longer biased by the narrated order (Experiment 2a), rather, we found that their memory of the narrated order was biased by the chronological order (Experiment 2b). In light of a previous study showing that memory of arbitrary action sequences would be biased to the canonically-ordered sequences known as scripts (Bower et al., 1979), our results generalized such bias to narrative events where chronological information was provided. The shift of the observed iconicity bias from the chronological order judgment to the narrated order judgments indicates that the iconicity assumption is implicated in both event understanding and event memory.

General Discussion

We demonstrated the presence of the iconicity bias in the inference of the chronological order and also in the memory of the narrated order of events. We showed that participants' judgments of the chronological order were biased by the narrated order (Experiment 1a) and their judgments of the narrated order were biased by the chronological order (Experiment 2b). These results provide direct evidence that the iconicity assumption (Claus & Kelter, 2006; Zwaan, 1996; Zwaan & Radvansky, 1998) would lead to bias in chronology inference such that participants would assume the chronological order to match the narrated order, as well as in temporal-order memory such that participants would assume the narrated order to match the chronological order.

The bias from the narrated order in the chronological order task was observed in Experiment 1a but not in Experiment 2a. A possible reason is that in Experiment 2a, the chronological positions of events were provided during encoding, which may help eliminate viewers' assumption that the two

orders are congruent. In contrast, the bias from the chronological order in the narrated order task was observed in Experiment 2b but not in Experiment 1b. A possible reason is that in Experiment 1b the narrated order task performance is high which might have caused a ceiling effect, whereas in Experiment 2b wherein the overall narrated order performance did not reach ceiling, we observed a bias from the chronological order.

The presence of iconicity assumption can be conceived as a kind of prior belief about the congruence between the chronological order and the narrated order. We adopt a Bayesian framework (posterior \propto likelihood \times prior) to illustrate the interplay between the two orders, that is, how prior belief would lead to the bias in judging the chronological order as well as the narrated order.

The chronological order task in Experiment 1a and 2a could be regarded as Bayesian inference. For instance, assuming event A is narrated earlier than event B, when judging whether A was chronologically earlier or later than B, one will combine prior knowledge on how likely A happened before B (when no narrated order information is available, modelled as the *prior* term), with the prior belief about how likely the event being chronologically earlier was also narrated earlier (modelled as the *likelihood* term), thus giving the *posterior odds*:

posterior likelihood prior

$$\frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix} \middle| D\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix} \middle| D\right)} = \frac{P\left(D \middle| \begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(D \middle| \begin{smallmatrix} B \\ A \end{smallmatrix}\right)} \times \frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix}\right)},$$

$$\text{while } \frac{P\left(D \middle| \begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(D \middle| \begin{smallmatrix} B \\ A \end{smallmatrix}\right)} = \frac{P(AB|D) * P(\text{Congruent}) + P(BA|D) * P(\text{Incongruent})}{P(AB|D) * P(\text{Incongruent}) + P(BA|D) * P(\text{Congruent})},$$

where vertically arranged $\begin{smallmatrix} A \\ B \end{smallmatrix}$ means A is chronologically earlier than B, and horizontally arranged AB means A is narrated earlier than B. D represents the observation that event A and B were narrated.

The *prior* term represents inference based on heuristic common knowledge about real-life experiences, for example, for a given person in a story, a master degree is more likely to be obtained *before* a PhD degree than vice versa. In this case it does not contain any information about the narrated order. The prior odds ratio may vary in each event pair according to the chronological distance or causal strength between the two events. The *likelihood* term models the influence from the narrated order. If the inference of the chronological order happens during online encoding, when the narrated order is determined, that is, $P(AB|D) = 1$, then the *likelihood* term could be simplified to $\frac{P(\text{Congruent})}{P(\text{Incongruent})}$, which simply represents the subjective belief about the congruence between the chronological order and the narrated order of events. In the incongruent condition, the contribution of the *likelihood* would be inverted to be $\frac{P(\text{Incongruent})}{P(\text{Congruent})}$ where $P(BA|D) = 1$. Assuming the distribution of the prior odds ratio is matched between trials in the congruent and the incongruent condition, the prior belief that favours what being chronologically earlier as also narrated earlier, that is, $\frac{P(\text{Congruent})}{P(\text{Incongruent})} > 1$, would result in the bias observed in Experiment 1a.

In Experiment 2a, when the chronological positions of events were provided during encoding, we did not observe any judgment bias caused by the narrated order. It is possible that when chronological information was available, participants would easily detect the non-linearity of the narrative, thus driving their prior belief to be uninformative such that $\frac{P(\text{Congruent})}{P(\text{Incongruent})} = 1$. Likewise, knowledge about the characteristics of the narratives' temporal structure may dynamically update or even alter the subjective belief or the default assumption on the chronological position of upcoming events (Dery & Koenig, 2015). An extreme scenario is that the prior belief could be inverted to favour what narrated later as chronologically earlier, that is, $\frac{P(\text{Congruent})}{P(\text{Incongruent})} < 1$, which could be caused by

gradually learning the temporal structure of the narrative (e.g., *Memento*¹, 2000) or by some heuristic understanding of narrative schemas (e.g., *Arrival*², 2016).

