Behavioral and Eye-Tracking Evidence for Disrupted Event Segmentation during Continuous Memory Encoding Due to Short Video Watching

Hongxiao Li [1,2], Jiashen Li [1,2], Xin Hao [1,2*], Wei Liu [1,2*]

1. Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education, Wuhan, China

2. Key Laboratory of Human Development and Mental Health of Hubei Province, School of Psychology, Central China Normal University, Wuhan, China

Correspondence:

Dr. Xin Hao, School of Psychology, Central China Normal University

Address: the 8th floor, Nanhu Complex Building, No 152 Luoyu Road, Wuhan.

Postal Code: 430079

E-mail: psyhaoxin@ccnu.edu.cn

Dr. Wei Liu, School of Psychology, Central China Normal University

Address: the 8th floor, Nanhu Complex Building, No 152 Luoyu Road, Wuhan.

Postal Code: 430079

E-mail: weiliu1991@ccnu.edu.cn

Abstract

The widespread use of short-video platforms (e.g., TikTok), with over one billion monthly active users worldwide in 2023, necessitates an examination of their effects on cognitive functions. We hypothesized that because of the frequent, user-driven shifts in content during short-video watching, it mainly impairs event segmentation—a critical cognitive process in continuous perception and memory formation. To investigate this, we integrated memory encoding-retrieval tasks, eye-tracking technology, self-report questionnaires, and advanced machine learning analyses of eye movements. Our experiments (Study1 N=113, Study2 N=60), which varied the types of videos (random vs. personalized) and memory stimuli (continuous movies vs. static images), demonstrated that watching randomly selected short videos, particularly when integrated into daily viewing patterns, detrimentally affects memory encoding, especially in the initial and middle phases of continuous memory encoding. However, this impairment was not evident in trial-based image encoding tasks, suggesting a selective disruption in continuous memory processes. Intersubject correlation (ISC) analysis of eye movements revealed decreased synchronization at event boundaries during movie watching after exposure to random short videos. Furthermore, Hidden Markov Model (HMM) analyses indicated that participants were more likely to experience fragmented event segmentation during continuous memory encoding following exposure to random short videos. These findings underscore the potentially harmful impact of short-video watching on event segmentation and memory formation, emphasizing the role of recommendation algorithms in modulating human cognition.

Keywords: Short Video; Event Segmentation; Eye-tracking; Intersubject Correlation; Hidden Markov Model

Introduction

The exponential growth of short video platforms—internationally known as TikTok and its Chinese counterpart, Douyin—has established them as dominant forces in social media, with 1.5 billion monthly active users as of 2023 and is expected to reach two billion by the end of 2024 (Retrieved from https://www.businessofapps.com/data/tik-tok-statistics/ on January 23, 2024). In China, the origin country of the platforms' parent company, ByteDance, accounts exceeding 1 billion, with users averaging over 2.5 hours of activity per day. There are two major differences between short videos and other types of video-based entertainment such as Netflix and YouTube. The first difference is the personalized recommendation algorithm, which is considered the primary cause of potential addictive behavior. The powerful recommendation system of TikTok/Douyin can accurately predict user interests based on their browsing history and tagged video classificationsn[1,2]**.** Using functional magnetic resonance imaging (fMRI) in healthy adults, Su and colleagues found that viewing personalized videos activates the default mode network (DMN) and modulates functional connectivity between large-scale brain networks[3,4]. The second distinction pertains to the video length: usually under 15 seconds and seldom exceeds one minute. Users' ability to swiftly switch videos based on preference results in even briefer video engagement. This leads to fragmented cognitive patterns, which markedly differ from the sustained cognitive processing required by daily life activities, such as watching a movie or attending a lecture. Given the platforms' widespread use, potential for addiction, and unique cognitive patterns, empirical studies are needed to investigate their consequences for cognitive functions.

Previous research on the effects of short video exposure on human behavior has primarily relied on self-report studies[5]. These have indicated that, like reports on smartphone and social media usage, TikTok exposure correlates positively with symptoms of depression and anxiety[6–9]. However, self-reported measures in the context of short video watching are constrained by potential memory biases and tendencies to respond in socially desirable ways. For instance, individuals may experience subjective time distortions when estimating the duration of their online activities[10]—this issue could arise similarly for estimates of time spent on short video platforms. An alternative research approach involves acutely exposing participants to short videos in a controlled laboratory setting and subsequently observing the carry-over effects on their behaviors. Recent laboratory studies have demonstrated that short video exposure adversely affects cognitive functions, including such as time perception[11], women's body satisfaction[12], prospective memory[13], attention[14]. Hence, we advocate for an integrative research methodology that combines self-reported surveys with objective assessments of cognitive functions and concurrent neural and/or biological data collection, to elucidate the potential underlying mechanisms mediating the relationship between short video watching and cognitive alterations.

In this study, we examined the impact of short video exposure on human episodic memory encoding, particularly anticipating disruptions in cognitive processing due to user-controlled video duration. We investigated whether such exposure impairs event segmentation-a spontaneous cognitive processes by which we divide continuous streams of information into discrete units[15–17]. Previous studies used continuous stimuli, such as commercial movies[18–20], sports footage[21], and auditory narratives[22], has examined this

cognitive process and its link with successful memory formation. The findings consistently indicate that optimal event segmentation during continuous information processing is associated with enhanced memory performance in subsequent retrieval tasks[23–26]. We hypothesize that the abrupt, user-initiated transitions characteristic of short video watching could disrupt event segmentation, a crucial mechanism in the formation of continuous memories. To test this hypothesis, we applied both continuous (i.e., Study1) and picture-based (i.e., Study2) memory encoding-retrieval paradigms. If short video exposure selectively impairs memory retrieval for continuous videos but not for static images, it would suggest a negative and selective impact on event segmentation. However, since memory performance is only an indirect indicator of event segmentation, a direct measurement method is needed to measure event segmentation during naturalistic memory encoding.

To this end, we adopted an eye-tracking technique as a novel alternative to neural methods like fMRI[27–32] and EEG[25] for measuring event segmentation during continuous memory formation and examining its potential alteration by preceding short video exposure. Leveraging the high temporal resolution and cost-efficiency of eye tracking[33,34] and building upon Clewett et al.'s findings that pupil dilation reflects changes in event structure elicited by static visual and auditory stimuli[35], we applied this method to identify signatures of event segmentation in continuous memory encoding. We tracked pupil size and eye movement patterns, integrating advanced computational approaches (i.e., *Hidden Markov Models* (HMM)[36] and *inter-subject correlation analysis* (ISC)[37]) —typically used in neuroimaging—into our eye-tracking data analysis. By applying the HMM—a data-driven event segmentation technique—to eye-tracking data, which includes pupil size and the coordinates

of fixation points over time, we determined the model-generated optimal event number, distinctive eye-tracking patterns specific to each event, and the precise timing of event transitions. We further employed ISC to assess the concordance of eye movements across participants during event processing, with prior studies indicating that higher intra-event eye movement synchronicity, as opposed to inter-event, is associated with enhanced memory retrieval performance[38].

In summary, we conducted two studies (Study 1, N=113; Study 2, N=60) to assess the potential negative effects of acute short-video exposure on memory encoding in healthy adults (**Figure 1**). In Study 1, we allocated participants into four groups to monitor their eye movements with an eye-tracking device during a continuous memory encoding task, involving approximately 20 minutes of the "*BBC-Sherlock*" episode. Subsequently, participants were instructed to freely recall the movie's details. We examined the impact of short-video exposure on memory encoding by varying the tasks preceding the encoding phase, thus probing the carry-over effects of such exposure on subsequent memory formation. Two groups initially engaged with 25 minutes of short videos, differentiated by the video selection method: one group viewed randomly generated videos from an completely new account (Short-Random group, Short-R), while the other group watched personalized recommendations from their own accounts (Short-Personalized group, Short-P). The other two groups viewed a 25-minute continuous video clip before the memory encoding task, differing in content: one (Long group) watched an unrelated documentary ("*BBC-Ocean*"), and the other (Schema group) watched a "*BBC-Sherlock*" segment related to the encoding material, thereby providing narrative context and serving as a manipulation check to confirm

the effects of experimental conditions on memory performance. This four-group design allowed us to evaluate the carry-over effects of video exposure, both short and long, on memory encoding and to dissect the role of personalized recommendation algorithms, a central characteristic of short-video platforms. We found both behavioral and eye-tracking evidence that short-video exposure could affect event segmentation and continuous memory encoding. Subsequently, Study 2 questioned the specificity of this effect to continuous memory encoding: two groups (Short vs. Long) performed trial-based image encoding tasks, preceded by 15 minutes of video viewing—short or long. Our findings indicated no significant impact of short-video exposure on various measures of memory performance, implying that the influence of short-video exposure may be confined to continuous memory encoding contexts.

