

“I” Always Cooperate: Self-Positivity Bias in Destination Memory for Cooperation and Cheating

Mengsi Li and Aiqing Nie

Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China

ABSTRACT

A growing body of research on episodic memory in interpersonal interaction divides it into two sub-types: source memory (i.e., memory for the match between inputted information and its sources) and destination memory (i.e., memory for the association between outputted information and its destinations). A topic related to this is memory for episodes involving cooperation and cheating. However, previous studies only examined source memory for cooperation and cheating, and neglected destination memory. The current study involved an event-related potential (ERP) experiment exploring destination memory and its retrieval-relevant ERP correlates in the social dilemma game. Participants interacted with virtual partners in the game, where both players were assigned either cooperation or cheating. Their destination memory for cooperative and cheating behaviors was later tested. Behaviorally, a clear self-positivity bias (i.e., enhanced memory for self-relevant positive information) was revealed: destination memory was more accurate for cooperation than cheating. The ERP correlates were also considered since they could reveal whether cooperation and cheating influenced different subprocesses of destination retrieval differently. Results indicated that the familiarity-driven FN400 and the recollection-based late positive complex (LPC) were more widely distributed or enhanced for cooperation than for cheating. In addition, late posterior negativity (LPN), which indexes efforts to reconstruct encoding episodes and evaluate retrieval outcome, lasted longer temporally for cooperation versus cheating, which might reflect greater difficulty in binding the self with morally unfavorable cheating behaviors. Altogether, both behavioral and ERP results constitute evidence for the self-positivity bias.

KEYWORDS

cheating
cooperation
destination memory
old/new effect
self-positivity bias

INTRODUCTION

Source Memory and Destination Memory in Interpersonal Interaction

The episodic memory system evolved to enable remembering of past happenings of day-to-day life and help with problem solving at present or in the future (Dickerson & Eichenbaum, 2010; Mahr & Csibra, 2018; Tulving, 2002). Daily encounters are suffused with interpersonal interactions. Episodic memory for our interactions with others thus has great significance in guiding our behaviors in the social world. The past decade has witnessed a surge in research on this topic, which can be viewed as a manifestation of researchers' endeavor to endow memory research with greater real-world relevance. Because of the bidirectional nature of interpersonal interaction, episodic memory for it can be further divided into two sub-types: *source memory* and *destination memory*. Source memory refers to the association between information

inputted from others and its deliverers/sources. In contrast, destination memory refers to the binding between information outputted by oneself and its receivers/destinations (El Haj, 2017; El Haj et al., 2016; El Haj et al., 2018; El Haj & Miller, 2018; Iliadou et al., 2019; Fischer et al., 2015; Gopie & MacLeod, 2009; Lindner et al., 2015; Shao et al., 2011).

The experimental paradigm to investigate episodic memory in interpersonal interactions can be summarized as follows. Participants are instructed to interact with real people or virtual partners (i.e., unknown/famous facial images) during encoding. They receive information from specific interactants in source encoding and deliver information to these people in destination encoding. In the subsequent source

Corresponding author: Aiqing Nie, Department of Psychology and Behavioral Sciences, Zhejiang University, 148 Tianmushan Road, Hangzhou, Zhejiang Province, China 310028
E-mail: nieaiq@126.com

memory test, participants are instructed to attribute inputted information to their associated sources, while in the destination memory test, they are supposed to link outputted information with their original destinations (El Haj, 2017; El Haj et al., 2016; El Haj et al., 2018; El Haj & Miller, 2018; Iliadou et al., 2019; Fischer et al., 2015; Gopie & MacLeod, 2009; Lindner et al., 2015; Shao et al., 2011).

Intuitively, source memory and destination memory have reverse directions of information transfer: input in terms of source memory and output in terms of destination memory (El Haj & Miller, 2018; Gopie & MacLeod, 2009). Yet, would such a difference in information transformation lead to divergent memory performance of the two? Using the above-mentioned experimental paradigm, some studies have provided an affirmative answer to this question by revealing greater performance in source memory over destination memory (El Haj et al., 2016; Fischer et al., 2015; Gopie & MacLeod, 2009). Gopie and MacLeod (2009) argued that such a result pattern appeared because in the destination versus source encoding conditions, individuals tend to connect information more closely with themselves than with the external world because they are more responsible for their behaviors and their behaviors are more informative of their own characteristics. This self-focus during destination encoding breaks the match between the information content and the exterior information episodes, resulting in more fallible destination compared with source memory. Gopie and MacLeod have also confirmed this argument in their studies. They found that destination memory was enhanced when attention was shifted from the self towards the interactive partner by letting participants say the name of the partner before telling them a fact (e.g., “Oprah Winfrey, the United States Postal Service handles 40% of the world’s mail volume”). However, contradictory findings also exist, as Lindner et al. (2015) revealed statistically indistinguishable source memory and destination memory performance for face-to-face interactions and argued that these two subtypes might just be two sides of the same coin. Taken together, it remains to be explored if destination memory could be discriminated from source memory.

Source Memory and Destination Memory for Cooperation and Cheating

As mentioned, whether destination memory and source memory rely on separate underpinnings is still under question. One way to answer this question is to manipulate features of the information transferred and see if destination memory and source memory are modulated differently by information features. Information exchanged during social interaction takes diverse forms (e.g., words, objects, and behaviors), and has different valences (i.e., positive, neutral, and negative). In the present study, we focused on behavior information of different valence categories (i.e., cooperative and cheating). Memory for cooperative and cheating behaviors is important in everyday life since it helps facilitate social cooperation and avoid being exploited. Source memory for cooperative and cheating behaviors of others have been well examined using the social dilemma game paradigm (Barclay, 2008; Bell et al., 2010; Bell et al., 2016; Li & Nie, 2021; Volstorf et al., 2011). In such a paradigm, participants interact with virtual interactants, deciding whether to cooperate with

or cheat them, and the interactants have choices of cooperation and cheating as well. In the subsequent source memory task, participants are shown with faces of interactants in the game and are asked to retrieve the associated behaviors. Studies applying such a paradigm demonstrated equivalent source memory for cooperative and cheating behaviors when the ratios of the two types of behaviors are equal (Bell et al., 2010; Bell et al., 2016; Li & Nie, 2021; Volstorf et al., 2011).

Despite the fact that social exchange is bidirectional and that theories of social exchange imply that maintenance of continuous social cooperation requires reciprocal altruists to track both the behaviors of their partners and their own behaviors (Cook et al., 2013), memory for one’s own cooperative and cheating behaviors (i.e., destination memory) has long been overlooked. We thus aimed to explore destination memory for cooperative and cheating behaviors applying a modified social dilemma game paradigm. Is destination memory affected by cooperation and cheating in a different way from source memory? Since previous research has shown destination memory to be more self-focused compared with source memory (Gopie & MacLeod, 2009), it is likely that destination memory for cooperation will be greater than cheating. This is because previous research has indicated a quite robust memory bias, termed the *self-positivity bias*, in that people are inclined to remember positive information connected with themselves better than negative information. Such a bias is supposed to be driven by a fundamental human need to protect a positive self-image and is thus advantageous for psychological wellbeing (Leary, 2007). It has already been revealed in autobiographical memory and episodic memory. For instance, when asked to recall self-related daily events, participants reported a significantly larger number of positive events than negative ones (Betz & Skowronski, 1997; Ritchie et al., 2017). In addition, a study confirmed the self-positivity bias in episodic memory (Zhang et al., 2018). In this study, participants were instructed to indicate whether positive and negative trait adjectives could describe themselves (self-reference) or another person (other-reference) during encoding, and in the subsequent source memory task, participants were required to make source classifications for words: whether a given word was previously encoded in the self-reference or the other-reference condition. Participants made more accurate source classifications for positive than negative words under the self-reference condition (Zhang et al., 2018). As defined, self-reference should be a core part of destination memory (El Haj et al., 2014; El Haj et al., 2016; El Haj & Miller, 2018; Fischer et al., 2015; Gopie & MacLeod, 2009). Thus, we aimed to extend previous research and examine whether behavior valence would modulate destination memory in a way conforming to the self-positivity bias (i.e., enhanced destination memory for cooperation versus cheating).

Old/New Effects in Source Memory

Memory retrieval can be further divided into different subprocesses, and different subprocesses of destination memory may be modulated by behavior valence in different ways. The technique of event-related potential (ERP) has revealed temporal dynamics of retrieval through the neural index of successful retrieval, termed *old/new effects*. Old/new effects are regarded as the difference in neural activities between

previously encoded (or old) information that is correctly retrieved and novel unstudied (or new) information that is correctly identified as novel, therefore reflecting successful retrieval (Bader & Mecklinger, 2017; Höljtje & Mecklinger, 2018; Park & Donaldson, 2019; Proverbio et al., 2019; Ross et al., 2018; Rugg & Curran, 2007; Ye et al., 2019). No study to date has examined old/new effects in destination memory, but there exist numerous studies exploring the effects in source memory (Addante et al., 2012; Leynes et al., 2017; Minor & Herzmann, 2019; Nardini & Leynes, 2020; Nie et al., 2019; Proverbio et al., 2019; Ventura-Bort et al., 2020; Ye et al., 2019). Destination memory and source memory are similar in the sense that they both demand the binding between item information and associated contextual details. We thus reviewed different components of old/new effects in source memory below. In the present study, we explored (a) whether these components could be recorded in destination memory, and (b) whether different components would be similarly affected by cooperative and cheating behaviors.

Typically, research has examined two types of old/new effects in source memory: Source-correct old/new effects, or the difference in neural correlates between correctly retrieved old items with correct source identifications and correctly classified new items, and source-incorrect old/new effect, or the difference in neural correlates between correctly identified old items with incorrect source identifications and correctly discriminated new items. Other studies also compared waveform difference between source-correct trials and source-incorrect trials (Addante et al., 2012; Leynes et al., 2013; Leynes et al., 2017; Leynes & Kakadia, 2013; Leynes & Nagovsky, 2016; Mollison & Curran, 2012; Nardini & Leynes, 2020; Ventura-Bort et al., 2020; Woroch & Gonsalves, 2010). Four components of old/new effects have been typically revealed in source memory, each corresponding to a unique subprocess that underlies source memory retrieval. Temporal and topographic features, and the psychological meaning of each component are detailed below.

There are two early components, termed the *mid-frontal effect* (FN400) and the *late positive complex* (LPC), respectively (Addante et al., 2012; Höljtje & Mecklinger, 2018; Leynes et al., 2017; Minor & Herzmann, 2019; Mollison & Curran, 2012; Nardini & Leynes, 2020; Proverbio et al., 2019; Ventura-Bort et al., 2020; Woroch & Gonsalves, 2010). They correspond to two functionally distinct mnemonic signals proposed by the dual-process model of episodic memory: familiarity and recollection. The FN400 is the putative index of familiarity, conceptualized as a fast-acting, relatively involuntary process that mostly aids pure recognition of previously encountered items. The LPC, by contrast, is the correlate of recollection, which happens later under conscious control and makes contextual information of prior occurrence accessible (Bader & Mecklinger, 2017; Curran, 2000; Höljtje & Mecklinger, 2018; MacKenzie & Donaldson, 2007; Migo et al., 2012; Nie et al., 2019; Proverbio et al., 2019; Ross et al., 2018; Rugg & Curran, 2007; Ye et al., 2019; Yonelinas et al., 2010). Regarding temporal and topographical features, the FN400 starts approximately from 300 ms and is usually largest around 400 ms poststimulus onset, which has more negative-going amplitudes for correctly rejected new information than for successfully retrieved old information at mid-frontal electrodes. The LPC arises later

at about 400-500 ms poststimulus onset and captures the old/new difference over the parietal region, manifesting as enhanced positivity for correctly classified old versus new information (Addante et al., 2012; Höljtje & Mecklinger, 2018; Leynes et al., 2017; Minor & Herzmann, 2019; Mollison & Curran, 2012; Nardini & Leynes, 2020; Proverbio et al., 2019; Ventura-Bort et al., 2020; Woroch & Gonsalves, 2010).