The narrated order task in Experiment 1b and 2b could also be regarded as Bayesian inference:

posterior likelihood prior

$$\frac{P(AB|D)}{P(BA|D)} = \frac{P(D|AB)}{P(D|BA)} \times \frac{P(AB)}{P(BA)},$$

$$\text{while } \frac{P(AB)}{P(BA)} = \frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix} \middle| D\right) * P(\text{Congruent}) + P\left(\begin{smallmatrix} B \\ A \end{smallmatrix} \middle| D\right) * P(\text{Incongruent})}{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix} \middle| D\right) * P(\text{Incongruent}) + P\left(\begin{smallmatrix} B \\ A \end{smallmatrix} \middle| D\right) * P(\text{Congruent})}.$$

The *likelihood* term now represents memory retrieval of the encoding event sequence. In this case it does not contain any information about the chronological order of events. Unlike the chronological order task, the narrated order task is an offline memory task. Thus, the role of the *prior*, as demonstrated in the model, is determined by both the prior belief about the congruence between the two orders and the subjective chronological representation of events. Note that when the prior belief $\frac{P(\text{Congruent})}{P(\text{Incongruent})}$ is kept constant, the model does not predict the interaction between the Narration Distance and the Congruence factor. However, it is possible that the prior belief might differ across different narration distance levels.

By treating the narrated order task as Bayesian inference, we also highlighted the reconstructive nature of human memory (Bartlett, 1932; Hemmer & Steyvers, 2009) that memory is subject to biases from general knowledge of everyday experiences. We separated the iconicity assumption from the role of prior knowledge such as schemas (Bonasia et al., 2018; van Kesteren, Ruiter, Fernández, & Henson, 2012) in memory judgment processes, and conceive it as a kind of prior belief that functions as non-memorial process in recollecting events (Scoboria et al., 2014).

In the current study, judgments of the chronological order and the narrated order were assessed at only one time point for each encoding material. Studies using naturalistic materials with varying delays have found that faster time-dependent decline rates in memory for perceptual details than memory for central details (Furman, Dorfman, Hasson, Davachi, & Dudai, 2007; Sekeres et al., 2016). It would be expected that performance in the narrated order task would decline rapidly over time, as the narrated order is the surface information during narrative comprehension (Gernsbacher, 1985), whereas performance in the chronological order task would be less susceptible or even remain stable across time, as the chronological order task relies more on prior knowledge. However, memory decline over time may disrupt the current chronological order task when participants are no longer able to identify the events depicted in the images. Assessing the bias in question with a reduced overall task performance would not be trivial; future studies would be required to assess how the prior belief might be altered as a function of delay length.

While we showed that human participants were able to reconstruct the global structure of a narrative which was disrupted during encoding (see also Kintsch, Mandel, & Kozminsky, 1977), event perception and understanding would be considerably hampered when the narrative's fine-grained structure is altered (Hymel, Levin, & Baker, 2016). Instead of scrambling the narrative in the coarse unit of discrete events, with their fine temporal structure preserved, future research could be focussed on specifying the temporal unit and structure of events and its impact on different aspects of event cognition (cf. Ferguson, Homa, & Ellis, 2016).

More broadly, we demonstrated the role of a specific kind of prior belief, namely the iconicity assumption, on event understanding and event memory. This iconicity assumption is implicit in real life first-person experiences when time flows in one direction and the two orders are always congruent.

Thus, retrieving real life event sequences could be treated as a combination of the two order tasks, with the *posterior odds* being:

$$\begin{array}{ccc} \textit{posterior} & \textit{likelihood} & \textit{prior} \\ \frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix} \middle| D\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix} \middle| D\right)} & = \frac{P\left(D \middle| \begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(D \middle| \begin{smallmatrix} B \\ A \end{smallmatrix}\right)} \times \frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix}\right)}, \end{array}$$

where diagonally arranged $\begin{smallmatrix} A \\ B \end{smallmatrix}$ means A is both chronologically earlier and narrated earlier than B.