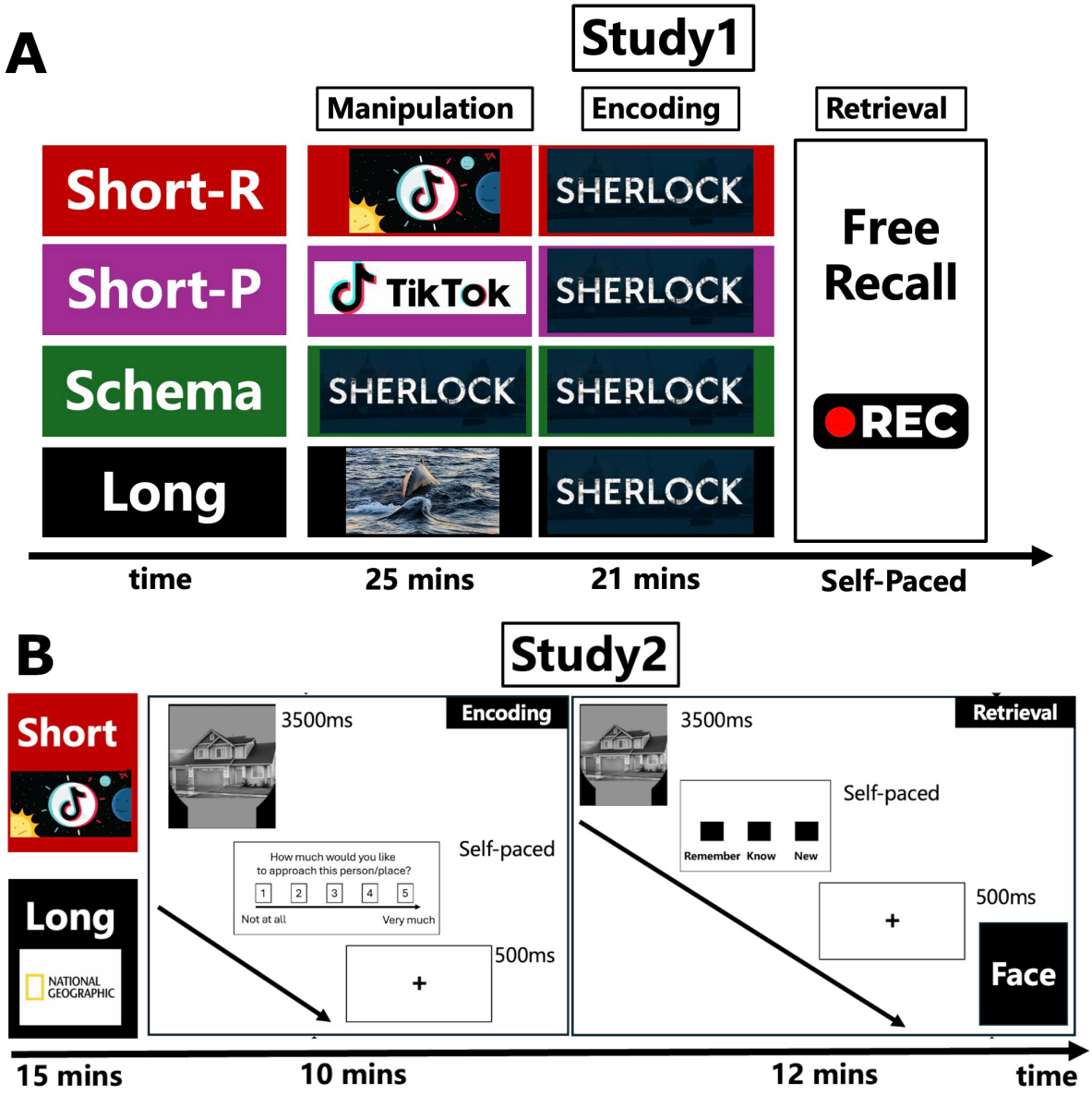

Figure 1: Experimental Design (A) Study 1 Outline. In the manipulation phase, participants were exposed to one of four conditions: randomly selected short videos (Short-R group), personalized short videos (Short-P group), a unrelated segment from the "BBC-Ocean" documentary (Long group), or a related segment from a Sherlock movie (Schema group). During the encoding phase, all participants viewed the same segment of the Sherlock movie while their eye movements were monitored, and memory was assessed through a subsequent free recall task. **(B) Study 2 Outline.** In the manipulation phase, participants watched either of short videos (Short group) or a documentary from National Geographic (Long group). During the encoding phase, participants rated the approachability of images (faces or houses) in a self-paced task, following the presentation of each image. In the retrieval phase, memory recognition was evaluated using a Remember/Know (RK) paradigm.

2. Methods

2.1 Participants

Study1 Naturalistic memory encoding

In Study1, 113 right-handed, healthy undergraduates (73 female, 40 males; average age: 20.42 years, standard deviation: 2.04) were recruited. All of them completed all experimental procedures and were included in the final data analysis. All had normal or corrected vision, no hearing deficits, Beck Depression Inventory scores under 11—signifying no present depression—and no known neurological or psychiatric conditions. They also reported not having seen the movie segment "Sherlock Season 1 Episode 1: A Study in Pink" prior to the study. Recruitment was via advertisements, with compensation provided for participation. Each gave written informed consent ahead of the experiments. These were randomly distributed into four groups to receive various experimental manipulations (see the experimental design section for details). The demographics of the groups were: Short-R (28 participants; 11 male, average age = 20.32, SD = 2.08), Short-P (30 participants; 8 male, average age = 20.58, SD = 1.99), Long (28 participants; 9 male, average age = 20.18, SD = 2.11), and Schema (27 participants; 12 male, average age = 20.77, SD = 1.96). No significant differences in age ($F=1.145$, $p=0.33$) or gender distribution ($\chi^2=2.28$, $p=0.51$) were observed between groups. All subjects watched the same movie clip and were monitored using the same eye-tracking equipment. The study protocols were approved by the University Institutional Review Board for Human Subjects.

Study2 Picture-based memory encoding

In Study 2, we recruited 60 right-handed, healthy undergraduate students (52 females, 8 males; mean age = 21.7, SD = 1.7), who completed all experimental procedures and were included in the final data analysis. Participants had normal or corrected-to-normal vision, no auditory impairments, Beck Depression Inventory scores below 11—indicating an absence of current depressive symptoms—and no reported neurological or psychiatric conditions. All were required to have previously used Douyin (i.e., the Chinese version of TikTok) and to have this application installed on their smartphones. During the manipulation phase, participants were instructed to use their phones to play videos (short or long) within the app. Participants were divided into two age- and gender-matched groups: a Short group (26 females, 4 males; mean age = 21.36, SD = 1.65) and a Long group (26 females, 4 males; mean age = 22.2, SD = 1.71). The study protocols received approval from the University's Institutional Review Board for Human Subjects.

Sample Size Justification

Given no previous study investigated the effect of short video watching on memory encoding, estimating the precise effect size before collecting data was not possible. The sample size was guided by earlier research on event boundaries that typically engaged 20-30 subjects[18,23,29], often supplemented by other techniques like EEG and fMRI. A relevant previous study, employing pupil dilation measurements to examine event segmentation, involved 34 participants and posited a need for at least 25 subjects to identify a substantial effect size (d=0.8) under conditions of an α of 0.05 and 80% power[35]. For our Study1, we targeted a minimum of 25 evaluable subjects per experimental condition, ensuring complete datasets of memory and eye movement metrics. All primary and secondary analyses

exceeded this participant count, with the minimal sample in the schema group with 27 participants included in the analyses. Study2, we achieved the final sample of 30 participants in our final analyses.

2.2 Materials and Design

2.2.1 Study1 Naturalistic memory encoding

Experimental design

The study was structured into three phases: the manipulation phase, the primary video viewing phase, and the free recall phase. Initially, participants were informed that they would engage with approximately 45 minutes of diverse video content, with the objective of retaining information for subsequent memory assessment. During the free recall phase, participants were instructed to recall and describe only the videos shown during the primary video viewing phase. Instructions emphasized remembering content from both the manipulation and primary video viewing phases to maintain focus throughout the experiment, thereby ensuring that the manipulation phase was not strategically disregarded.

In the manipulation phase, aimed at influencing the event segmentation during subsequent video viewing, participants in the Short-Random (R) group were exposed to 400 short TikTok clips across themes such as humans, animals, objects, and scenes, each theme contributing 100 clips of 1 to 5 seconds, curated from a neutral account to avoid bias from personalized algorithms. Conversely, the Short-Personalized (P) group viewed clips tailored by recommendation algorithms from their profiles, ensuring the total viewing time matched the 20 minutes allocated to other groups. The Long group engaged with a 20-minute segment

from the "BBC Ocean" documentary, while the Schema group viewed a 20-minute portion from "BBC Sherlock" that introduced elements of the primary video content.

During the primary viewing phase, participants from all groups watched a 25-minute segment from "BBC Sherlock," selected due to its prevalent use in earlier event segmentation studies[19,20,23,25], which indicated its capability to evoke distinct, event-specific neural representations.

In the free recall phase, participants recounted in detail the content from the primary phase, emphasizing the depth of recall. The manipulation phase content, while not directly evaluated, was included in assessments to maintain participant focus throughout the experiment. The video presentations for both phases were controlled using Experiment Builder software (SR Research), and initial responses were captured via smartphones and later transcribed for detailed memory analysis.

Human-labelled event boundaries and memory measures

Utilizing "Sherlock Season 1 Episode 1 'A Study in Pink'," previously employed in segmentation studies[19,20,23], allowed us to adopt established event boundaries without modification, specifically those identified by Chen [20], covering 20 events within a 25-minute framework. Audio recordings from the free recall task were transcribed and assessed by two independent raters. The evaluation of each participant's recall was based on three metrics: (1) the remember score, reflecting the number of events recalled; (2) the raw detail score, representing the average number of details recalled per event; and (3) the corrected detail score, calculated as the ratio of recalled details to the total details per event, averaged across all events. Additionally, we investigated the serial position effect by calculating the recall

probability for each event, considering its position in the original movie's storyline. Initial evaluations by two raters were independent; disagreements were discussed until a unified consensus was reached, which guided further analysis.