The *late posterior negativity* (LPN) and the *right-frontal effect* (RFE) are two late components that are more closely related with source memory retrieval (compared with retrieval of item information). The LPN reflects continued evaluation of retrieval products and effort to reconstruct previously experienced encoding episodes when the retrieval demand is not satisfied by recollection. It is in an opposite form to the FN400 and the LPC, in that it has more negative-going ERPs for old items whose sources are correctly retrieved than for correctly rejected new information. It starts from 600 ms at the earliest and distributes over posterior regions (Barrick & Dillon, 2018; Hellerstedt & Johansson, 2016; Leynes et al., 2013; Leynes et al., 2017; Leynes & Kakadia, 2013; Leynes & Nagovsky, 2016; Mecklinger et al., 2016; Nardini & Leynes, 2020; Sommer et al., 2018; Nie et al., 2013). The RFE indexes general monitoring or evaluation of retrieval outcomes. It shows more positive-going ERPs for old information with correct sources than for new information from 800 ms or a later onset over the frontal region, and the right-frontal region in particular (Cruse & Wilding, 2009; Hayama et al., 2008; Leynes et al., 2013; Leynes et al., 2017; Leynes & Kakadia, 2013; Leynes & Nagovsky, 2016; Nardini & Leynes, 2020; Nie et al., 2019; Ye et al., 2019).

Based on the ERP research in source memory retrieval reviewed above, in our study, we examined whether the FN400, LPC, LPN, and RFE could be recorded for destination-correct old/new effects and destination-incorrect old/new effects. To be specific, destination-correct old/new effects measure the difference in neural activation between correctly retrieved old items with correct destination identifications (destination-correct) and correctly classified new items. Destination-incorrect old/new effects reflect the difference in neural activation between correctly identified old items with incorrect destination identifications (destination-incorrect) and correctly discriminated new items. We also explored if waveform difference between destination-correct and destination-incorrect trials could be obtained in the latency window of any of the four components.

The Current Study

The goal of the current study was to provide evidence for the self-reference process in destination memory (unique in destination memory but not in source memory) by revealing that behavior valence modulates destination memory in a pattern consistent with the self-positivity bias. To this end, influences of cooperation and cheating on destination memory and corresponding ERP old/new effects were examined in a modified social dilemma game. Behaviorally, we anticipated destination memory for cooperation to be greater than for cheating. For retrieval-relevant old/new effects, we first predicted that the four components of destination-correct old/new effects (i.e., the FN400, LPC, LPN, and RFE) could be recorded. For destination-incorrect tri-

als, the LPC, LPN, and RFE might be less reliable since they are more related to retrieval of contextual details. Then we expected the recorded effects to behave distinctly in response to cooperation and cheating. Specifically, we expected there to be enhanced FN400 and LPC for cooperation than cheating if the self-positivity bias held true. The LPN and RFE might also act differently under the cooperative and cheating conditions. However, more specific predictions regarding these two components were not made, as previous findings were insufficient. We also considered the waveform difference between destination-correct and destination-incorrect trials, and we expected it to be modulated by behavior valence. Yet, no specific a priori hypothesis was made due to lack of previous research.

METHODS

Participants

Thirty-five undergraduate and graduate students (23 females) with a mean age of 21 years (range: 18–32) participated in the current study for course credit. All were right-handed native Chinese speakers, had normal or corrected-to-normal visual acuity, and reported no history of neurological impairments or psychiatric disorders. Each participant gave written informed consent prior to the experiment and was debriefed after the experiment. Data from two participants were removed from analyses since they provided insufficient artifact-free trials (< 16) under more than one condition, and data from three other participants were excluded because of technical problems during recording, leaving 30 effective participants in data analyses. The G*Power 3.1 software (Faul et al., 2009) was used to determine the sample size. A minimum sample size of 28 would allow for detecting a medium size effect ($f = 0.25$) under standard criteria (two-tailed $\alpha = 0.05$, $1-\beta = 0.80$) if considering the one-way repeated-measures analysis of variance (ANOVA) that included the factor of behavior type (cooperative, neutral, cheating), with an assumed correlation ρ of .50 between the levels of behavior type. The effect size was chosen based on previous literature exploring source memory for cooperation and cheating (Barclay, 2008; Bell et al., 2010; Bell et al., 2016; Volstorf et al., 2011). These studies showed a small-to-medium effect size ($f = 0.10-0.25$) as defined by Cohen (1988). The medium effect size was chosen here to reduce time and cost required by an ERP study.

Design

This experiment was a single factor, within-subject design with the factor of behavior type (cooperative, neutral, and cheating). Behavior types were determined by behaviors assigned to participants in the social dilemma game.

Materials

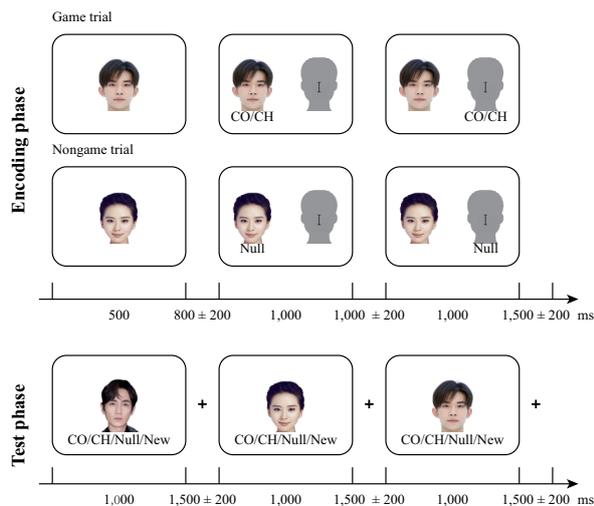
A pool of 450 colored facial images of Chinese celebrities (formal stimuli) and 94 colored images of unknown individuals (34 for practice and 60 as fillers in the formal experiment) were used. Expressions on all the faces were neutral or slightly positive, and there was an even

mix of females and males for both famous and unknown faces. Most of the famous faces were obtained from a previous study (Nie & Yu, 2021), and lab assistants helped collect the other famous and unknown faces from the Internet. The famous face images came from singers, actors, hosts, athletes, politicians, entrepreneurs, artists, and so forth. Two groups of 40 college students who did not participate in the experiment were recruited to evaluate the popularity and trustworthiness of the famous face images. For popularity, the first group judged how familiar they were with each famous face image on a five-point scale (1 = *totally unfamiliar*, 3 = *unsure*, and 5 = *totally familiar*), resulting in an average popularity score of 3.9 ± 0.814 . The second group was required to estimate on a five-point scale how trustworthy they thought each famous face image was (1 = *totally untrustworthy*, 3 = *unsure*, and 5 = *totally trustworthy*), and the average score of trustworthiness was 3.247 ± 0.340 . The 94 unknown face images were used because of the difficulty in obtaining more famous face images. We processed all the above facial images into 270×360 pixels, 72 px/in., and 50.8 cd/m² with white background using Adobe Photoshop CC 2018.

The formal images were divided into 15 blocks of 30 images each, counterbalanced by gender, age, and occupation. Twenty-four formal facial images of each block were set as to-be-remembered items during encoding. The 24 encoding trials were divided further into 16 (2/3) game trials and 8 (1/3) nongame trials. Images in half of the game trials were accompanied with cooperative self behaviors while those in the other half were with cheating self behaviors. Those in the nongame trials were accompanied with neutral self behaviors. Each block contained another four filler trials, two appearing at the beginning and two at the end of each block, to eliminate the influence of primacy and recency effects. Behaviors associated with the fillers were half cooperative and half cheating, randomly assigned by the programming. During retrieval, each block included the 24 studied images and six unstudied new ones, while fillers were excluded.

Procedure

Upon arriving at the laboratory, participants were comfortably seated in a quiet, electrically shielded room with a dim light 70 cm in front of a SAMSUNG Sync Master CRT monitor (refresh frequency = 100 Hz, screen resolution = 1024×768 pixels). All participants took part in practice trials with feedback to get familiar with the instructions and experimental protocols first, and then took part in 15 blocks of the formal experiment with identical protocols as the practice, but with no feedback. Every block contained three phases: encoding, distraction, and test. There was always a 5 min rest period between two successive blocks to reduce interferences across blocks and for participants to relax. It took nearly three hours to finish a whole experiment, including time spent in preparing the ERP equipment, the practice and the formal experiment trials, and breaks. Stimulus presentation was controlled and behavioral responses were recorded with MATLAB software v2010b (Psychtoolbox). All stimuli were presented against a white background. Visual angles for facial images and the head portrait representing the participants were 5.880° (horizontal) and 7.841° (vertical). Participants were reminded to avoid eye blinks and head movements during stimu-

**FIGURE 1.**

Schematic illustration of the encoding phase (including a game trial and a nongame trial) and the test phase. CO = cooperate, CH = cheat.

lus presentation. Index and thumb fingers of the two hands were used for key presses during encoding, while index and middle fingers of both hands were used for key presses during the test trials. Assignment of hands was counterbalanced among participants. Figure 1 depicts the schematic illustration of the experimental procedure with sample facial images and descriptions for behavior instructions.

ENCODING PHASE

The current study applied a social dilemma game as the encoding task, which was modified based on tasks in previous source memory research (Bell et al., 2016). Participants were told that they would interact with celebrities in a game where the celebrities and themselves would be assigned with cooperative and cheating behaviors randomly by the programming in each round of the game (or in some rounds, no behavior would be assigned to them or their partners). It was stressed to all participants that the points they won or lost in each round of the game would be determined by the behaviors of both players, according to rules (see Table 1). Mutual cooperation benefited both sides while cheating of both sides caused a loss to both. In case when one side cooperated while the other cheated, the cooperator lost while the cheater gained points. Nongame trials (where no behavior was assigned to either player) did not result in a change in points. Participants were told that the higher total score they acquired, the more course credits they would be offered for the participation in the experiment. However, all participants would, in fact, gain even scores and hence would receive equal course credits. This manipulation was not revealed to the participants before the experiment and was instead explained to them afterwards. Participants were asked to judge the instructed behaviors by key presses, to try their best to connect the instructed behaviors with themselves, and to engage emotionally by imagining they were in real interactions with the celebrities. For the memory task, they were informed to pay attention to both the partners' and their own behaviors in the game as

TABLE 1.

Scoring Rule in the Social Dilemma Game.

		Interactant's behavior	
		Cooperate	Cheat
Participant's behavior	Cooperate	5, 5	-10, 10
	Cheat	10, -10	-5, -5

Note. Payoff outcomes in the game trial were determined by behaviors of both sides. Scores of participants are presented in bold font, while scores of the interactants are shown in regular font.

their brain activity would be recorded continuously, although only their own but not the partners' behaviors would be tested later.