The *likelihood* term represents evidence from memory retrieval of the encoding order, that is,

$$\frac{P\left(D \middle| \begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(D \middle| \begin{smallmatrix} B \\ A \end{smallmatrix}\right)} = \frac{P(D|AB)}{P(D|BA)},$$

and the *prior* term represents evidence from inference based on prior knowledge, that is,

$$\frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix}\right)} = \frac{P\left(\begin{smallmatrix} A \\ B \end{smallmatrix}\right)}{P\left(\begin{smallmatrix} B \\ A \end{smallmatrix}\right)}.$$

The relatively high performance in both tasks suggests retrieving temporal order of naturalistic experiences could be both supported by actual memory retrieval and inference based on prior knowledge (Friedman, 1993; Jacques, Rubin, LaBar, & Cabeza, 2008; Lehn et al., 2009). The iconicity assumption functions as a bridge linking these distinct cognitive processes, and is the basis for learning event schemas through real-life experiences.

Acknowledgements

We would like to thank Boqiang Zhang for his suggestion for the experimental material for Experiment 1, Xing Tian for his laboratory support during data collection, and Fabrice Berna for helpful comments on an earlier version of the manuscript.

Notes

1. In the film *Memento*, a series of coloured sequence is shown in reverse order. If the viewers have identified this temporal pattern of the narrated events, their prior belief should favour that the next event to be narrated within the coloured sequence is chronologically earlier than the preceding coloured event.
2. In the film *Arrival*, a series of flashforward scenes depicting a mother's pre-cognition of her daughter's birth, childhood and death is easily misinterpreted as flashbacks (Bordwell, 2016). In this case, the viewers' belief might also favour that the daughter-related scenes are chronologically earlier than the previous narrated events, that is, these scenes depict *memory* of the mother.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Baker, L. (1978). Processing temporal relationships in simple stories: Effects of input sequence. *Journal of Verbal Learning and Verbal Behavior*, *17*(5), 559–572. [https://doi.org/10.1016/S0022-5371\(78\)90337-7](https://doi.org/10.1016/S0022-5371(78)90337-7)
- Bartlett, F. (1932). *Remembering: An Experimental and Social Study*. Cambridge: Cambridge University. Retrieved from <http://books.google.com/books?hl=en&lr=&id=WG5ZcHGTrm4C&oi=fnd&pg=PR9&dq=bartlett+remembering&ots=BAeWcuInfI&sig=rFLXsRDPpEtodXcTFIMScqhAPec>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). <https://doi.org/10.18637/jss.v067.i01>
- Bonasia, K., Sekeres, M. J., Gilboa, A., Grady, C. L., Winocur, G., & Moscovitch, M. (2018). Prior Knowledge Modulates the Neural Substrates of Encoding and Retrieving Naturalistic Events at Short and Long Delays. *Neurobiology of Learning and Memory*. <https://doi.org/10.1016/j.nlm.2018.02.017>

- Bordwell, D. (2013). *Narration in the Fiction Film*. Routledge.
- Bordwell, D. (2016, November 23). ARRIVAL: When is Now? Retrieved January 22, 2018, from www.davidbordwell.net/blog/2016/11/23/arrival-when-is-now/
- Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220. [https://doi.org/10.1016/0010-0285\(79\)90009-4](https://doi.org/10.1016/0010-0285(79)90009-4)
- Briner, S. W., Virtue, S., & Kurby, C. A. (2012). Processing Causality in Narrative Events: Temporal Order Matters. *Discourse Processes*, *49*(1), 61–77. <https://doi.org/10.1080/0163853X.2011.607952>
- Cameron, A. (2008). *Modular narratives in contemporary cinema*. Basingstoke [England] ; New York: Palgrave Macmillan.
- Chatman, S. B. (1980). *Story and Discourse: Narrative Structure in Fiction and Film*. Cornell University Press.
- Claus, B., & Kelter, S. (2006). Comprehending narratives containing flashbacks: Evidence for temporally organized representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1031–1044. <https://doi.org/10.1037/0278-7393.32.5.1031>
- Clewett, D., & Davachi, L. (2017). The ebb and flow of experience determines the temporal structure of memory. *Current Opinion in Behavioral Sciences*, *17*(Supplement C), 186–193. <https://doi.org/10.1016/j.cobeha.2017.08.013>
- Cutting, J. E. (2016). Narrative theory and the dynamics of popular movies. *Psychonomic Bulletin & Review*, 1–31. <https://doi.org/10.3758/s13423-016-1051-4>
- Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: the role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. <https://doi.org/10.1016/j.tics.2014.12.004>
- Dery, J. E., & Koenig, J.-P. (2015). A Narrative-Expectation-Based Approach to Temporal Update in Discourse Comprehension. *Discourse Processes*, *52*(7), 559–584. <https://doi.org/10.1080/0163853X.2014.966293>
- Eichenbaum, H. (2013). Memory on time. *Trends in Cognitive Sciences*, *17*(2), 81–88. <https://doi.org/10.1016/j.tics.2012.12.007>
- Ekstrom, A. D., Copara, M. S., Isham, E. A., Wang, W., & Yonelinas, A. P. (2011). Dissociable networks involved in spatial and temporal order source retrieval. *NeuroImage*, *56*(3), 1803–1813. <https://doi.org/10.1016/j.neuroimage.2011.02.033>