2.2.2 Study2 Picture-based memory encoding

Experimental design

In Study 2, we organized the research into three distinct phases: the manipulation phase, the memory encoding phase, and the retrieval phase. During the manipulation phase, participants used the Douyin app to watch 15 minutes of either short videos or a long video, specifically the documentary "The Enchanting Creatures of the Motuo Forest - Inside the World's Largest Canyon" from National Geographic. Participants were instructed not to use the comment or like functions of the app and to disable notifications from other apps to avoid interruptions. The memory encoding and retrieval phases involved assessing memory with both social and non-social stimuli, using a total of 180 images. Ninety images depicted human faces and were sourced from the Chinese Academy of Sciences' Sinicized Face Emotion Image System. From a pool of 600 grayscale images, we selected 90 faces (45 males, 45 females) displaying positive, neutral, and negative emotions—30 faces per emotion category. Sixty of these faces were employed for memory encoding, and 30 served as distractors during the recognition task. For non-social stimuli, we curated 90 images (30 each of houses, castles, and offices) from pixabay.com. These were similarly divided, with 60 used for encoding and 30 for interference, all presented within an oval mask against a black backdrop.

Encoding

Participants underwent an initial encoding session where they were tasked with rating the approachability of 60 faces and 60 non-social images. This activity was structured into two segments separated by a 20-second interval, with each image displayed for 3500 ms followed by a self-paced approachability rating on a scale from 1 (not approachable) to 5 (very approachable). This task masked the true intent of memory encoding, requiring detailed attention to each stimulus. An interstimulus interval of 500 ms was maintained.

Retrieval

Memory retention was unexpectedly assessed using the remember/know (RK) paradigm[39]. Participants were presented with the previously encoded images alongside 30 new faces and 30 new non-social images across two blocks, totaling 180 stimuli. They were asked to classify each image as remembered, known, or new. Each image was shown for 3500 ms, followed by a self-paced response, with trials separated by a 500 ms interval. A brief practice session using unrelated cartoon faces and buildings preceded the testing to clarify the distinction between 'remember' and 'know' responses.

Memory measures

Overall recognition memory accuracy was quantified by the net difference between the overall hit rate and false alarm rate [i.e., Overall recognition memory = hit rate (remember + know) − false alarm rate (remember + know)]. Separate metrics for recollection and familiarity were calculated: recollection was defined as the net hit rate for remembered items minus the false alarm rate for those classified as remembered [i.e., hit rate (remember) − false alarm rate (remember)], and familiarity was calculated as the net rate of known responses, adjusted for false alarms [i.e., (hit rate know / (1 - hit rate remember)) - (false alarm rate

know / (1 - false alarm rate remember))]. These measures were used to evaluating the impact of experimental variables on long-term memory retention. Given that the images originated from two principal categories—faces and houses—we first computed the Overall Recognition, Recollection, and Familiarity scores separately for each category. This analysis aimed to determine if the act of watching short videos exerted a selective impact on certain categories. Subsequently, we computed all memory metrics without regard to image category. In summary, we derived nine memory metrics for comparative analyses between groups of participants who viewed either short or long videos prior to the encoding phase, and for correlational studies with questionnaire-based assessments of daily short video usage.

2.3 Questionnaire measure of short video use

Participants from both Study 1 and Study 2 completed a 20-item Short Video Use Scale at the conclusion of their respective experiments. This data collection occurred post memory encoding and retrieval phases to prevent participants from inferring the study's objectives, which could potentially influence their memory performance. While prior research has focused on YouTube addiction, studies on "TikTok Use" (or "Douyin Use") are scarce. The scale employed for assessing YouTube addiction is a modified version of the Internet Addiction Test (IAT)[40]. In this study, we used the Chinese version of Young's 20-item IAT, replacing "the Internet" with "Short Videos". We defined "Short Videos" at the beginning of the questionnaire as encompassing any short video platform, including TikTok (International Version), Douyin (the Chinese version of TikTok), KuaiShou, and the Tencent Short Video Platform. The questionnaire uses a Likert-type scale ranging from 1 (rarely) to 5 (always),

where higher scores indicate greater severity of problems caused by short video usage. This instrument has proven effective in prior studies examining problematic TikTok use[3,4].

2.4 Eye-tracking data acquisition

All participants, regardless of their group allocation, were accommodated in a uniform environment where they were exposed to stimuli on the same model of monitor, with their eye movements being uniformly recorded. The experimental setting was modified to mitigate any influence from external light by using curtains to block out natural sunlight, ensuring that indoor lighting maintained consistent luminance throughout the day. The precision in monitoring eye movements was achieved through a desk-mounted EyeLink 1000 system (SR Research Ltd., Mississauga, Ontario, Canada), using a 35-mm lens at a high sampling rate of 1000 Hz. A head and chin rest provided by SR Research minimized head movements. Before the onset of the experiment, participants underwent a briefing on the eye movement calibration and the eye tracker setup, which included adjustments to the camera for optimal eyelash visibility and modifications to the infrared sensitivity threshold. Calibration involved directing participants to sequentially focus on the four extremities of the monitor's display and was initiated by asking them to concentrate on nine randomly ordered black dots across the screen, placed at predetermined locations. The calibration's extent adapted based on the stimulus task's screen area, using a 9-point grid. Each point had to be focused upon by the participants until its disappearance, without predicting its trajectory. Post-calibration, the accuracy was verified by reevaluating participants' ability to fixate on the same nine points, this time assessing the consistency between recorded and actual positions to quantify visual

discrepancies. The analysis software provided metrics on the average and maximal visual inaccuracies for each fixation. The experimental procedures started once the calibration was deemed satisfactory, with participants instructed to maintain silence during the presentation of stimuli.

2.5 Eye-tracking data preprocessing and analysis

Preprocessing

We used Matlab to preprocess the eye-tracking data. Given the predetermined duration of the video content, we standardized the length of the data sequences while preserving their integrity. We also corrected for artifacts due to blinks by replacing these instances, culminating in a complete dataset for each participant that included coordinates of screen fixations (X, Y) and pupil diameters.

Interject-correlation analyses (ISC) of eye movements

To investigate how event boundaries influence the synchronization of eye movement patterns and the effects of prior exposure to short videos, we calculated the inter-subject correlation (ISC) of eye movements during specific temporal segments around event boundaries (pre-event, during-event, and post-event) for each participant. The analysis was conducted by calculating Pearson's correlation coefficients for vertical gaze positions between each subject and all other viewers, averaging these values to derive an ISC for each individual, and repeating across all participants to establish individual ISC scores. This procedure was also applied to pupil size and horizontal gaze data. A composite ISC was then calculated as the mean of these values: $ISC = (ISC_x + ISC_y + ISC_{pupil}) / 3$. ISC assessments for each event boundary were conducted under three distinct scenarios: 10 seconds before the

boundary, 5 seconds both before and after the boundary, and 10 seconds following the boundary. We analyzed ISC variability for each participant across these conditions using repeated-measures ANOVA and conducted between-group comparisons using ANOVA to determine if ISC scores varied among different groups.

Eye-tracking-based Hidden Markov Models (HMM)

We adapted Hidden Markov Models, previously used in fMRI[19] and EEG[25] studies, to analyze eye-tracking data. This method processed three time series: X and Y coordinates of screen fixations and pupil size. Our data-driven event segmentation model could discern specific eye movement patterns and the structural dynamics of events, including the identification of event boundaries. Initially, we explored individual differences in event structures without presetting the number of events (K). We employed the HMM across K values ranging from 15 to 25, allowing the model to select the optimal K based on eye movement data. Previous fMRI studies, using the same movie clip, established a ground truth of 20 events[20]. Consequently, K values below 20 might indicate overlooked responses near annotated boundaries, whereas values above 20 could suggest an over-segmentation of event cognition, with additional boundaries emerging from the eye movement patterns. We further examined the correlation between individual K values and memory performance to affirm the relationship between event segmentation and event memory. As hypothesized, K values exceeding the established ground truth correlated with poorer memory performance. Subsequently, we set the model to a predefined k of 20 to elucidate latent states and the structure of observed events. Using a leave-one-out approach, we initially applied this model with a fixed k to ascertain optimal event boundaries for each participant. We evaluated the

model's fit by comparing the HMM-determined boundaries against those from a null model using within-event similarity metrics. This comparison was statistically analyzed against a distribution generated from 1000 randomized event sequences to assess the significance of the model's fit. A significant p-value indicated a robust alignment between the observed eye movement patterns and the HMM-derived event boundaries.

2.6 Data and code availability

Processed behavioral and eye-tracking data have been made available on the OSF platform (https://osf.io/n7k45/?view_only=673d89b23720410799b5daec1f5d511e). Raw data can be accessed by request to the corresponding authors. Due to copyright constraints, original video materials from the study are not available. The analytical scripts and models used in the study are available in the same repository where the data is hosted. Data analyses were performed using MATLAB and Python, with the HMM developed using the BrainIAK toolbox [41] and ISC inspired by Pauline's study [22].