To study destination memory, we controlled not only the behaviors of the participants' partners (as source memory studies have usually done) but also those of the participants. In the game trials, interactants were assigned with either cooperation or cheating by the programming first, and participants were instructed to cooperate with half of the interactants who cooperated (or cheated), and to cheat the other half. Accordingly, the ratios of cooperation and cheating were equivalent for both players. This was how all participants were designed to receive equal points in the game. While being debriefed, no participant reported explicit detection of such manipulation. In nongame trials, behaviors of both players were presented as "Null".

During encoding, each face image was presented three times. Each trial started with a fixation cross at the center of the screen for 500 ± 100 ms. A face image then appeared for the first time at the center of the screen for 500 ms, followed by a white blank screen for $1,000 \pm 200$ ms. Subsequently, the same face image was presented again along with the behavior assigned to it. During the second presentation, the image was placed at the center of the left side of the screen. Meanwhile, there was a gray head portrait representing the participants at the center of the right side of the screen, equaling in size to the face image. A behavior instruction, either "Cooperate" or "Cheat," was displayed beneath the face image of the interactant. The second presentation of the face image and the gray head portrait lasted for 1,000 ms, and was followed by a white blank screen of $1,000 \pm 200$ ms. Participants were told to pay attention to the partner's behavior and try to engage emotionally as if they were encountering real interactions, although their memory for the partner's behavior would not be tested later. Then, the same face image was presented for the third time together with the behavior assigned to the participants. Again, the same face image was shown on the left of the screen while the head portrait was shown on the right for 1,000 ms, and the behavior instruction of "Cooperate" or "Cheat" was placed under the head portrait to indicate the instructed behaviors for participants. During the third presentation of the face image, participants were required to judge whether they were instructed to cooperate or cheat by pressing the "F" or "J" keys on the keyboard, try to connect the behavior with themselves, and remember this behavior as their own choice. Then, a white blank screen was shown for $1,500 \pm 200$ ms. The protocols of the nongame trials were identical to the game trials, except for the behaviors assigned and the key presses:

The behaviors assigned to the interactants and the participants were both shown as “Null,” and participants were instructed to press the space bar on the keyboard when judging their behavior type during the third face image presentation.

DISTRACTION

There was a 1 min distraction task sandwiched between the encoding and the test in each block to prevent recency effects, during which participants solved several easy arithmetic problems.

TEST PHASE

During retrieval, each block began with a fixation cross at the center of the screen for 500±100 ms. Afterwards, a test face probe was presented for 1,000 ms, followed by a 1,500±200 ms blank screen. Participants should decide whether a cooperative, neutral, or cheating behavior was assigned to them when interacting with the given face image during encoding, or whether the face was new, by pressing the “D,” “F,” “J,” or “K” keys on the keyboard.

Electrophysiological Recording

Scalp electroencephalographic activity (EEG) was continuously recorded from 32 Ag/AgCl electrodes extended from the international 10/20 system using the Neuroscan software and hardware (Scan, SynAmps, Compumedics, El Paso, USA). Two electrodes placed at the outer canthi of the two eyes recorded the horizontal electro-oculograms, and two affixed above and below the left eye recorded the vertical electro-oculograms. The EEG recording was amplified with a gain of 500 and digitized at a sampling rate of 500 Hz per channel. All channels were referenced to the right mastoid and were band-pass filtered from 0.05 to 40 Hz online. All electrode impedances were maintained lower than 5 kΩ.

RESULTS

We used IBM SPSS Statistics version 22 to perform behavioral and ERP data analyses. Repeated-measures ANOVA were corrected using the Greenhouse-Geisser method when violating sphericity. The *F* ratios were reported with Greenhouse-Geisser adjusted *p* values and degrees of freedom, the effect sizes of partial eta-squared (η_p^2), and coefficients (ϵ). All inferential analyses applied an α level of .05 (two-tailed). Only statistically significant results of interest are reported. We refer to old faces with correct destination classifications and those with incorrect

destination classifications as “destination-correct” and “destination-incorrect” respectively, and correctly rejected new faces as “new.”

Behavioral Analyses and the Data

We first analyzed responses during encoding to review the participants’ compliance with the instructed behaviors assigned to them in the social dilemma game. To accomplish this, we counted different types of participants’ actual behaviors in response to different types of instructed behaviors assigned to them (shown in Table 2). Overall, participants’ actual behaviors were highly compliant with the instructed behaviors (around 97% in compliance) under all the three conditions.

To compare destination memory performance across the three behavior conditions, and more specifically, to see if there existed a self-positivity bias (greater destination memory for cooperative versus cheating behaviors), we conducted three repeated-measures ANOVAs with the factor of behavior type (cooperative, neutral, and cheating) separately on (a) destination-correct rates, (b) destination-incorrect rates, and (c) unbiased destination-correct rates. Destination-correct rates referred to the proportion of old cooperative trials that were given cooperative destination classifications to the total number of old cooperative trials. Destination-incorrect rates equaled to the number of old cooperative trials with incorrect destination classifications (neutral or cheating) divided by the total number of old cooperative trials, thus reflecting destination memory errors. The rates for neutral and cheating conditions were calculated similarly. Unbiased destination-correct rates were calculated based on destination-correct rates, while taking into consideration the guessing bias. Unbiased source-correct rates (or unbiased hit rates of source memory) have been used in previous source memory literature (Suzuki & Suga, 2010; Wagner, 1993). In the present study, we computed destination-correct rates referring to the method in source memory research: The unbiased destination-correct rate was the product of the destination-correct rate and a response-bias term. The response-bias term accounted for the guessing bias. For example, in the cooperative condition, it was the proportion of old faces in the cooperative condition correctly classified as cooperative to the total number of faces (including new faces and all types of old faces) classified as cooperative. Table 3 shows the relevant data.

For destination-correct rates, the ANOVA revealed a statistically significant effect of behavior type, $F(1.998, 57.933) = 24.410, p < .001, \eta_p^2 = 0.457, \epsilon = 0.999$. Pairwise analysis suggested higher rates for cooperative versus neutral behaviors, $p < .001, 95\% \text{ CI } [0.132, 0.282]$, cooperative versus cheating behaviors, $p = .019, 95\% \text{ CI } [0.012, 0.165]$, and cheating

TABLE 2.

Compliance With Instructed Behaviors in the Social Dilemma Game. Participants’ Four Types of Actual Behaviors (Cooperative, Neutral, Cheating, vs. No Response) Categorized by the Three Types of Instructed Behaviors (Cooperative, Neutral, vs. Cheating) are Counted

	Actual cooperative		Actual neutral		Actual cheating		No response	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Instructed cooperative	97.6%	5.9%	1.2%	5.7%	0.5%	0.8%	0.8%	1.9%
Instructed neutral	1.3%	5.7%	96.9%	6.9%	0.3%	0.6%	1.5%	3.6%
Instructed cheating	0.9%	0.8%	0.1%	0.3%	97.2%	5.8%	1.8%	5.8%

Note. Ratios of compliant responses (e.g., actual behaviors that are consistent with the instructed cooperative behaviors) are marked in italics.

TABLE 3.

Destination-Correct Rates (Proportions of Items Accompanied by Successful Destination Identifications) and Destination-Incorrect Rates (Proportions of Items Accompanied by Unsuccessful Destination Identifications), and Unbiased Destination-Correct Rates (Destination-Correct Rates Corrected for Guessing Biases) as a Function of the Factor Behavior Type (Cooperative, Neutral, vs. Cheating)

	Cooperative		Neutral		Cheating	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Destination-correct	0.596	0.120	0.389	0.133	0.507	0.131
Destination-incorrect	0.316	0.091	0.500	0.111	0.399	0.125
Unbiased destination-correct	0.313	0.105	0.151	0.065	0.253	0.091

versus neutral behaviors, $p = .001$, 95% CI [0.044, 0.193]. In terms of destination-incorrect rates, the ANOVA also confirmed a statistically significant effect of behavior type, $F(1.893, 54.892) = 21.083$, $p < .001$, $\eta_p^2 = 0.421$, $\epsilon = 0.946$. As demonstrated by corresponding pairwise analyses, the rates were lower in cooperative versus neutral behaviors, $p < .001$, 95% CI [-0.251, -0.118], in cooperative versus cheating behaviors, $p = .040$, 95% CI [-0.163, -0.003], and in cheating versus neutral behaviors, $p = .003$, 95% CI [-0.170, -0.032]. The ANOVA for unbiased destination-correct rates demonstrated an effect of behavior type as well, $F(1.578, 45.775) = 73.219$, $p < .001$, $\eta_p^2 = 0.716$, $\epsilon = 0.789$, pair-wise analysis of which further indicated statistically significantly greater unbiased destination-correct rates for cooperative versus neutral behaviors, $p < .001$, 95% CI [0.121, 0.205], cooperative versus cheating behaviors, $p < .001$, 95% CI [0.033, 0.087], and cheating versus neutral behaviors, $p < .001$, 95% CI [0.070, 0.135]. To summarize, destination-correct and unbiased destination-correct rates both decreased across cooperative, cheating, and neutral behaviors, while destination-incorrect rates increased across cooperative, cheating, and neutral behaviors. The above data was in accordance with the self-positivity bias.

ERP Analyses and the Data

Offline EEG data were processed using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes in MATLAB. The EEG voltages were algebraically rereferenced to the average of the left and the right mastoids. Before averaging, independent component analysis (ICA) in EEGLAB was used to correct ocular artifacts. Following ocular correction, artifact detection in ERPLAB was applied to reject trials in which ERP amplitudes exceeded $\pm 80 \mu\text{V}$. The continuous EEG data were time-locked to face onsets, segmented from -100 to 1,500 ms, and baseline corrected over the prestimulus interval. The ERPs were averaged from destination-correct cooperative, neutral, and cheating trials, destination-incorrect cooperative, neutral, and cheating trials, and new trials. To reiterate, old trials with correct destination classifications and those with incorrect destination classifications were termed "destination-correct" and "destination-incorrect" trials, respectively, and new trials correctly discriminated as new were termed "new." Trial counts for all the above conditions per participants

included in grand average ERPs had a minimum of 16, in line with the criterion adopted by former ERP research (Minor & Herzmann, 2019; Nardini & Leynes, 2020; Ye et al., 2019). The mean number of trials used for averaging were 62, 40, and 54 for destination-correct cooperative, neutral, and cheating trials, respectively; 33, 52, and 40 for destination-incorrect cooperative, neutral, and cheating trials, respectively; and 67 for new trials.

Based on visual inspection of the current waveforms and extant ERP literature focusing on old/new effects (Addante et al., 2012; Cruse & Wilding, 2009; Hayama et al., 2008; Hölting & Mecklinger, 2018; Leynes et al., 2013; Leynes et al., 2017; Leynes & Kakadia, 2013; Leynes & Nagovsky, 2016; Minor & Herzmann, 2019), mean amplitudes were measured in four time segments: 300–450, 450–600, 600–1,000, and 1,000–1,400 ms. The regions of interest (ROIs) consisted of three channel groups each over frontal (F3, Fz, and F4), central (C3, Cz, and C4), parietal (P3, Pz, and P4), and occipital (O1, Oz, and O2) scalp regions. Four components of old/new effects were to be considered, including the FN400, LPC, LPN, and RFE. For theoretical consideration, the FN400 referred to old/new effects between 300 and 450 ms over frontal region, while the LPC referred to old/new effects between 450 and 600 ms over parietal region. However, the effects were also recorded in other scalp regions within the two time segments, so these additional effects were also reported. The latency windows of 600–1,000 and 1,000–1,400 ms were to capture the LPN and REF. The LPN focused on the parietal and occipital regions in both 600–1,000 and 1,000–1,400 ms, and the REF concerned the frontal region in 1,000–1,400 ms. For cooperative, neutral, and cheating cases, we investigated destination-correct old/new effects (destination-correct versus new), destination-incorrect old/new effects (destination-incorrect versus new), as well as the difference between destination-correct and destination-incorrect trials, referring to previous research on old/new effects in source memory (Addante et al., 2012; Leynes et al., 2017; Minor & Herzmann, 2019; Nardini & Leynes, 2020; Proverbio et al., 2019; Ventura-Bort et al., 2020; Ye et al., 2019).