- Ferguson, R., Homa, D., & Ellis, D. (2016). Memory for Temporally Dynamic Scenes. *The Quarterly Journal of Experimental Psychology*, 0(ja), 1–42.
<https://doi.org/10.1080/17470218.2016.1174721>
- Foos, P. W. (1992). Constructing Schemata While Reading Simple Stories. *The Journal of General Psychology*, 119(4), 419–425. <https://doi.org/10.1080/00221309.1992.9921183>
- Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, 113(1), 44–66.
<https://doi.org/10.1037/0033-2909.113.1.44>
- Furman, O., Dorfman, N., Hasson, U., Davachi, L., & Dudai, Y. (2007). They saw a movie: Long-term memory for an extended audiovisual narrative. *Learning & Memory*, 14(6), 457–467.
<https://doi.org/10.1101/lm.550407>
- Genette, G. (1983). *Narrative Discourse: An Essay in Method*. Cornell University Press.
- Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, 17(3), 324–363. [https://doi.org/10.1016/0010-0285\(85\)90012-X](https://doi.org/10.1016/0010-0285(85)90012-X)
- Hemmer, P., & Steyvers, M. (2009). A Bayesian Account of Reconstructive Memory. *Topics in Cognitive Science*, 1(1), 189–202. <https://doi.org/10.1111/j.1756-8765.2008.01010.x>
- Henderson, B. (1983). Tense, Mood, and Voice in Film (Notes after Genette). *Film Quarterly*, 36(4), 4–17. <https://doi.org/10.2307/3697090>
- Hymel, A., Levin, D. T., & Baker, L. J. (2016). Default processing of event sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 42(2), 235–246.
<https://doi.org/10.1037/xhp0000082>
- Jacques, P. S., Rubin, D. C., LaBar, K. S., & Cabeza, R. (2008). The short and long of it: Neural correlates of temporal-order memory for autobiographical events. *Journal of Cognitive Neuroscience*, 20(7), 1327–1341. Retrieved from
<http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2008.20091>
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446.
<https://doi.org/10.1016/j.jml.2007.11.007>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with More than One Random Factor: Designs, Analytic Models, and Statistical Power. *Annual Review of Psychology*, 68(1), null.
<https://doi.org/10.1146/annurev-psych-122414-033702>

- Kintsch, W., Mandel, T. S., & Kozminsky, E. (1977). Summarizing scrambled stories. *Memory & Cognition*, 5(5), 547–552. <https://doi.org/10.3758/BF03197399>
- Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3), 451–460. <https://doi.org/10.1037/0278-7393.32.3.451>
- Lehn, H., Steffenach, H.-A., van Strien, N. M., Veltman, D. J., Witter, M. P., & Haberg, A. K. (2009). A Specific Role of the Human Hippocampus in Recall of Temporal Sequences. *Journal of Neuroscience*, 29(11), 3475–3484. <https://doi.org/10.1523/JNEUROSCI.5370-08.2009>
- Mandler, J. M. (1986). On the comprehension of temporal order. *Language and Cognitive Processes*, 1(4), 309–320. <https://doi.org/10.1080/01690968608404680>
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Reason*, 4(2), 61–64. Retrieved from <http://pcl.missouri.edu/sites/default/files/morey.2008.pdf>
- Ohtsuka, K., & Brewer, W. F. (1992). Discourse organization in the comprehension of temporal order in narrative texts. *Discourse Processes*, 15(3), 317–336. <https://doi.org/10.1080/01638539209544815>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>
- Ranganath, C., & Hsieh, L.-T. (2016). The hippocampus: a special place for time. *Annals of the New York Academy of Sciences*, 1369(1), 93–110. <https://doi.org/10.1111/nyas.13043>
- Rinck, M., Hähnel, A., & Becker, G. (2001). Using temporal information to construct, update, and retrieve situation models of narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 67–80. <https://doi.org/10.1037/0278-7393.27.1.67>
- Scoboria, A., Jackson, D. L., Talarico, J., Hanczakowski, M., Wysman, L., & Mazzoni, G. (2014). The role of belief in occurrence within autobiographical memory. *Journal of Experimental Psychology: General*, 143(3), 1242–1258. <https://doi.org/10.1037/a0034110>
- Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. *Learning & Memory*, 23(2), 72–82. <https://doi.org/10.1101/lm.039057.115>

- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). *afex: Analysis of Factorial Experiments*. R package version 0.21-2. <https://CRAN.R-project.org/package=afex>
- van Kesteren, M. T. R., Ruitter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35(4), 211–219. <https://doi.org/10.1016/j.tins.2012.02.001>
- Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1196. <https://doi.org/10.1037/0278-7393.22.5.1196>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>