3. Results

3.1 Exposure to randomly selected, but not personalized short videos, and its interaction with daily viewing habits on event memory impairment

In study1, we assessed participants' memory performance by measuring the number of events they could recall and their ability to recall details within these events. Furthermore, we calculated corrected detail scores to account for variations in the number of details per event. These memory metrics were then correlated with daily viewing habits assessed through TikTok scores, a questionnaire-based measure (**Figure 2, left panels**). Our analysis revealed that frequent daily use of short videos was marginally and negatively associated with the number of recalled events ($r = -0.17$, $p = 0.07$) but exhibited no correlation with detail recall (raw scores: $p = 0.12$; adjusted scores: $p = 0.41$). Subsequent group comparisons on these memory metrics identified significant differences across three measures (**Figure 2, middle panels**): recalled events ($F = 5.83$, $p < 0.001$, $\eta^2 = 0.138$), detail scores ($F = 3.28$, $p = 0.024$, $\eta^2 = 0.08$), and corrected detail scores ($F = 3.00$, $p = 0.03$, $\eta^2 = 0.07$). Notably, the Schema group exhibited enhanced memory performance, recalling more events (Long: $t = 3.75$, $p < 0.001$, Cohen's d = 1.01; Short-R: $t = 3.03$, $p = 0.004$, Cohen's d = 0.81; Short-P: $t = 4.21$, $p < 0.001$, Cohen's d = 1.11) and details—regardless of measurement type (Detail scores: Short-R: $t = 2.51$, $p = 0.015$, Cohen's d = 0.67; Short-P: $t = 2.66$, $p = 0.01$, Cohen's d = 0.70; and corrected scores: Long: $t = 1.86$, $p = 0.06$, Cohen's d = 0.50; Short-R: $t = 2.31$, $p = 0.02$, Cohen's d = 0.62; Short-P: $t = 2.47$, $p = 0.01$, Cohen's d = 0.65). No significant differences were observed among the other three groups (all $p > 0.10$). Lastly, we explored group-specific correlations between TikTok scores and memory metrics (**Figure 2, right panels**). A

strong negative correlation was noted in the Short-R group between TikTok scores and the number of recalled events ($r = -0.42$, $p = 0.02$), a trend absent in other groups (Short-P: $r = 0.27$, $p = 0.13$; Long: $r = -0.04$, $p = 0.81$; Schema: $r = -0.04$, $p = 0.81$). No other group-specific correlations were observed (all $p > 0.10$).

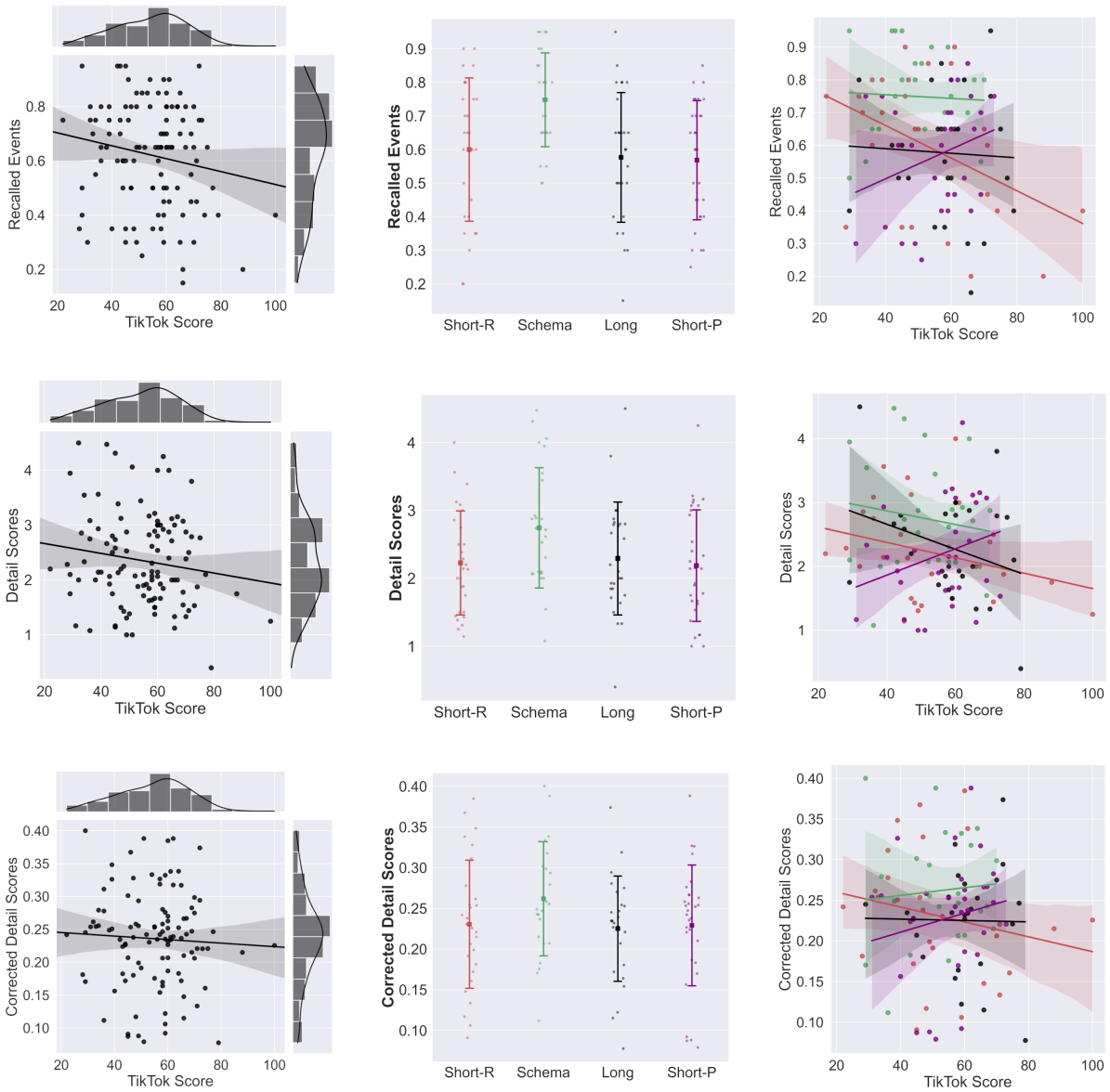

Figure 2: Impact of Acute Short Video Viewing and its Interaction with Daily Viewing Habits on Subsequent Event Memory. Left Panels: Correlation analysis reveals that higher TikTok scores, a self-reported measure reflecting the habitual viewing of short videos, are marginally significantly and negatively associated with the number of recalled events (r = -0.17, p = 0.07), but show no correlation with the detail recall (raw scores: p = 0.12; adjusted scores: p = 0.41). **Middle Panels:** Memory performance comparisons across four groups indicate that the Schema group exhibited advantages in three memory metrics. Neither the Short-R nor the Short-P groups demonstrated inferior memory performance relative to the Long group across the measures. **Right Panels:** Within the Short-R group, a strong negative correlation was observed between TikTok scores and the number of recalled events (r = -0.42, p = 0.02), a trend not evident in other groups.

3.2 Impact of short video watching on memory encoding during the initial and middle phases of movie watching

We used the free recall method to assess the position-specific impact of short video viewing on memory encoding. Without memory cues from experimenters or on-screen prompts, events were typically recalled in their correct temporal sequence (**Figure 3A**). Upon analyzing the serial positioning effect across all participants, we observed variations in recall probability based on series position (**Figure 3B**). We classified the positions into five segments, each comprising four consecutive positions. For instance, the first segment encompasses positions 1 to 4, whereas the second segment includes positions 5 to 8. Both the initial and final segments exhibited elevated memory recall probabilities ($M_{1\text{-}4}$=0.77, $Std_{1\text{-}4}$=0.41; $M_{17\text{-}20}$=0.66, std=0.47), while the middle segments demonstrated lower recall probabilities ($M_{9\text{-}12}$=0.39, $Std_{9\text{-}12}$=0.48; $M_{13\text{-}16}$=0.46, $std_{13\text{-}16}$=0.49). Subsequently, we conducted a group-specific analysis, identifying positions where the Short-R or Short-P groups demonstrated a lower recall probability compared to the Long group (**Figure 3C**). No evidence was found to support a disruptive effect of short video viewing (all p>0.1). We then merged the Short-R and Short-P groups into a combined SHORT group, and the Long and Schema groups into a LONG group and reevaluated the between-group differences (**Figure 4D**). We observed a significantly lower recall probability at positions 1 (χ^2=7.86, p_1=0.004) and position 6 (χ^2=7.32, p_6=0.006), and within the continuous time window from positions 10 to 13 (χ^2=4.41, p_{10}=0.03; χ^2=3.2, p_{11}=0.07; χ^2=5.31, p_{12}=0.02; χ^2=4.67, p_{13}=0.03) in the SHORT group compared to the LONG group (**Figure 4E**), indicating a position-selective effect of short video watching on memory encoding.