Figure 2 depicts the grand average waveforms for the cooperative condition: destination-correct cooperative, destination-incorrect cooperative, and new faces. Figure 3 plots the topographical maps of ERP difference waves for the cooperative condition: destination-correct cooperative minus new faces, destination-incorrect cooperative minus new faces, and destination-correct cooperative minus destination-incorrect cooperative faces. Figures 4 and 5 illustrate the grand average waveforms and the topographical maps for the neutral condition. Figures 6 and 7 present those for the cheating condition.

DESTINATION-CORRECT OLD/NEW EFFECTS

To explore whether the four components of destination-correct old/new effects (i.e., the FN400, LPC, LPN, and RFE) could all be obtained, and if any component was sensitive to behavior type, we conducted a $4 \times 4 \times 3$ (face type [cooperative, neutral, cheating, and new] \times region [frontal, central, parietal, and occipital] \times hemisphere [left, medial, and right]) repeated-measures ANOVA on mean amplitudes in each latency window comparing destination-correct trials of

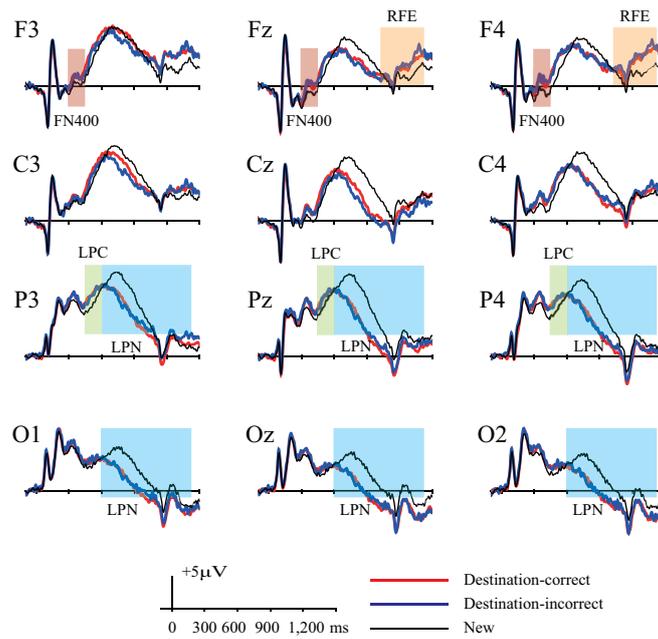


FIGURE 2.

Grand average waveforms of destination-correct, destination-incorrect, and new faces under the cooperative behavior condition, with the color bars indicating old/new effects in the four examined time segments.

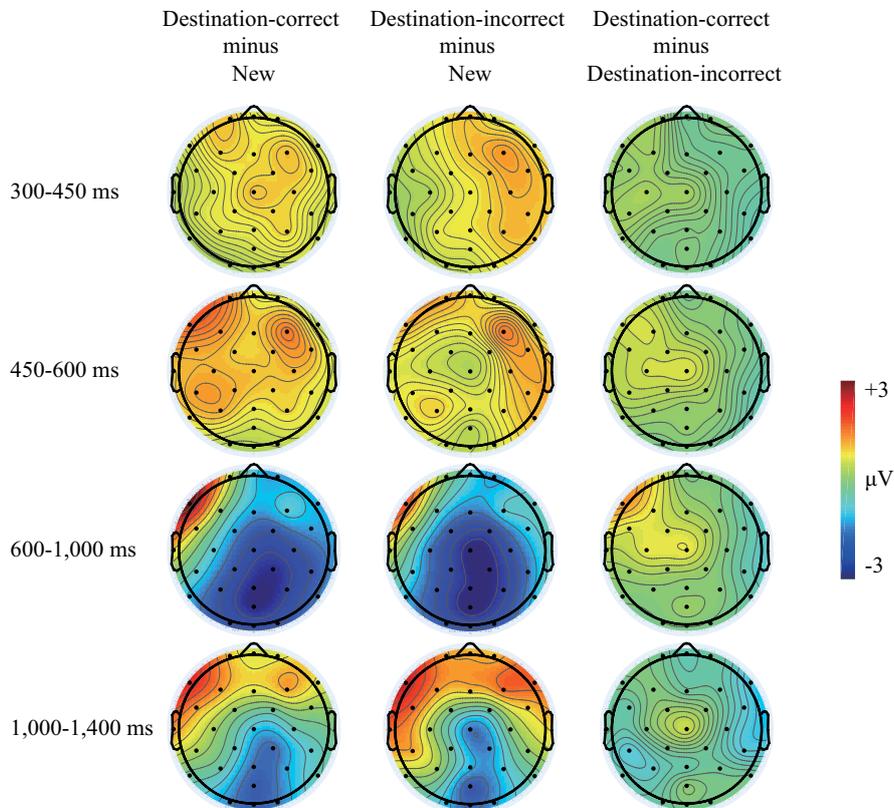


FIGURE 3.

Topographic maps in the cooperative behavior condition, derived from difference waveforms between destination-correct and new faces, destination-incorrect and new faces, and destination-correct and destination-incorrect faces, for the four examined time segments.

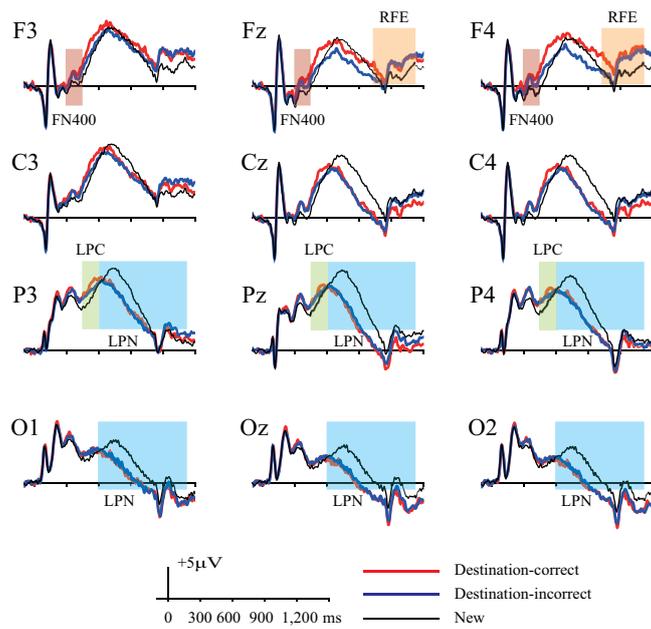


FIGURE 4. Grand average waveforms of destination-correct, destination-incorrect, and new faces in the neutral behavior condition, with the color bars indicating old/new effects in the four examined time segments.

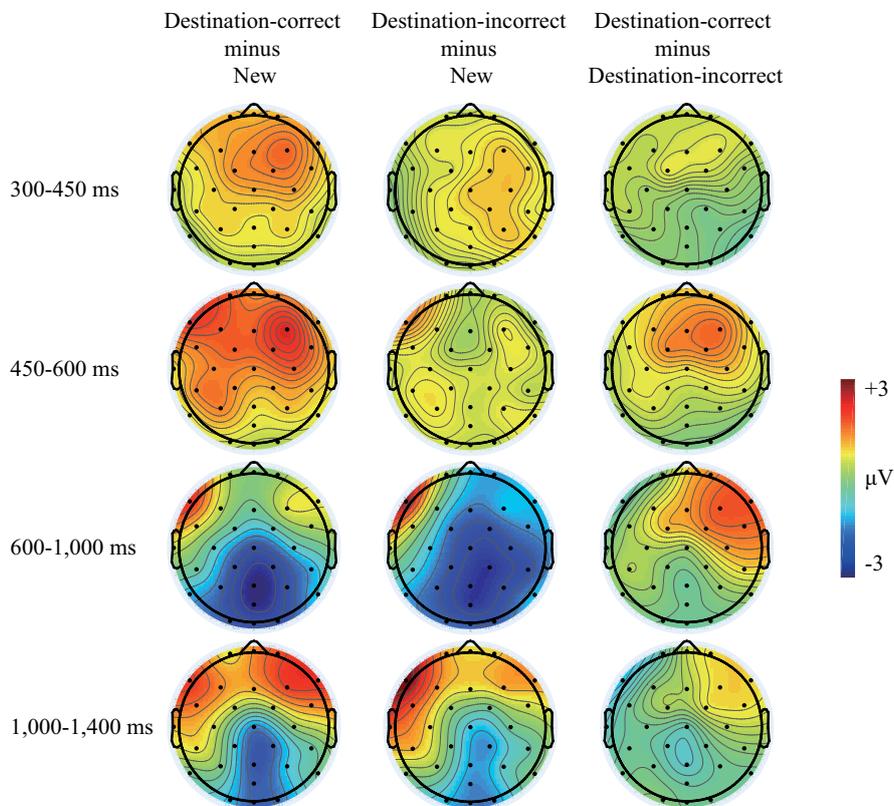


FIGURE 5. Topographic maps in the neutral behavior condition, derived from difference waveforms between destination-correct and new faces, destination-incorrect and new faces, and destination-correct and destination-incorrect faces, for the four examined time segments.

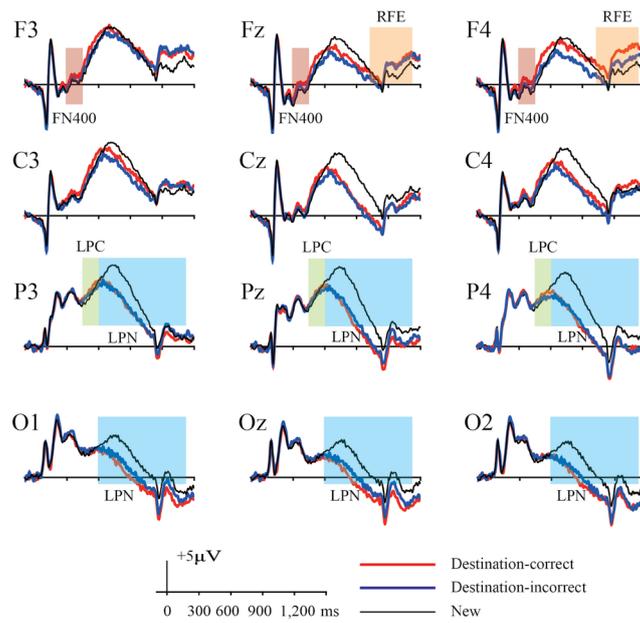


FIGURE 6.

Grand average waveforms of destination-correct, destination-incorrect, and new faces in the cheating behavior condition, with the color bars indicating old/new effects in the four examined time segments.

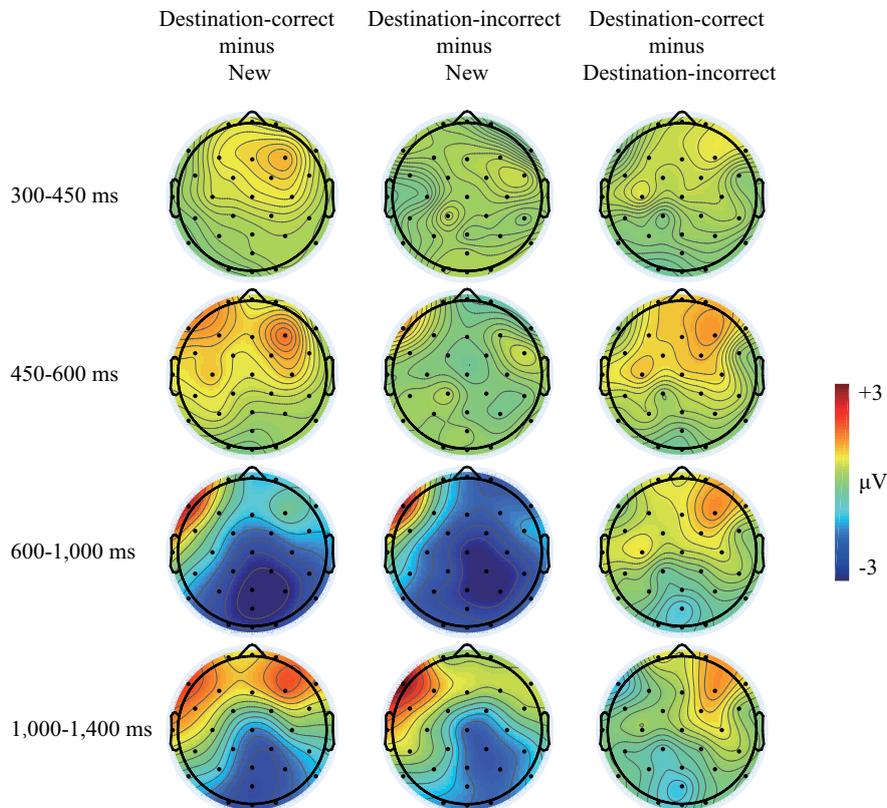


FIGURE 7.

Topographic maps in the cheating behavior condition, derived from difference waveforms between destination-correct and new faces, destination-incorrect and new faces, and destination-correct and destination-incorrect faces, for the four examined time segments.

the three behavior types and the new trials. We report only the data relevant to the factor of face type.

300–450 ms (FN400). The ANOVA yielded a statistically significant main effect of face type, $F(2.981, 86.454) = 6.370, p = .001, \eta_p^2 = 0.180, \epsilon = 0.994$, and a statistically significant two-way interaction between face type and region, $F(4.327, 125.495) = 3.023, p = .018, \eta_p^2 = 0.094, \epsilon = 0.481$. Post-hoc comparisons for the two-way interaction indicated that over the frontal region, new faces elicited enhanced negativity compared to destination-correct cooperative, neutral, and cheating faces, $p = .002, 95\% \text{ CI } [-1.641, -0.308]; p = .001, 95\% \text{ CI } [-2.391, -0.540]; p = .010, 95\% \text{ CI } [-1.748, -0.178]$, while over the central and parietal regions, new faces only elicited more negative-going amplitudes than destination-correct cooperative and neutral faces, central: $p = .015, 95\% \text{ CI } [-1.643, -0.130]$ and $p = .005, 95\% \text{ CI } [-2.029, -0.283]$; parietal: $p = .042, 95\% \text{ CI } [-1.549, -0.018]$ and $p = .014, 95\% \text{ CI } [-1.661, -0.138]$. Amplitudes elicited by destination-correct cooperative, neutral, and cheating faces did not statistically significantly differ over all the regions. These data suggest destination-correct FN400 for cooperative and neutral faces across the frontal, central, and parietal regions, and for cheating faces over the frontal region only.

450–600 ms (LPC). As shown by the ANOVA, there was a statistically significant main effect of face type, $F(2.916, 84.563) = 4.053, p = .010, \eta_p^2 = 0.123, \epsilon = 0.972$, a two-way interaction of face type by region, $F(4.283, 124.261) = 4.932, p = .001, \eta_p^2 = 0.145, \epsilon = 0.476$, and a three-way interaction, $F(9.842, 285.431) = 1.876, p = .049, \eta_p^2 = 0.061, \epsilon = 0.547$. Post-hoc comparisons for the two-way interaction revealed that over the frontal region, destination-correct cooperative faces, $p = .043, 95\% \text{ CI } [0.023, 2.157]$, and destination-correct neutral faces, $p = .002, 95\% \text{ CI } [0.572, 3.054]$, yielded larger positivity compared with new faces; over the central region, only amplitudes for destination-correct neutral faces were more positive-going than those for new faces, $p = .015, 95\% \text{ CI } [0.193, 2.425]$; over all regions, amplitudes for destination-correct cooperative, neutral, and cheating faces were comparable. Post-hoc comparisons for the three-way interaction continued to demonstrate that at the left-frontal site (F3), amplitudes for destination-correct cooperative and neutral faces were more positive-going than those for new faces, while at the right-frontal (F4), destination-correct cooperative, neutral, and cheating faces all elicited more positive-going amplitudes compared to new ones, $p = .006, 95\% \text{ CI } [0.298, 2.363]; p < .001, 95\% \text{ CI } [1.095, 3.436]; p = .015, 95\% \text{ CI } [0.201, 2.577]$; at left- and right-central sites (C3 and C4), it was further recorded that amplitudes for destination-correct neutral faces were more positive-going compared with those for new ones, $p = .018, 95\% \text{ CI } [0.169, 2.483]$, and $p = .007, 95\% \text{ CI } [0.331, 2.688]$; at all sites, amplitudes for destination-correct cooperative, neutral, and cheating faces were statistically indistinguishable. Therefore, we obtained destination-correct LPC for cooperative faces over the frontal region and for neutral faces over the frontal and central regions, while destination-correct LPC for cheating faces was restricted to F4.

600–1,000 ms (LPN). According to the ANOVA, the main effect of face type, $F(2.578, 74.756) = 11.970, p < .001, \eta_p^2 = 0.292, \epsilon = 0.859$, the two-way interaction of face type by region, $F(3.899, 113.069) =$

$11.240, p < .001, \eta_p^2 = 0.279, \epsilon = 0.433$, and the three-way interaction, $F(7.992, 231.759) = 2.756, p = .006, \eta_p^2 = 0.087, \epsilon = 0.444$, all reached statistical significance. As revealed by the post-hoc comparisons for the two-way interaction, over both parietal and occipital regions, new faces were with more positive-going amplitudes compared with destination-correct cooperative, neutral, and cheating faces, parietal: $p < .001, 95\% \text{ CI } [1.163, 3.953], p < .001, 95\% \text{ CI } [1.219, 3.466]$, and $p < .001, 95\% \text{ CI } [1.660, 4.320]$; occipital: $p < .001, 95\% \text{ CI } [1.117, 3.258], p < .001, 95\% \text{ CI } [1.351, 2.950]$, and $p < .001, 95\% \text{ CI } [1.507, 3.725]$. Furthermore, post-hoc comparisons for the three-way interaction indicated that the magnitudes of the effects for all the three behavior categories seemed to be largest at the medial-parietal (Pz) site, $p < .001, 95\% \text{ CI } [1.417, 4.617]; p < .001, 95\% \text{ CI } [1.786, 4.483]; p < .001, 95\% \text{ CI } [1.925, 5.032]$; amplitudes elicited by destination-correct cooperative, neutral, and cheating faces were statistically equivalent at all sites. These data suggest statistically equivalent destination-correct LPN over the parietal and occipital regions for cooperative, neutral, and cheating faces in the epoch of 600–1,000 ms.

1,000–1,400 ms (LPN/RFE). Face type interacted with region statistically significantly in this time segment, $F(4.244, 123.087) = 14.011, p < 0.001, \eta_p^2 = 0.376, \epsilon = 0.472$. Post-hoc comparisons for this interaction confirmed statistically significantly greater positivity for new versus destination-correct cheating faces over the parietal region, $p = .040, 95\% \text{ CI } [0.055, 3.318]$, and statistically significantly enhanced positivity for new versus destination-correct cooperative, neutral, and cheating faces over the occipital region, $p < .001, 95\% \text{ CI } [0.621, 2.619]; p = .005, 95\% \text{ CI } [0.365, 2.683]; p = .001, 95\% \text{ CI } [0.797, 3.581]$; amplitudes for destination-correct faces of the three face types did not differ from each other statistically significantly over all regions. Accordingly, the destination-correct RFE was not recorded for any behavior type; cooperative and neutral faces showed destination-correct LPN over the occipital region only, while cheating faces reported the LPN over both the parietal and occipital regions.

DESTINATION-INCORRECT OLD/NEW EFFECTS

To explore whether the four components of old/new effects (i.e., the FN400, LPC, LPN, and RFE) could all be recorded for destination-incorrect trials, and whether any component was modulated by behavior type, we conducted a $4 \times 4 \times 3$ (face type [cooperative, neutral, cheating, and new] \times region [frontal, central, parietal, and occipital] \times hemisphere [left, medial, and right]) repeated-measures ANOVA on mean amplitudes in each latency window comparing destination-incorrect trials of the three behavior types with the new trials. Only data relevant to the factor of face type was reported.

300–450 ms (FN400). There was only a statistically significant main effect of face type, $F(2.651, 76.889) = 3.172, p = .034, \eta_p^2 = 0.099, \epsilon = 0.884$. Post-hoc comparisons showed that amplitudes were more negative-going for new versus destination-incorrect cooperative and neutral faces, $p = .002, 95\% \text{ CI } [-1.593, -0.042]; p = .001, 95\% \text{ CI } [-1.581, -0.004]$; no other contrasts of the four levels of face type reached statistical significance. Based on these results, cooperative and neutral instead of cheating trials recorded destination-incorrect FN400.

450–600 ms (LPC). No effect relevant to face type reached statistical significance within this time segment, so destination-incorrect LPC was not obtained for all the three behavior types.

600–1,000 ms (LPN). The ANOVA recorded a main effect of face type, $F(2.695, 78.145) = 14.064$, $p < .001$, $\eta_p^2 = 0.327$, $\epsilon = 0.898$. Face type also interacted with region, $F(3.725, 108.035) = 3.388$, $p = .014$, $\eta_p^2 = 0.105$, $\epsilon = 0.414$. Post-hoc comparisons for the two-way interaction showed that amplitudes for new faces were more positive-going than those for destination-incorrect faces of all the three behavior types (cooperative, neutral, and cheating) over the parietal, $p < .001$, 95% CI [1.337, 3.809]; $p < .001$, 95% CI [1.041, 3.642]; $p < .001$, 95% CI [1.443, 4.200], and occipital regions, $p < .001$, 95% CI [1.043, 3.153]; $p < .001$, 95% CI [0.917, 2.866]; $p < .001$, 95% CI [1.009, 3.002]. Post-hoc comparisons also discovered that amplitudes for faces of all the three behavior types were comparable. The LPN was thus obtained over the parietal and occipital regions in destination-incorrect trials and was unmodulated by behavior type.