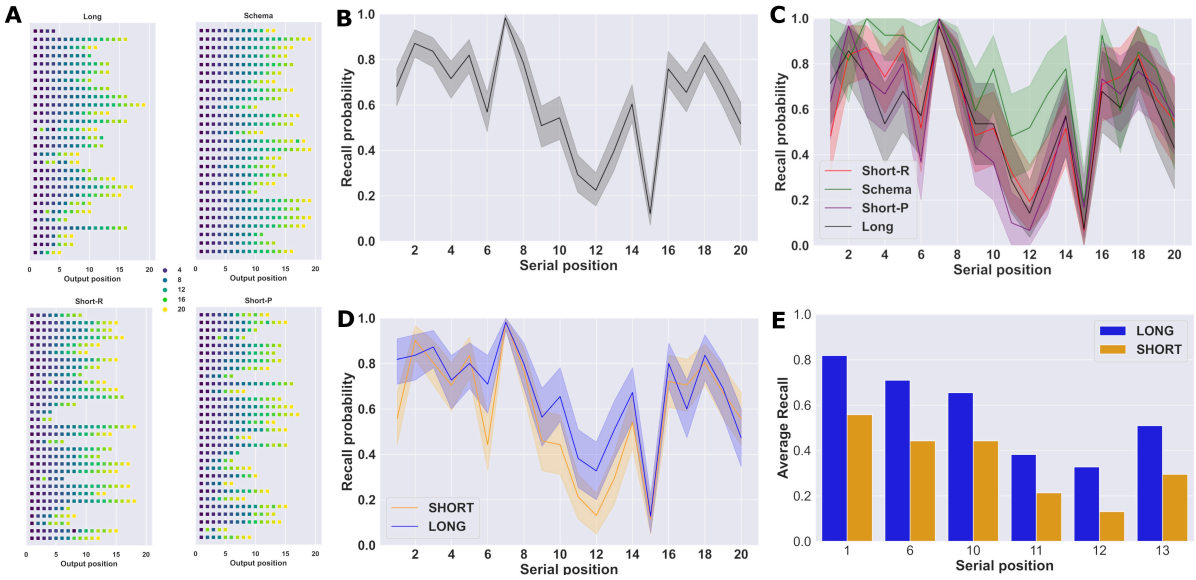

Figure 3: Selective Disruptive Effects of Short Video Watching on Event Memory. **(A)** Recall by each individual, color-coded by the event's serial position in the movie. Events toward the story's end are in yellow, while those at the beginning are in purple. Each row represents a participant, and each panel, an experimental group. **(B)** The average recall probability as a function of serial position illustrates the serial positioning effect: events at the beginning and end of the film are more memorable than those in the middle. **(C)** Comparison of recall probabilities between the four groups **(D)** Comparison of recall probabilities between the SHORT (Short-R and Short-P) and LONG (Long and Schema) groups. **(E)** Serial positions showing significantly lower recall probability in the SHORT group compared to the LONG group.

3.3 Disrupted event segmentation after short video watching indexed by reduced inter-subject correlation at boundaries

We employed eye-tracking-based inter-subject correlation (ISC) analysis to investigate the impact of acute short video viewing on subsequent event segmentation during continuous video watching. This ISC calculation incorporated both vertical and horizontal gaze coordinates, alongside pupil sizes, for each participant. ISC was examined under three conditions (**Figure 4A**): 10 seconds preceding the event boundary (pre-boundary condition), 5 seconds before and after the boundary (boundary condition), and 10 seconds after the boundary (post-boundary condition).

A between-group ANOVA revealed significant differences in boundary ISC among four groups ($F=4.14$, $p=0.008$, $\eta^2=0.11$; **Figure 4B**). Post-hoc analyses indicated that the Short-R group exhibited lower boundary ISC compared to the Long group ($t=1.94$, $p=0.05$, Cohen's $d=0.28$), Schema group ($t=2.28$, $p=0.02$, Cohen's $d=0.28$), and Short-P group ($t=2.88$, $p=0.006$, Cohen's $d=0.79$). Additionally, within-group dynamic changes from pre-boundary to boundary were analyzed (**Figure 4C**). The Long, Schema, and Short-P groups demonstrated an increase in ISC from pre-boundary to boundary (Long: $t=8.29$, $p<0.001$, Cohen's $d=1.65$; Schema: $t=5.26$, $p<0.001$, Cohen's $d=1.03$; Short-P: $t=13.52$, $p<0.001$, Cohen's $d=2.82$), whereas the Short-R group showed a decrease ($t=-8.13$, $p<0.001$, Cohen's $d=-1.48$).

We characterized the increase in ISC from pre-boundary to boundary as a "boundary response" and assessed its association with memory performance. We observed significant positive correlations between the boundary response and the detail score, which measures

how accurately participants could recall movie details (r=0.24, *p*=0.01). This finding suggests that the boundary response, absent in the Short-R group, plays a crucial role in linking event segmentation to successful memory encoding. Additionally, the number of events participants recalled (i.e., remember score) demonstrated positive correlations with pre-boundary ISC (r= 0.27, *p*=0.006), boundary ISC (r=0.25, *p*= 0.009), and post-boundary ISC (r = 0.25, *p* = 0.008). We depicted all possible correlations between ISC values and memory performance within the correlation matrix (**Figure 4D**).

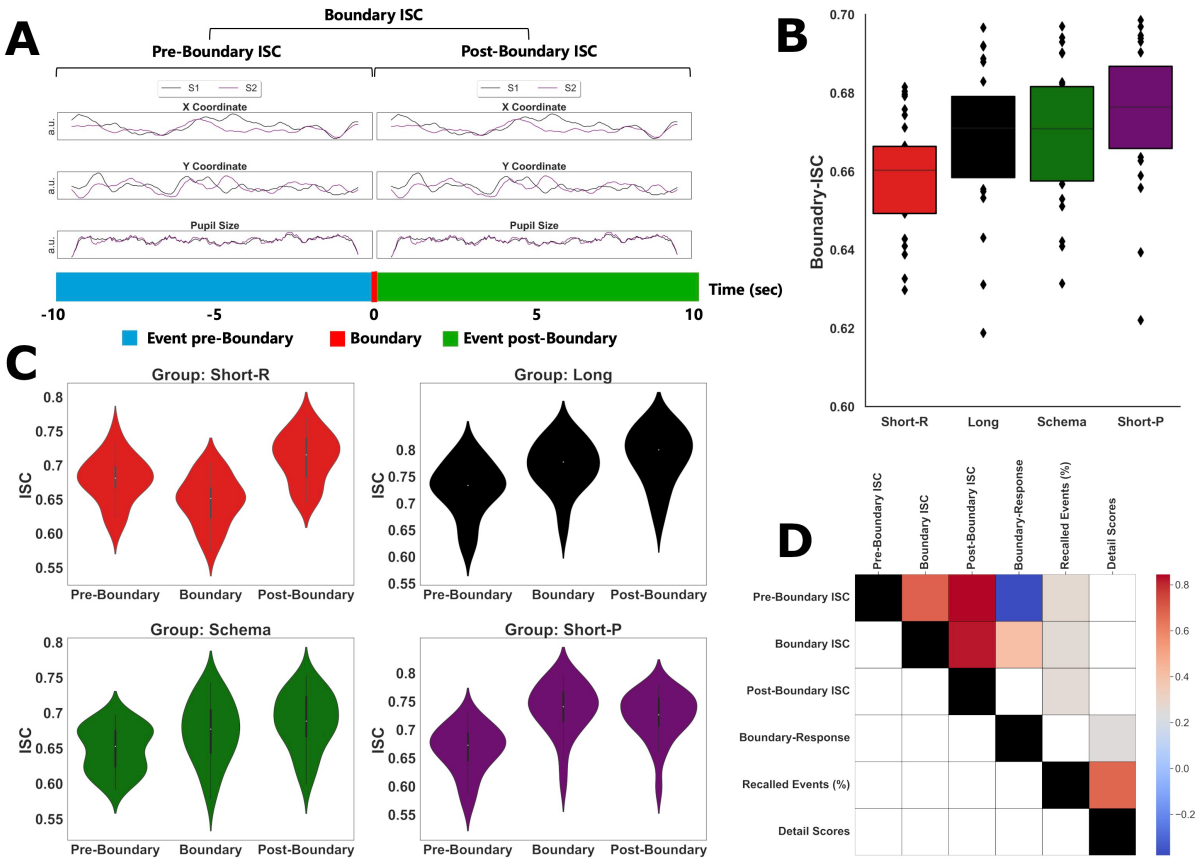

Figure 4 Intersubject Correlation (ISC) Analysis and Event Segmentation During Memory Encoding. (A) Displays synchronized eye movements—positions (X and Y coordinates) and pupil sizes—of participants S1 and S2 while they observed two distinct events separated by an event boundary. ISC is quantified as the mean correlation between both vertical and horizontal gaze positions and pupil sizes across participants. Three ISC metrics were analyzed: pre-boundary ISC, calculated during the final 10 seconds before the event boundary; boundary ISC, derived from 5 seconds of eye-tracking data at the boundary; and post-boundary ISC, measured during the initial 10 seconds after the event boundary. **(B)** Participants in the Short-R group, who watched randomly selected short videos, had diminished synchronization of eye movements around event boundaries compared to other groups. **(C)** All groups, except the Short-R group, exhibited higher boundary ISC relative to pre-boundary ISC, a phenomenon defined as the Boundary Response. **(D)** Boundary Response during encoding was predictive of subsequent detailed recall (i.e., detail scores), and that all three ISC metrics were predictive of the ability to recall the number of events (i.e., recalled events (%)).