1,000–1,400 ms (LPN/RFE). There was a statistically significant two-way interaction of face type by region only, $F(4.263, 123.616) = 7.403$, $p < .001$, $\eta_p^2 = 0.203$, $\epsilon = 0.474$. Post-hoc comparisons for the interaction confirmed statistically significantly more positive-going amplitudes for new versus destination-incorrect cheating faces only over the occipital region, but not for new versus destination-incorrect cooperative or neutral faces, $p = .015$, 95% CI [0.228, 2.935]. In addition, amplitudes for the three kinds of destination-incorrect faces were not statistically significantly different from each other. Therefore, the RFE was not acquired in destination-incorrect faces, while the destination-incorrect LPN for cheating rather than cooperative and neutral faces lasted in this latency window over the occipital region.

ERP DIFFERENCE BETWEEN DESTINATION-CORRECT AND DESTINATION-INCORRECT ITEMS

To examine the difference in mean amplitudes between destination-correct and destination-incorrect trials, we conducted three $2 \times 4 \times 3$ (memory judgment [destination-correct and destination-incorrect] \times region [frontal, central, parietal, and occipital] \times hemisphere [left, medial, and right]) repeated-measures ANOVAs on mean amplitudes in each latency window for the cooperative, neutral, and cheating conditions, respectively. Only statistically significant effects regarding memory judgment are reported.

300–450 ms (FN400). No statistically significant effects were acquired for both the cooperative and neutral conditions in this latency window. For the cheating condition, there was a statistically significant interaction between memory judgment and region, $F(1.713, 49.665) = 3.824$, $p = .035$, $\eta_p^2 = 0.117$, $\epsilon = 0.571$, while post-hoc comparisons suggested no statistically significant differences in mean amplitudes for destination-correct and destination-incorrect faces across all the four regions. Consequently, the difference between destination-correct and destination-incorrect faces did not reach statistical significance in the time segment of the FN400.

450–600 ms (LPC). Regarding cooperative behaviors, no statistically significant effect was obtained. For neutral behaviors, a statistically

significant interaction of memory judgment by region, $F(1.530, 44.378) = 11.648$, $p < .001$, $\eta_p^2 = 0.287$, $\epsilon = 0.510$, and a significant three-way interaction were reported, $F(5.216, 151.256) = 2.921$, $p = .014$, $\eta_p^2 = 0.092$, $\epsilon = 0.885$. Post-hoc comparisons for the two-way interaction confirmed that destination-correct faces elicited more positive-going amplitudes than destination-incorrect faces over the frontal and central regions, $p = .004$, 95% CI [0.490, 2.350]; $p = .045$, 95% CI [0.021, 1.727]. Post-hoc comparisons for the three-way interaction continued to show that this difference was statistically significant at all the three electrodes (F3, Fz, and F4) over the frontal region, $p = .045$, 95% CI [0.023, 2.084]; $p = .003$, 95% CI [0.558, 2.489]; $p = .001$, 95% CI [0.788, 2.580], but only at the right-central electrode (C4) over the central region, $p = .011$, 95% CI [0.274, 1.914]. With regard to the cheating behaviors, only the two-way interaction between memory judgment and region reached statistical significance, $F(1.689, 48.969) = 7.344$, $p = .003$, $\eta_p^2 = 0.202$, $\epsilon = 0.563$. Post-hoc comparisons further indicated significantly larger positivity for destination-correct versus destination-incorrect faces over the frontal region, $p = .041$, 95% CI [0.048, 2.035]. These results suggested significant difference between destination-correct and destination-incorrect faces for the neutral and the cheating conditions over the frontal region in the latency window of the LPC.

600–1,000 ms (LPN). No effect appeared to be statistically significant for the cooperative condition. For the neutral condition, memory judgment interacted with region significantly, $F(1.647, 47.751) = 9.642$, $p = .001$, $\eta_p^2 = 0.250$, $\epsilon = 0.549$, and a statistically significant three-way interaction was reported as well, $F(4.102, 118.944) = 3.499$, $p = .009$, $\eta_p^2 = 0.108$, $\epsilon = 0.684$. Regarding the two-way interaction, post-hoc comparisons revealed a statistically significant difference between amplitudes for destination-correct faces and those for destination-incorrect faces over the frontal region, $p = .013$, 95% CI [0.301, 2.301]. Regarding the two-way interaction, post-hoc comparisons further demonstrated that such difference was recorded at the medial-frontal (Fz), right-frontal (F4), and right-central (C4) sites, $p = .015$, 95% CI [0.269, 2.337]; $p = .001$, 95% CI [0.831, 2.938]; $p = .013$, 95% CI [0.288, 2.281]. As for the cheating condition, only the two-way interaction of memory judgment by region reached statistical significance, $F(1.802, 52.244) = 9.391$, $p = .001$, $\eta_p^2 = 0.245$, $\epsilon = 0.601$. Yet, pairwise comparisons did not reveal any statistically significant differences between conditions. Thus, within 600–1,000 ms, the difference between destination-correct and destination-incorrect faces was recorded over the frontal and central regions for the neutral condition only.

1,000–1,400 ms (LPN/RFE). No effect was revealed for both the cooperative and the neutral conditions. A statistically significant two-way interaction was obtained for cheating faces between memory judgment and region, $F(1.778, 51.558) = 4.780$, $p = .015$, $\eta_p^2 = 0.142$, $\epsilon = 0.593$, but post-hoc comparisons failed to confirm a statistically significant difference between amplitudes elicited by faces with different memory judgments over any region. Consequently, no difference between destination-correct and destination-incorrect faces was observed in 1,000–1,400 ms for any of the three behavior circumstances.

DISCUSSION

Using a modified social dilemma game paradigm to explore destination memory in an ERP experiment, the current study explored the influence of cooperation and cheating on destination memory performance and old/new effects. The main goal of our study was to answer the following question: Would individuals remember the destinations of their positive behaviors better than the destinations of their negative behaviors? We assumed that if so, then the self-positivity bias could be confirmed as evidence for the self-reference process in destination memory. The behavioral performance of destination memory turned out to be modulated by cooperation and cheating in the way we expected: There was a mnemonic advantage for self-related cooperative versus cheating behaviors. For ERP data, we analyzed destination-correct and destination-incorrect old/new effects, and the waveform difference between destination-correct and destination-incorrect trials. For a brief summary, we recorded reliable FN400, LPC, and LPN, but not RFE for destination-correct old/new effects, and FN400 and LPN for destination-incorrect old/new effects. Specifically, there was a sign that destination-correct FN400 and LPC distributed more widely for cooperative and neutral behaviors than the effects for cheating behaviors. Destination-correct LPN was comparable across behavior types in terms of both magnitude and distribution in 600–1,000 ms, but distributed more widely in 1,000–1,400 ms for cheating behaviors compared with neutral and cooperative behaviors. For destination-incorrect old/new effects, the FN400 was only recorded under the cooperative and neutral conditions. The LPN was statistically indistinguishable during 600–1,000 ms across behavior conditions, but only lasted until 1,400 ms in the cheating condition. Furthermore, destination-correct and destination-incorrect trials were found to elicit amplitudes of different magnitude in the latency window of the LPC and LPN. In 450–600 ms, the difference in mean amplitudes between destination-correct and destination-incorrect trials was observed over the frontal region under the neutral and cheating conditions. In 600–1,000 ms, the difference was only acquired for neutral behaviors over the frontal and central regions. In the following sections, we will first discuss the self-positivity bias in the behavioral data, and then provide interpretations on how cooperation and cheating modulated different subprocesses of retrieval.

Modulation of Cooperation and Cheating on Destination Memory Performance Accords with the Self-Positivity Bias

How do we remember our own positive and negative behaviors towards specific targets? The current study revealed that individuals' destination memory for cooperative behaviors exceeded that for cheating behaviors in social exchange situations. Put differently, individuals remembered the targets of their cooperative behaviors better than those of their cheating behaviors. The data pattern is in accordance with the self-positivity bias: more accurate destination memory decisions (greater destination-correct and unbiased destination-correct rates) and less destination classification errors (lower destination-incorrect

rates) in the cooperative versus cheating condition. The self-positivity bias has been revealed in a wide range of cognitive tasks including attribution (Mezulis et al., 2004), word processing (Cai et al., 2016; Fields & Kuperberg, 2015; Li et al., 2016; Nowicka et al., 2018), and memory (Betz & Skowronski, 1997; Ritchie et al., 2017; Sedikides & Green, 2009; Zhang et al., 2018). Regarding memory, prior research indicated that people recollected unpleasant daily events related to themselves more poorly than pleasant events (Betz & Skowronski, 1997; Ritchie et al., 2017), and performed worse in source retrieval when negative versus positive trait adjectives were encoded in a self-reference manner (Zhang et al., 2018). The self-positivity bias is driven by strong motivation for self-protection (Leary, 2007). In the specific context of our study, participants might have tried to maintain a positive self-schema by avoiding building a connection of self-threatening cheating behaviors with the self. Taken together, the memory system has developed a robust self-protective strategy.

The self-positivity bias in our study confirms self-reference as a critical aspect of destination memory. Destination memory was proposed in contrast to source memory (El Haj et al., 2014; El Haj et al., 2016; El Haj & Miller, 2018; Gopie & MacLeod, 2009; Lindner et al., 2015; Marsh & Hicks, 2002). Researchers have managed to discriminate these two subtypes of episodic memory but found inconsistent results. Some confirmed lower destination memory versus source memory (El Haj et al., 2016; Fischer et al., 2015; Gopie & MacLeod, 2009) and attributed it to the higher degree of self-focus in destination memory, which might disrupt the connection between information and its external destinations and result in more fallible destination memory versus source memory (Gopie & MacLeod, 2009). Nevertheless, there was also evidence that source memory and destination memory were statistically indistinguishable (Lindner et al., 2015). We therefore tried to examine the self-reference aspect of destination memory by investigating whether information valence modulated destination memory in a self-positivity pattern, and found reliable behavioral evidence for this bias. Moreover, although we did not include a source memory task in the current study, previous studies applying the social dilemma game showed equivalent source memory for cooperation and cheating (Bell et al., 2010; Bell et al., 2016; Li & Nie, 2021; Volstorf et al., 2011). Consequently, source memory and destination memory are fundamentally distinct in their levels of self-reference.

Three Components of Old/New Effects Were Recorded in Destination Retrieval

As stated above, source memory and destination memory differ in terms of their levels of self-referential processing. Still, they are both parts of the episodic memory system and share common characteristics. We provided evidence for their resemblance via the retrieval-relevant old/new effects. Based on analyses of destination-correct and destination-incorrect old/new effects, we recorded three out of four classical old/new effects of source memory in the current destination memory task: the familiarity-driven FN400, the recollection-based LPC, and the LPN which represents evaluation and reconstruction

processes. Although no previous study has concerned old/new effects in destination memory, the current data pattern is not surprising. Essentially, destinations and sources in interpersonal interaction, and other source details like color, time, location, and so forth, are all certain kinds of contextual details. Both source memory and destination memory retrieve the studied items and their associated context (El Haj et al., 2016; Fischer et al., 2015; Gopie & MacLeod, 2009; Lindner et al., 2015; Shao et al., 2011). The FN400, reflects the familiarity process that assists the clarification of studied items. The LPC and LPN relate more closely to the retrieval of contextual details. For the RFE, which indexes the general post-retrieval monitoring process, it was not obtained in the current study. Yet, future studies should investigate whether its absence is because of the nature of destination memory or other aspects of the design of the current task. Furthermore, old/new effects in destination-correct items differed from those in destination-incorrect items in that the LPC did not exist in the destination-incorrect case. The LPC reflects the recollection of contextual details (Addante et al., 2012; Herron, 2017; Leynes et al., 2017; Minor & Herzmann, 2019; Nardini & Leynes, 2020; Ventura-Bort et al., 2020), so it is reasonable that when the retrieval of item-destination match is unsuccessful, the LPC decreases or even disappears.