3.4 Disrupted event segmentation after short video watching revealed by HMM

We applied the HMM-based event segmentation model, specifically evaluating optimal K values and p-values, to examine the event segmentation process during naturalistic memory encoding. For each participant, the HMM event segmentation model identified the optimal K value from a candidate range of 15 to 25 and pinpointed the specific locations of each boundary using time series data of pupil size and vertical (y) and horizontal (x) gaze coordinates. Aligning these HMM-generated event boundaries on a time-by-time similarity matrix can capture patterns of within-event similarities (**Figure 5A**).

Firstly, we confirmed the cognitive relevance of the HMM-based optimal K value. Across participants, although the optimal K value was not linked to the number of events recalled (i.e., "remember" score; F=0.23, p=0.87), it associated with the number of details recalled ("detail" score; F=7.36, p<0.001; **Figure 5B**). A linear correlation analysis revealed that the optimal K value negatively predicted the detail score (r=-0.42, p<0.001). Subsequent between-group analyses centered on the optimal K value and p-value, assessing the fit between eye-tracking data and model predictions based on HMM-generated event boundaries. We observed significant differences in optimal K values between groups (F=7.25, p<0.001; **Figure 5C**). Specifically, the Short-R group exhibited significantly higher K values than the long group (t=2.18, p=0.034, Cohen's d=0.28), indicating more fragmented event segmentation during movie watching. Contrarily, K values for the Short-P group were significantly lower than those of the Long group (t=-2.9, p=0.006, Cohen's d=-0.31). Furthermore, we analyzed p-values as indicators of event segmentation. A lower p-value suggested a better fit between HMM predictions on locations of event boundaries and eye-

tracking pattern similarity. Experimental manipulations prior to movie watching influenced p-values (F=4.14, *p*=0.008; **Figure 5D**). No significant differences were found between the two short video groups (Short-R vs Short-P: t=0.71, *p*=0.48, Cohen's d=0.19). Both the Schema and Long groups demonstrated significantly lower p-values than both the Short-P (Schema: t=-2.69, *p*=0.01, Cohen's d=-0.77; Long: t=-2.32, *p*=0.02, Cohen's d=-0.67) and Short-R (Schema: t=-2.62, *p*=0.01, Cohen's d=-0.70; Long: t=-2.13, *p*=0.03, Cohen's d=-0.57) groups.

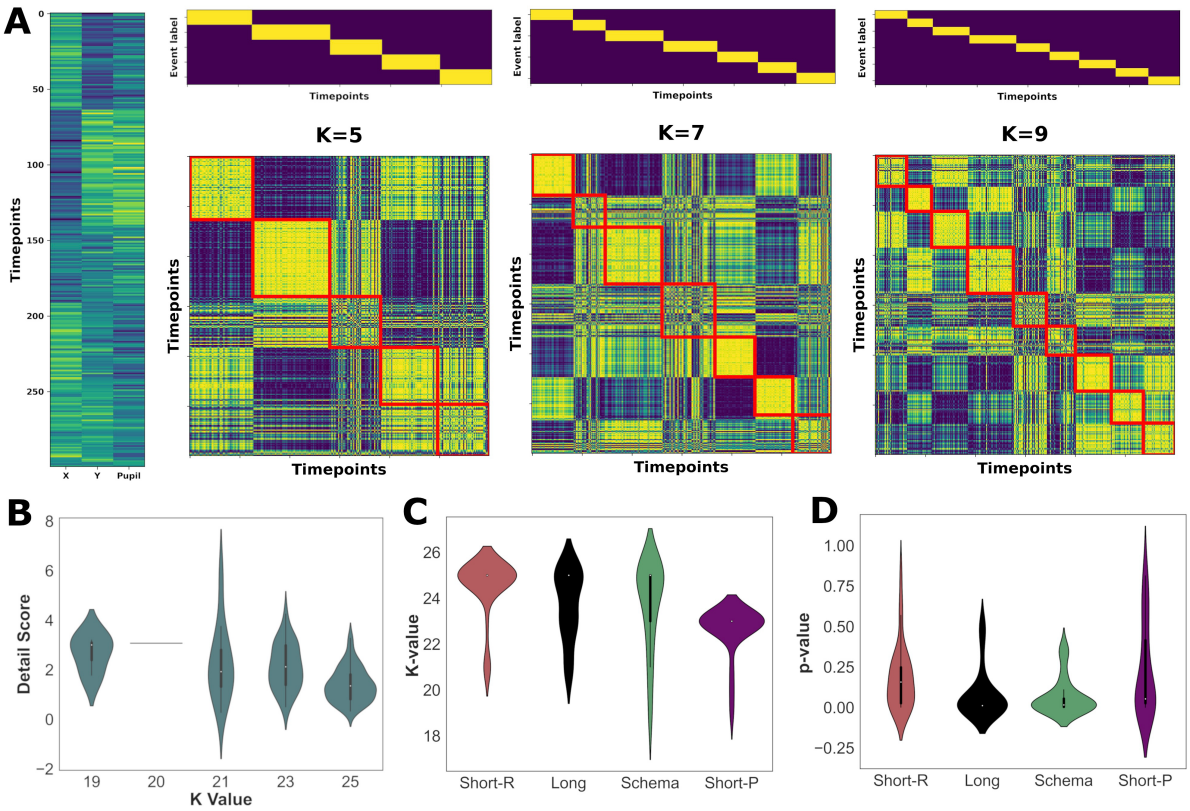

Figure 5: Application of Hidden Markov Models (HMM) in Event Segmentation of Eye Movements During Memory Encoding. (A) HMM's ability to discern individual differences in the event structure from eye-tracking data is highlighted. Inputs to the HMM include time series data of pupil size, vertical (y), and horizontal (x) gaze coordinates (left panel). The HMM model then determines the optimal K value, representing the ideal number of events, from a set of potential values. Examples of three event segmentation models at K values of 5, 7, and 9 are presented. Upper panels display event labels at each time point, while lower panels show a time-by-time correlation matrix of eye-movement patterns. Red boundaries indicate the calculation of within-event pattern similarity. **(B)** A higher K value is associated with poorer recall of movie details among participants. **(C)** This panel compares the average optimal K values across four groups. **(D)** Analysis of individual p-values to determine if event boundaries identified by the HMM result in significantly greater within-event similarity compared to a null model, across four groups.

3.5 No evidence that watching short video disrupts memory encoding of static pictures

We conducted Study 2 with an independent sample of 60 healthy young adults to examine the influence of short video watching on memory encoding and its dependency on the type of stimuli. Our hypothesis was that the picture-based memory encoding task, requiring only brief attention spans and no need for spontaneous event segmentation, would show reduced sensitivity to short video exposure. All participants completed a picture-based encoding-retrieval paradigm using faces and houses as stimuli, along with a self-reported questionnaire assessing daily-life short video usage (i.e., TikTok score). During the manipulation phase, participants were randomly assigned to either watch 15 minutes of short videos or a long video, followed by the encoding phase. The retrieval phase began immediately afterward, and performance was evaluated through separate measures of recognition, familiarity, and recollection for different stimulus types, and collectively for all pictures. No significant differences were found in any of the memory metrics investigated (all $p>0.10$; **Figure 6**, left panels). Furthermore, across all participants, no significant correlations were observed between TikTok scores and memory metrics (all $p>0.10$; **Figure 6**, right panels). Complete statistical results for each group comparison and correlation analysis are available in **Table 1**.

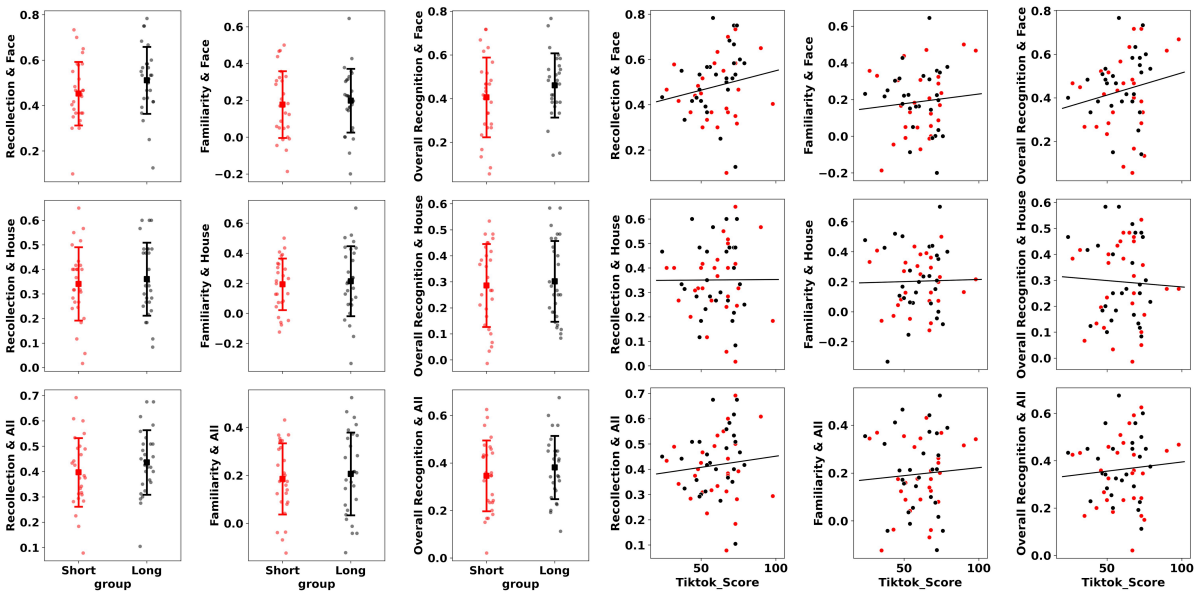

Figure 6 No evidence of impact from acute or daily life short video watching on memory encoding of static pictures. Memory encoding was assessed using a picture-based encoding-retrieval paradigm involving human faces and houses as stimuli. Memory performance was quantitatively assessed through measures of recognition, familiarity, and recollection. The frequency of daily-life short video consumption was quantified via a self-reported questionnaire designed to capture usage intensity (i.e., "TikTok Score"). Left panels: Between-group analysis of memory performance across different stimuli (faces and houses) and memory assessment metrics (recognition, familiarity, and recollection). Right panels: Correlation analysis between TikTok scores and memory performance, encompassing all types of stimuli and memory measures.