Modulation of Cooperation and Cheating on Familiarity and Recollection in Destination Memory Reveals a Trend of Self-Positivity Bias

As revealed by destination-correct old/new effects, the FN400 and LPC are more widely distributed under the cooperative and neutral conditions than the cheating condition. For the FN400, it was recorded over frontal, central, and parietal regions in the cooperative and neutral conditions, but only over the frontal region in the cheating condition. The LPC appeared over the frontal region in the cooperative and the neutral conditions, while it was distributed restrictively at the right-frontal site in the cheating condition. In addition, the destination-incorrect FN400 was obtained only in the cooperative and neutral conditions. The wider distribution of the destination-correct FN400 and LPC for cooperative and neutral behaviors and the absence of destination-incorrect FN400 for cheating behaviors did reveal a sign of the self-positivity bias. It was likely that cooperation enhanced familiarity and recollection in comparison to cheating in destination retrieval to some degree. Referring to theoretical explanations of the FN400 and LPC (Addante et al., 2012; Hölting & Mecklinger, 2018; Leynes et al., 2017; Minor & Herzmann, 2019; Mollison & Curran, 2012; Nardini & Leynes, 2020; Proverbio et al., 2019; Ventura-Bort et al., 2020; Woroch & Gonsalves, 2010), these ERP results indicated that participants experienced a deeper sense of knowing when re-encountering faces that they had previously cooperated with, and were more likely to recollect their cooperative but not cheating behaviors in greater detail. However, it must be pointed out that the modulation of cooperation and cheating on the magnitude of old/new effects will serve as more compelling evidence for the self-positivity bias, and thus replication of the study is needed.

To date, our study is the first to explore old/new effects in destination memory for cooperation and cheating, but there was also some evidence for the self-positivity bias in previous ERP studies. Martínez-Galindo and Cansino (2017) applied a betting game to investigate old/new effects in an ERP experiment, in which participants played card games with unknown faces during encoding and decided whether they have won or lost the game when they interacted with given faces in the test. They recorded the FN400 and LPC for faces encoded under the winning rather than the losing context. Although these authors did not explain their results from the self-positivity perspective, the winning/losing context was the context where “I” won/lost, so their data might also be treated as ERP evidence for the self-positivity bias in memory.

Destination Memory for Cheating Versus Cooperation Demands Additional Reconstruction and Evaluation Processes

Regarding the LPN, which is an index of evaluation of retrieval outcome and reconstruction of encoding episodes (Barrick & Dillon, 2018; Leynes et al., 2013; Leynes et al., 2017; Leynes & Kakadia, 2013; Leynes & Nagovsky, 2016; Mecklinger et al., 2016; Nardini & Leynes, 2020; Sommer et al., 2018), it was unaffected by cooperation and cheating in both destination-correct and destination-incorrect items during 600–1,000 ms. Yet, during 1,000–1,400 ms, the LPN in destination-correct items distributed more widely for cheating versus cooperation since it appeared across the parietal and occipital regions in the cheating condition, but over the occipital region only in the cooperative and the neutral conditions. In the same time window for destination-incorrect items, the LPN was more lasting for cheating versus cooperation, since it was only recorded for cheating trials but not for cooperative and neutral ones. Altogether, the current data suggest that destination retrieval for cheating involves additional evaluation and reconstruction processes compared with cooperation. Leynes and Nagovsky (2016) proposed in their review that the LPN was more likely to be recorded in difficult source retrieval task. For instance, Strunk et al. (2017) reported the LPN for source memory in old instead of young participants, although the latter age group outperformed the former in the source memory task. In reference to this, we may infer cheating behaviors might be harder to be integrated with the self during encoding, and consequently the retrieval of self-related cheating behaviors might be more difficult than self-related cooperative ones, so that more cognitive resources would be needed to reconstruct encoding episodes and evaluate retrieval outcome. This could also be understood under the framework of the self-positivity bias, although more evidence is needed.

Diagnostic Recollection and Undiagnostic Recollection in Destination Memory for Cooperation and Cheating

Apart from old/new effects for destination-correct and destination-incorrect items, we also considered the waveform difference between destination-correct and destination-incorrect items. The waveform

difference was observed in 450–600 ms over the frontal region for the cheating condition and in 450–1,000 ms over the frontal and central regions for the neutral condition. Based on the source monitoring framework, successful retrieval of specified contextual details (details that are required to be retrieved by the experiment and thus diagnostic for the current task) relies on diagnostic recollection, while unsuccessful retrieval of specified contextual details results from undiagnostic recollection (Leynes & Mok, 2017; Mitchell & Johnson, 2009). Leynes and Mok showed that under self-reference encoding, diagnostic recollection and undiagnostic recollection did not differ in ERP amplitudes. They argued that this was because when items were encoded under the self-reference manner, they were more likely to be related to internal cues of generated by the participants. Thereby, during retrieval, participants tended to recollect these internal contextual details spontaneously together with task-specified external contextual details, even if those were not required to be reported and undiagnostic. In other words, instead of reflecting retrieval of no contextual information, destination-incorrect old/new effects might reflect retrieval of task-irrelevant contextual information. Our study shows further that information valence might alter the difference between diagnostic and undiagnostic recollection. In particular, under the cooperative condition, the difference between them disappeared. It is possible that positive behaviors are easier to be integrated with the self, so that during destination encoding, both destination-correct and destination-incorrect items generate various internal contextual cues that would later be recollected. By contrast, neutral and cheating behaviors might be less likely to be associated with the self, such that destination-incorrect items are accompanied with less internal cues. It would be interesting to investigate further how information valence modulates the difference between diagnostic and undiagnostic recollection under self-reference and nonself-reference encoding.

Limitations and Future Directions

First, we considered only destination memory for cooperation and cheating. It would be better if a source memory test was also included, allowing for a direct comparison between the two memory tasks. Second, we focused merely on moral valence of cooperative and cheating behaviors while neglecting the outcome valence of the game. In fact, the participants' cooperative behaviors can either lead to gain or loss for themselves. It would be more rigorous to explore whether participants remember their behaviors in a way that gives priority to moral valence or outcome valence. A prior study investigating source memory indicated that when remembering others' behaviors, participants recollected morality of behaviors better than their personal consequences (Bell et al., 2014). It is worth considering whether the rule still holds true when individuals try to remember their own behaviors. Another intriguing direction to consider is how individual differences may affect the self-positivity bias. Cheating behaviors might not be that self-threatening for individuals who frequently cheat others in life, and therefore, destination memory for cooperation of these individuals might not be greater than that for cheating. Last but not least, we cannot deny that controlling participants' behaviors in the social dilemma

game is not ideal. Yet, to study destination memory, we must ensure equal ratios of participants' cooperative, neutral, and cheating behaviors. Future studies can consider creative ways to study destination memory for social behaviors while allowing participants to act on their own will. It will also be a meaningful extension of the current destination memory research if future studies compare destination memory for behaviors consistent versus inconsistent with one's own will, since in real life, sometimes we may also need to follow others' instructions, and thus, our behaviors might not always be our independent choices.

CONCLUSION

In sum, this is the first study to our knowledge to investigate the sensitivity of destination memory and its retrieval-relevant neural correlates to information valence (social behavior valence). By combining the social dilemma game and the destination memory paradigm, we provided evidence for the self-positivity bias in destination memory from both behavioral and neural perspectives, although ERP evidence was less compelling compared with the behavior patterns. Behaviorally, destination memory for cooperation was more accurate than for cheating. Regarding neural patterns, cooperation enhances familiarity and recollection processes to a certain degree, while cheating relies on additional reconstruction and evaluation processes. These results suggest that cooperative behaviors are easier to be bonded with the self and thus easier to be subsequently retrieved, in line with the self-positivity bias.

ACKNOWLEDGEMENTS

This work is supported by the Projects of Humanities and Social Sciences, Ministry of Education (Grant No. 21YJA190005 and 17YJA190010), of Zhejiang Provincial Natural Science Foundation of China (Grant No. LY21C090002), Zhejiang Federation of Humanities and Social Sciences Circles (Grant No. 2021N78), National Natural Science Foundation of China (Grant No. 31300831), the Fundamental Research Funds for the Central Universities, and MOE Frontier Science Center for Brain Science & Brain-Machine Integration, Zhejiang University.

The authors declare that they have no conflicts of interest.

The experiment protocols were in line with the ethical standards laid by the Declaration of Helsinki and was approved by the Research Ethics Committee of Zhejiang University.

DATA AVAILABILITY

The datasets generated and/or analyzed during the current study, and the study materials applied, are available from the corresponding author upon reasonable request.

REFERENCES

- Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *NeuroImage*, 62, 439–450. doi: 10.1016/j.neuroimage.2012.04.031
- Bader, R., & Mecklinger, A. (2017). Separating event-related potential ef-