Table1 Statistical Analysis of Memory Performance in Study 2

	Group		T-test[1]		Pearson correlation[2]	
Metric[3]	**Short**	**Long**	**t**	***p***	**r**	***p***
Recollection$_{face}$	0.45(0.13)	0.50(0.14)	1.316	0.19	0.18	0.14
Recollection$_{house}$	0.33(0.14)	0.35(0.14)	0.519	0.60	-0.005	0.96
Recollection$_{all}$	0.39(0.13)	0.42(0.12)	1.013	0.31	0.10	0.44
Familarity$_{face}$	0.18(0.17)	0.19(0.18)	0.242	0.80	0.07	0.54
Familarity$_{house}$	0.34(0.79)	0.21(0.22)	-0.818	0.41	0.14	0.25
Familarity$_{all}$	0.26(0.41)	0.20(0.16)	-0.680	0.49	0.15	0.22
OverallRecognition$_{face}$	0.40(0.18)	0.44(0.16)	0.850	0.39	0.19	0.17
OverallRecognition$_{house}$	0.27(0.16)	0.30(0.15)	0.548	0.315	-0.06	0.61
OverallRecognition$_{all}$	0.34(0.15)	0.37(0.13)	0.814	0.419	0.07	0.58

1. Between-group t-tests (t and *p* values) were conducted to explore potential differences in memory metrics between the Short and Long groups.

2. Pearson correlations (r and *p* values) were performed to identify any associations between memory metrics and questionnaire-based measures of daily short video usage (i.e., TikTok score).

3. Overall recognition memory = hit rate (remember + know) − false alarm rate (remember + know). Recollection = hit rate (remember) − false alarm rate (remember). Familiarity = (hit rate know/(1 − hit rate remember)) − (false alarm rate know/(1 − false alarm rate remember)). Subscript indicates that the memory metric was calculated based on pictures of faces, houses, or all pictures used.

4. Discussion

Our findings provide robust evidence that watching randomly selected short videos markedly impairs event segmentation during subsequent continuous memory encoding. This cognitive disruption was demonstrated using advanced eye-tracking analyses, including intersubject correlation (ISC) and Hidden Markov Model (HMM) event segmentation model. Specifically, we observed diminished synchronization of eye movements at event boundaries and increased fragmentation in event segmentation following exposure to random, but not personalized, short videos. Importantly, these adverse effects were confined to continuous memory tasks, such as movie watching, and were not observed in discrete, trial-based tasks like picture encoding. These findings indicate that rapid, user-driven content shifts inherent to short-video platforms specifically disrupt event segmentation, thereby altering memory encoding and retrieval processes. This highlights the significant impact of digital media consumption on key aspects of human cognition.

Our behavioral analyses from Study 1 found that acute exposure to short videos, coupled with habitual daily usage, impairs event memory encoding. This suggests that neither self-reported daily use nor acute exposure to short videos independently influences memory encoding. Although previous studies have examined the cognitive impacts of short video watching[11–14], none have explored its relationship with human episodic memory encoding. A related study demonstrated that using search engines reduces recall rates of the information itself (i.e., Google effect)[42]. With the rise of short video platforms as a dominant social media form, there is growing interest in their effects on memory. Based on these findings, we recommend that future research should combine investigations of habitual short video

consumption with acute exposure in laboratory settings. Using our free recall measure for event memory, we observed a selective impact on the initial and middle phases of continuous memory encoding, typically weakened due to the serial position effect[43]. We interpret these results as indicating that participants found it challenging to shift their attention from short videos to the current movie initially and struggled to maintain focus on encoding information during the middle of the movie. In a parallel Study 2, we found that neither routine daily viewing nor acute exposure to short videos prior to picture encoding affected memory performance. This outcome supports the resilience of picture-based memory encoding to interference from prior short video watching and highlights the advantages of using more ecologically valid stimuli in studies of human cognition [44,45], as the less ecologically valid trial-based paradigm failed to capture the effects of short video watching.

Our eye-tracking analysis employed advanced HMM and ISC methods to examine event segmentation during continuous memory encoding and to identify group differences in this cognitive process. Participants who viewed short videos prior to memory encoding demonstrated disrupted event segmentation, as evidenced by abnormally high numbers of segmented events and decreased synchronization of eye movements near event boundaries. Importantly, these eye-tracking metrics of event segmentation predicted subsequent memory performance, highlighting their cognitive relevance. Both neural markers of event segmentation, derived from HMM and ISC, have demonstrated comparable cognitive relevance[18,23,38] and have been used to investigate developmental changes[29,46], anticipation[27], schema[47] in event segmentation. Our study is the first to apply Event Segmentation Theory (EST)[15–17] to a real-life situation, specifically the cognitive consequences of short video

watching. According to the EST, individuals spontaneously segment their ongoing experiences into discrete events, a process crucial for the long-term storage of event representations. EST posits that event boundaries arise at points of prediction error, leading to the abandonment of the current event model and the initiation of a new one. The hippocampus plays a pivotal role in signaling these prediction errors [48,49]. Our findings suggest that viewing short videos induces internal event segmentation, which is distinct from external event boundaries. These internal boundaries, not driven by prediction errors, negatively impact memory formation.

Our study makes a unique contribution to the literature on the cognitive consequences of short video watching, specifically by directly comparing the effects on continuous versus discrete memory encoding. Our eye-tracking analysis indicates that event segmentation—essential for continuous perception and memory formation—is compromised after exposure to randomly selected videos, whereas personalized videos do not exhibit this impairment. These findings go beyond the general debate on whether short video watching negatively affects memory by pinpointing specific cognitive processes (i.e., event segmentation) that may be impacted. Additionally, we distinguished the distinct effects of random versus personalized short videos on memory encoding and event segmentation. Notably, randomly selected videos, rather than personalized recommendations, have significant effects. This result highlighted how short videos were generated as the boundary condition of effects of short video watching on human memory encoding. Previous brain imaging studies have shown that, compared to random videos, watching personalized short videos can activate and modulate distinct large-scale brain networks[3,4]. However, these studies provided only

neuroimaging evidence for the distinction between the two types of short videos without exploring their potential differing cognitive impacts. Our study offers experimental evidence from both behavioral and eye-tracking analyses that randomly selected, rather than personalized, short videos negatively impact memory encoding. Consequently, we propose that the effects of short video watching should be considered from two perspectives: randomly selected short videos impair event segmentation and consequently memory encoding, while personalized short videos primarily lead to addiction[50].

Our study has several limitations that warrant acknowledgment. Firstly, no eye-tracking data were collected when participants engaged in the picture-based memory encoding task (i.e., Study 2). However, this was due to the methods used for analyzing eye movements being suitable only for capturing cognitive processes during naturalistic memory encoding tasks. Consequently, it was not possible to compare eye-tracking metrics between Studies 1 and 2. Secondly, we recruited participants who are typical daily users of short videos; specifically, we did not recruit those with a severe addiction to short videos. Future studies could include a cohort of heavy users in similar study design. Finally, we used commercially available movies as stimuli in Study 1, which may have masked memory impairments in participants who viewed short videos prior to memory encoding, due to the high attractiveness and emotional content of the movies. Future research could employ more emotionally neutral and educationally relevant video clips, such as recordings from real-life educational lectures[38,51], as memory encoding material.

In conclusion, our findings demonstrate that watching randomly selected short videos significantly impairs event segmentation during continuous memory encoding. Importantly,

these effects are absent in discrete memory tasks, underscoring the unique challenges posed by processing continuous information. Given that self-report studies have predominated in short video research[5], we advocate for interdisciplinary efforts, particularly involving cognitive psychologists and neuroscientists, to further investigate and elucidate the changes in event segmentation caused by short video watching across varied demographics, with a special focus on children and adolescents.