- fects for conceptual fluency and episodic familiarity. *Journal of Cognitive Neuroscience*, 29, 1402–1414. doi: 10.1162/jocn_a_01131
- Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition*, 107, 817–828. doi: 10.1016/j.cognition.2007.11.013
- Barrick, E. M., & Dillon, D. G. (2018). An ERP study of multidimensional source retrieval in depression. *Biological Psychology*, 132, 176–191. doi: 10.1016/j.biopsycho.2018.01.001
- Bell, R., Buchner, A., & Musch, J. (2010). Enhanced old-new recognition and source memory for faces of cooperators and defectors in a social-dilemma game. *Cognition*, 117, 261–275. doi: 10.1016/j.cognition.2010.08.020
- Bell, R., Sasse, J., Möller, M., Czernochowski, D., Mayr, S., & Buchner, A. (2016). Event-related potentials in response to cheating and cooperation in a social dilemma game. *Psychophysiology*, 53, 216–228. doi: 10.1111/psyp.12561
- Bell, R., Schain, C., & Echterhoff, G. (2014). How selfish is memory for cheaters? Evidence for moral and egoistic biases. *Cognition*, 132, 437–442. doi: 10.1016/j.cognition.2014.05.001
- Betz, A. L., & Skowronski, J. J. (1997). Self-events and other-events: Temporal dating and event memory. *Memory & Cognition*, 25, 701–714. doi: 10.3758/bf03211313
- Cai, H., Wu, L., Shi, Y., Gu, R., & Sedikides, C. (2016). Self-enhancement among westerners and easterners: A cultural neuroscience approach. *Social Cognitive and Affective Neuroscience*, 11, 1569–1578. doi: 10.1093/scan/nsw072
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates Inc.
- Cook K. S., Cheshire C., Rice E. R. W., Nakagawa S. (2013). Social exchange theory. In: DeLamater J., Ward A. (eds), *Handbook of social psychology* (pp. 57–101). Dordrecht: Springer.
- Cruse, D., & Wilding, E. L. (2009). Prefrontal cortex contributions to episodic retrieval monitoring and evaluation. *Neuropsychologia*, 47, 2779–2789. doi: 10.1016/j.neuropsychologia.2009.06.003
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 28, 923–938. doi: 10.3758/BF03209340
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Dickerson, B. C., & Eichenbaum, H. (2010). The episodic memory system: Neurocircuitry and disorders. *Neuropsychopharmacology*, 35, 86–104. doi: 10.1038/npp.2009.126
- El Haj, M. (2017). Stereotypes influence destination memory in normal aging. *Experimental Aging Research*, 43, 355–366. doi: 10.1080/0361073X.2017.1333821
- El Haj, M., Allain, P., & Kessels, R. P. C. (2014). The cognitive and neuroanatomical underpinnings of destination memory. *Translational Neuroscience*, 5, 147–151. doi: 10.2478/s13380-014-0219-5
- El Haj, M., Caillaud, M., Verny, C., Fasotti, L., & Allain, P. (2016). Destination and source memory in Huntington's disease. *Journal of Neuropsychology*, 10, 77–89. doi: 10.1111/jnp.12057
- El Haj, M., & Miller, R. (2018). Destination memory: The relationship between memory and social cognition. *Psychological Research*, 82, 1027–1038. doi: 10.1007/s00426-017-0891-5
- El Haj, M., Saloppé, X., & Nandrino, J. L. (2018). Destination memory and deception: When I lie to Barack Obama about the moon. *Psychological Research*, 82, 600–606. doi: 10.1007/s00426-016-0840-8
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. doi: 10.3758/BRM.41.4.1149
- Fields, E. C., & Kuperberg, G. R. (2015). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Social Cognitive and Affective Neuroscience*, 10, 1202–1209. doi: 10.1093/scan/nsv004
- Fischer, N. M., Schult, J. C., & Steffens, M. C. (2015). Source and destination memory in face-to-face interaction: A multinomial modeling approach. *Journal of Experimental Psychology: Applied*, 21, 195–204. doi: 10.1037/xap0000046
- Gopie, N., & MacLeod, C. M. (2009). Destination memory: Stop me if I've told you this before. *Psychological Science*, 20, 1492–1499. doi: 10.1111/j.1467-9280.2009.02472.x
- Hayama, H. R., Johnson, J. D., & Rugg, M. D. (2008). The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: Specific or non-specific? *Neuropsychologia*, 46, 1211–1223. doi: 10.1016/j.neuropsychologia.2007.11.021
- Hellerstedt, R., & Johansson, M. (2016). Competitive semantic memory retrieval: Temporal dynamics revealed by event-related potentials. *PLoS One*, 11, e150091. doi: 10.1371/journal.pone.0150091
- Herron, J. E. (2017). ERP evidence for the control of emotional memories during strategic retrieval. *Cognitive, Affective, & Behavioral Neuroscience*, 17, 737–753. doi: 10.3758/s13415-017-0509-9
- Höltje, G., & Mecklinger, A. (2018). Electrophysiological reward signals predict episodic memory for immediate and delayed positive feedback events. *Brain Research*, 1701, 64–74. doi: 10.1016/j.brainres.2018.07.011
- Iliadou, P., Kladi, A., Frantzidis, C. A., Gilou, S., Tepelena, I., Gialaouzidis, M., ... Moraitou, D. (2019). The pattern of mu rhythm modulation during emotional destination Memory: Comparison between mild cognitive impairment patients and healthy controls. *Journal of Alzheimer's Disease*, 71, 1201–1215. doi: 10.3233/JAD-190311
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344. doi: 10.1146/annurev.psych.58.110405.085658
- Leynes, P. A., Askin, B., & Landau, J. D. (2017). Visual perspective during remembering: ERP evidence of familiarity-based source monitoring. *Cortex*, 91, 157–168. doi: 10.1016/j.cortex.2016.12.003
- Leynes, P. A., Crawford, J. T., Radebaugh, A. M., & Taranto, E. (2013). Event-related potential evidence of accessing gender stereotypes to aid source monitoring. *Brain Research*, 1491, 176–187. doi: 10.1016/j.brainres.2012.11.013
- Leynes, P. A., & Kakadia, B. (2013). Variations in retrieval monitoring during action memory judgments: Evidence from event-related

- potentials (ERPs). *International Journal of Psychophysiology*, *87*, 189–199. doi: 10.1016/j.ijpsycho.2013.01.004
- Leynes, P. A., & Mok, B. A. (2017). Encoding focus alters diagnostic recollection and event-related potentials (ERPs). *Brain and Cognition*, *117*, 1–11. doi: 10.1016/j.bandc.2017.06.011
- Leynes, P. A., & Nagovsky, I. (2016). Influence of encoding focus and stereotypes on source monitoring event-related-potentials. *Brain Research*, *1630*, 171–182. doi: 10.1016/j.brainres.2015.11.017
- Li, M., & Nie, A. (2021). Do we prioritise memory for cheaters? Rebuttal evidence from old/new effects in episodic memory. *Journal of Cognitive Psychology*, *33*, 247–271. doi: 10.1080/20445911.2021.1894157
- Li, S., Xu, K., Xu, Q., Xia, R., Ren, D., & Zhou, A. (2016). Positive bias in self-appraisals from friend's perspective. *NeuroReport*, *27*, 694–698. doi: 10.1097/WNR.0000000000000599
- Lindner, I., Drouin, H., Tanguay, A. F. N., Stamenova, V., & Davidson, P. S. R. (2015). Source and destination memory: Two sides of the same coin? *Memory*, *23*, 563–576. doi: 10.1080/09658211.2014.911329
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. doi: 10.3389/fnhum.2014.00213
- MacKenzie, G., & Donaldson, D. I. (2007). Dissociating recollection from familiarity: Electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *NeuroImage*, *36*, 454–463. doi: 10.1016/j.neuroimage.2006.12.005
- Mahr, J. B., & Csibra, G. (2018). Why do we remember? The communicative function of episodic memory. *Behavioral and Brain Sciences*, *41*, 1–63. doi: 10.1017/S0140525X17000012
- Marsh, R. L., & Hicks, J. L. (2002). Comparisons of target output monitoring and source input monitoring. *Applied Cognitive Psychology*, *16*, 845–862. doi: 10.1002/acp.859
- Martínez-Galindo, J. G., & Cansino, S. (2017). Emotional context during encoding modulates recognition electrophysiological activity. *Experimental Brain Research*, *235*, 169–179. doi: 10.1007/s00221-016-4780-8
- Mecklinger, A., Rosburg, T., & Johansson, M. (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. *Neuroscience & Biobehavioral Reviews*, *68*, 621–638. doi: 10.1016/j.neubiorev.2016.06.024
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*, 711–747. doi: 10.1037/0033-2909.130.5.711
- Migo, E. M., Mayes, A. R., & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. *Consciousness and Cognition*, *21*, 1435–1455. doi: 10.1016/j.concog.2012.04.014
- Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source memory during encoding and retrieval. *Brain Research*, *1718*, 32–45. doi: 10.1016/j.brainres.2019.05.001
- Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*, 638–677. doi: 10.1037/a0015849
- Mollison, M. V., & Curran, T. (2012). Familiarity in source memory. *Neuropsychologia*, *50*, 2546–2565. doi: 10.1016/j.neuropsychologia.2012.06.027
- Nardini, C. L., & Leynes, P. A. (2020). Encoding focus does not affect recollection of action memories: Event related potential (ERP) and modeling evidence. *International Journal of Psychophysiology*, *147*, 9–17. doi: 10.1016/j.ijpsycho.2019.09.001
- Nie, A., Guo, C., Liang, J., & Shen, M. (2013). The effect of late posterior negativity in retrieving the color of Chinese characters. *Neuroscience Letters*, *534*, 223–227. doi: 10.1016/j.neulet.2012.11.043
- Nie, A., Xiao, Y., Liu, S., Zhu, X., & Zhang, D. (2019). Sensitivity of reality monitoring to fluency: Evidence from behavioral performance and event-related potential (ERP) old/new effects. *Medical Science Monitor*, *25*, 9490–9498. doi: 10.12659/MSM.917401
- Nie, A., & Yu, Y. (2021). External (vs. internal) facial features contribute most to repetition priming in facial recognition: ERP evidence. *Perceptual and Motor Skills*, *128*, 15–47. doi: 10.1177/0031512520957150
- Nowicka, M. M., Wójcik, M. J., Kotłowska, I., Bola, M., & Nowicka, A. (2018). The impact of self-esteem on the preferential processing of self-related information: Electrophysiological correlates of explicit self vs. other evaluation. *PLoS One*, *13*, e200604. doi: 10.1371/journal.pone.0200604
- Park, J. L., & Donaldson, D. I. (2019). Detecting the neural correlates of episodic memory with mobile EEG: Recollecting objects in the real world. *NeuroImage*, *193*, 1–9. doi: 10.1016/j.neuroimage.2019.03.013
- Proverbio, A. M., Vanutelli, M. E., & Viganò, S. (2019). Remembering faces: The effects of emotional valence and temporal recency. *Brain and Cognition*, *135*, 103584. doi: 10.1016/j.bandc.2019.103584
- Ritchie, T. D., Sedikides, C., & Skowronski, J. J. (2017). Does a person selectively recall the good or the bad from their personal past? It depends on the recall target and the person's favourability of self-views. *Memory*, *25*, 934–944. doi: 10.1080/09658211.2016.1233984
- Ross, R. S., Smolen, A., Curran, T., & Nyhus, E. (2018). MAO-A phenotype effects response sensitivity and the parietal old/new effect during recognition memory. *Frontiers in Human Neuroscience*, *12*, 53. doi: 10.3389/fnhum.2018.00053
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257. doi: 10.1016/j.tics.2007.04.004
- Sedikides, C., & Green, J. D. (2009). Memory as a self-protective mechanism. *Social and Personality Psychology Compass*, *3*, 1055–1068. doi: 10.1111/j.1751-9004.2009.00220.x
- Shao, Z., Du, M., Wang, Y., & Li, X. (2011). Does target memory without decision component still have a deeper processing level than source memory? *Journal of Psychological Science*, *34*, 337–342. doi: 10.16719/j.cnki.1671-6981.2011.02.030

- Suzuki, A., & Suga, S. (2010). Enhanced memory for the wolf in sheep's clothing. *Cognition*, *117*, 224–229. doi: 10.1016/j.cognition.2010.08.004
- Sommer, K., Vita, S., & De Pascalis, V. (2018). The late posterior negativity in episodic memory: A correlate of stimulus retrieval? *Biological Psychology*, *133*, 44–53. doi: 10.1016/j.biopsycho.2018.01.016
- Strunk, J., James, T., Arndt, J., & Duarte, A. (2017). Age-related changes in neural oscillations supporting context memory retrieval. *Cortex*, *91*, 40–55. doi: 10.1016/j.cortex.2017.01.020
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25. doi: 10.1146/annurev.psych.53.100901.135114
- Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2020). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*, *145*, 106606. doi: 10.1016/j.neuropsychologia.2017.12.015
- Volstorf, J., Rieskamp, J., & Stevens, J. R. (2011). The good, the bad, and the rare: Memory for partners in social interactions. *PLoS One*, *6*, e18945. doi: 10.1371/journal.pone.0018945
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*, 3–28. doi: 10.1007/bf00987006
- Woroch, B., & Gonsalves, B. D. (2010). Event-related potential correlates of item and source memory strength. *Brain Research*, *1317*, 180–191. doi: 10.1016/j.brainres.2009.12.074
- Ye, J., Nie, A., & Liu, S. (2019). How do word frequency and memory task influence directed forgetting: An ERP study. *International Journal of Psychophysiology*, *146*, 157–172. doi: 10.1016/j.ijpsycho.2019.10.005
- Yonelinas, A. P., Aly, M., Wang, W., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*, 1178–1194. doi: 10.1002/hipo.20864
- Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*, 236–244. doi: 10.1027/1618-3169/a00040

RECEIVED 22.11.2020 | ACCEPTED 27.12.2021