Reference

1. Zhang, X., Wu, Y. & Liu, S. Exploring short-form video application addiction: Socio-technical and attachment perspectives. *Telematics and Informatics* **42**, 101243 (2019).
2. Bobadilla, J., Ortega, F., Hernando, A. & Gutiérrez, A. Recommender systems survey. *Knowl Based Syst* **46**, 109–132 (2013).
3. Su, C., Zhou, H., Wang, C., Geng, F. & Hu, Y. Individualized video recommendation modulates functional connectivity between large scale networks. *Hum Brain Mapp* **42**, 5288–5299 (2021).
4. Su, C. *et al.* Viewing personalized video clips recommended by TikTok activates default mode network and ventral tegmental area. *Neuroimage* **237**, 118136 (2021).
5. Montag, C., Marciano, L., Schulz, P. J. & Becker, B. Unlocking the brain secrets of social media through neuroscience. *Trends Cogn Sci* **27**, 1102–1104 (2023).
6. Yao, N., Chen, J., Huang, S., Montag, C. & Elhai, J. D. Depression and social anxiety in relation to problematic TikTok use severity: The mediating role of boredom proneness and distress intolerance. *Comput Human Behav* **145**, 107751 (2023).
7. McCashin, D. & Murphy, C. M. Using TikTok for public and youth mental health – A systematic review and content analysis. *Clin Child Psychol Psychiatry* **28**, 279–306 (2023).
8. Chao, M., Lei, J., He, R., Jiang, Y. & Yang, H. TikTok use and psychosocial factors among adolescents: Comparisons of non-users, moderate users, and addictive users. *Psychiatry Res* **325**, 115247 (2023).
9. Qu, D. *et al.* The longitudinal relationships between short video addiction and depressive symptoms: A cross-lagged panel network analysis. *Comput Human Behav* **152**, 108059 (2024).
10. Parry, D. A. *et al.* A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nat Hum Behav* **5**, 1535–1547 (2021).
11. Yang, Y. *et al.* Time distortion for short-form video users. *Comput Human Behav* **151**, 108009 (2024).
12. Joiner, R. *et al.* The effect of different types of TikTok dance challenge videos on young women's body satisfaction. *Comput Human Behav* **147**, 107856 (2023).
13. Chiossi, F., Haliburton, L., Ou, C., Butz, A. M. & Schmidt, A. Short-Form Videos Degrade Our Capacity to Retain Intentions: Effect of Context Switching On Prospective Memory. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–15 (ACM, New York, NY, USA, 2023). doi:10.1145/3544548.3580778.
14. Chen, Y., Li, M., Guo, F. & Wang, X. The effect of short-form video addiction on users' attention. *Behaviour & Information Technology* **42**, 2893–2910 (2023).
15. Kurby, C. A. & Zacks, J. M. Segmentation in the perception and memory of events. *Trends Cogn Sci* **12**, 72–79 (2008).
16. Zacks, J. M. & Tversky, B. Event structure in perception and conception. *Psychol Bull* **127**, 3–21 (2001).
17. Zacks, J. M. Event Perception and Memory. *Annu Rev Psychol* **71**, 165–191 (2020).

18. Hasson, U., Furman, O., Clark, D., Dudai, Y. & Davachi, L. Enhanced Intersubject Correlations during Movie Viewing Correlate with Successful Episodic Encoding. *Neuron* **57**, 452–462 (2008).
19. Baldassano, C. *et al.* Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* **95**, 709-721.e5 (2017).
20. Chen, J. *et al.* Shared memories reveal shared structure in neural activity across individuals. *Nat Neurosci* **20**, 115–125 (2017).
21. Antony, J. W. *et al.* Behavioral, Physiological, and Neural Signatures of Surprise during Naturalistic Sports Viewing. *Neuron* **109**, 377-390.e7 (2021).
22. Pérez, P. *et al.* Conscious processing of narrative stimuli synchronizes heart rate between individuals. *Cell Rep* **36**, 109692 (2021).
23. Liu, W., Shi, Y., Cousins, J. N., Kohn, N. & Fernández, G. Hippocampal-Medial Prefrontal Event Segmentation and Integration Contribute to Episodic Memory Formation. *Cerebral Cortex* **32**, 949–969 (2022).
24. Cohn-Sheehy, B. I. *et al.* The hippocampus constructs narrative memories across distant events. *Current Biology* **31**, 4935-4945.e7 (2021).
25. Silva, M., Baldassano, C. & Fuentemilla, L. Rapid Memory Reactivation at Movie Event Boundaries Promotes Episodic Encoding. *The Journal of Neuroscience* **39**, 8538–8548 (2019).
26. Reagh, Z. M., Delarazan, A. I., Garber, A. & Ranganath, C. Aging alters neural activity at event boundaries in the hippocampus and Posterior Medial network. *Nat Commun* **11**, 3980 (2020).
27. Lee, C. S., Aly, M. & Baldassano, C. Anticipation of temporally structured events in the brain. *Elife* **10**, (2021).
28. Geerligs, L., van Gerven, M. & Güçlü, U. Detecting neural state transitions underlying event segmentation. *Neuroimage* **236**, 118085 (2021).
29. Yates, T. S. *et al.* Neural event segmentation of continuous experience in human infants. *Proceedings of the National Academy of Sciences* **119**, (2022).
30. Zacks, J. M. *et al.* Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* **4**, 651–655 (2001).
31. Ben-Yakov, A. & Henson, R. N. The hippocampal film editor: Sensitivity and specificity to event boundaries in continuous experience. *Journal of Neuroscience* **38**, 10057–10068 (2018).
32. Hahamy, A., Dubossarsky, H. & Behrens, T. E. J. The human brain reactivates context-specific past information at event boundaries of naturalistic experiences. *Nat Neurosci* (2023) doi:10.1038/s41593-023-01331-6.
33. Ryan, J. D. & Shen, K. The eyes are a window into memory. *Curr Opin Behav Sci* **32**, 1–6 (2020).
34. Kragel, J. E. & Voss, J. L. Looking for the neural basis of memory. *Trends Cogn Sci* **26**, 53–65 (2022).
35. Clewett, D., Gasser, C. & Davachi, L. Pupil-linked arousal signals track the temporal organization of events in memory. *Nat Commun* **11**, 4007 (2020).
36. Eddy, S. R. What is a hidden Markov model? *Nat Biotechnol* **22**, 1315–1316 (2004).

37. Nastase, S. A., Gazzola, V., Hasson, U. & Keysers, C. Measuring shared responses across subjects using intersubject correlation. *Soc Cogn Affect Neurosci* (2019) doi:10.1093/scan/nsz037.
38. Madsen, J., Júlio, S. U., Gucik, P. J., Steinberg, R. & Parra, L. C. Synchronized eye movements predict test scores in online video education. *Proceedings of the National Academy of Sciences* **118**, (2021).
39. Rajaram, S. Remembering and knowing: Two means of access to the personal past. *Mem Cognit* **21**, 89–102 (1993).
40. Černja, I., Vejmelka, L. & Rajter, M. Internet addiction test: Croatian preliminary study. *BMC Psychiatry* **19**, 388 (2019).
41. Kumar, M. *et al.* BrainIAK tutorials: User-friendly learning materials for advanced fMRI analysis. *PLoS Comput Biol* **16**, e1007549 (2020).
42. Sparrow, B., Liu, J. & Wegner, D. M. Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *Science (1979)* **333**, 776–778 (2011).
43. FEIGENBAUM, E. A. & SIMON, H. A. A THEORY OF THE SERIAL POSITION EFFECT. *British Journal of Psychology* **53**, 307–320 (1962).
44. Holleman, G. A., Hooge, I. T. C., Kemner, C. & Hessels, R. S. The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity. *Front Psychol* **11**, (2020).
45. Liu, W., Guo, J. & Li, H. Using artworks to understand human memory and its neural mechanisms. *New Ideas Psychol* **74**, 101095 (2024).
46. Reagh, Z. M., Delarazan, A. I., Garber, A. & Ranganath, C. Aging alters neural activity at event boundaries in the hippocampus and Posterior Medial network. *Nat Commun* **11**, 3980 (2020).
47. Masís-Obando, R., Norman, K. A. & Baldassano, C. Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *Elife* **11**, (2022).
48. Wittmann, B. C., Bunzeck, N., Dolan, R. J. & Düzel, E. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *Neuroimage* **38**, 194–202 (2007).
49. Shohamy, D. & Adcock, R. A. Dopamine and adaptive memory. *Trends Cogn Sci* **14**, 464–472 (2010).
50. Qin, Y., Omar, B. & Musetti, A. The addiction behavior of short-form video app TikTok: The information quality and system quality perspective. *Front Psychol* **13**, (2022).
51. Amalric, M., Roveyaz, P. & Dehaene, S. Evaluating the impact of short educational videos on the cortical networks for mathematics. *Proceedings of the National Academy of Sciences* **120**, (2023).

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments and Funding

W.L was supported by the National Natural Science Foundation of China (grant No. 32300879), Humanities and Social Sciences Fund, Ministry of Education (grant No. 22YJCZH109), Natural Science Foundation of Hubei (grant No. 2022CFB793), Open Research Fund of the State Key Laboratory of Cognitive Neuroscience and Learning (CNLYB2103), Open Research Fund of the Key Laboratory of Adolescent Cyber Psychology and Behavior (grant No. CCNUCYPSYLAB2022B10), the Major Program of the National Social Science Foundation of China (grant No. 22&ZD187) and self-determined research funds of CCNU from the colleges'basic research and operation of MOE(CCNU22QN020).

Contributions

W.L. and X.H conceived the Study; J.S.L and H.X.L collected the data; W.L. and J.S.L analyzed the data; W.L. and H.X.L prepared the first draft. W.L. and X.H reviewed and edited the manuscript, provided supervision, and obtained funding.