DEPARTMENT of PSYCHOLOGY

Memory and Inferences of Cheating Behavior: Age and Interindividual Differences in Trust Impact Decision Bias

Hannah Tucker

Master's Thesis (30 hp)
Spring 2023

Supervisor: Mikael Johansson

Acknowledgements

The creation of this thesis and all its components has been a challenging but hugely satisfying endeavour. Throughout this process I have relied on both academic and emotional support of academic community, from professors, fellow students, and researchers worldwide dedicated to nurturing the scholars of tomorrow.

I would like to thank my supervisor Mikael Johansson especially. The inception of this project can be traced back to a particularly captivating cognitive neuroscience lecture, underscoring the vital role passionate educators play in their students' success. He has consistently inspired me and offered invaluable input for my designs, all the while nurturing my confidence and independence. Truly, I could not have asked for a more suitable supervisor and mentor.

I would also like to express my appreciation for Tim Brabender, a fellow postgraduate student whom I have both commiserated and celebrated with during this process. Our scholarly volley has continuously provided fresh perspectives, crucial support and scaffolding, as well as a safe place to ask silly questions (especially about programming and coding).

Finally, I would like to thank my family, who have always encouraged me to pursue this eccentric career, and who have always believed in me. I cannot imagine navigating this path without the social and emotional safety net they have provided.

Abstract

The social contract theory hypothesis of a memory enhancement for cheaters remains a continued debate, but few studies have addressed the potential inferential role of such an enhancement. Moreover, little research has been conducted on interindividual characteristics and how these may impact cheating sensitivity. This study conducted a preference-by-association experiment which creates associative memory links between exposure stimuli, associate stimuli, and behavioral descriptors. Linear mixed modelling was used to investigate the interrelationships between decision bias, behavioral descriptors, dispositional trust, previous experience of cheating, age, and gender, in both low and high memory performance trials. There was no general memory enhancement for cheaters or cooperators, but there was a positive decision bias toward cooperators once age was controlled for. Moreover, decision bias in favor of both cheaters and cooperators was positively associated with age. There were limited gender differences in decision bias for exposure and associate trials. While previous experiences of cheating did not appear to influence generalization of bias, negative decision bias against cheater associates did emerge if dispositional trust was controlled for in high memory performance trials. More trusting participants were also more likely to trust cheater associated stimuli than low trust participants. Finally, while there was a significant negativity bias in old-new recognition, this effect was negligible in generalization. Theoretical implications for social contract theory and cognitive neuroscience are discussed, as well as practical applications.

Keywords: *Cheating sensitivity, social contract theory, integrative encoding, general trust scale, decision bias, generalization*

Background

Cooperation is often considered a hallmark of human behavior and forms a pillar for human society (Axelrod & Hamilton, 1981). The impact of cooperation is omnipresent in human history, and has been suggested to have promoted social learning, the development of complex language systems, enabled societal division of labor, and more (Henrich & Henrich, 2007). Mutualist collaboration yields a group-level benefit, but tends to have two drawbacks on an individual level; a cooperator would benefit more from being selfish, e.g. keeping hunted or gathered food to themselves rather than sharing, and the system is vulnerable to exploitation, i.e. an individual may be taken advantage of without anything in return (Ibrahim, 2022; Axelrod & Hamilton, 1981). Thus, cooperation alone is not inherently adaptive and the persistence of cooperative behavior among human and non-human animals remains puzzling (Trivers, 1971; Axelrod & Hamilton, 1981).

Social contract theory

Social contract theory is a prominent theory within evolutionary psychology which suggests that cooperative social acts are conditional on whether such an act is likely to be reciprocated, rather than pure altruism (Cosmides & Tooby, 1992). By this logic, there must be a rule which dictates when it is appropriate to cooperate (Axelrod & Hamilton, 1981; Cosmides & Tooby, 1992), which is referred to as a social contract (Cosmides & Tooby, 1997). Violation of a social contract is the failure to reciprocate cooperative social acts while benefitting from them, i.e. cheating. These decision rules are hypothesized to be operated by specialized neural mechanisms which automatically and efficiently guide behavior (Cosmides & Tooby, 1992; van Lier et al, 2013).

To avoid exploitation by cheaters, individuals must possess mechanisms to detect and remember cooperative and noncooperative behaviors of others. Specifically, there ought to be a memory enhancement for cheaters which facilitates recognizing an individual and recalling their behavioral history (Cosmides & Tooby, 1992). This mechanism is also hypothesized to provide inferences in novel situations to protect individuals against cheating (Cosmides & Tooby, 1992). Studies which have explored this hypothesized memory advantage have had mixed findings. One key study initially provided evidence for an item memory enhancement for stimuli which had been paired with cheating descriptions in comparison to trustworthy stimuli (Mealey et al, 1996). However, this has consistently failed to be replicated (Barclay &

Lalumiére, 2006; Mehl & Buchner, 2008; Bell et al, 2012; Kroneisen & Bell, 2013). For example, in a longitudinal study, Mehl and Buchner (2008) failed to find any significant difference in recognition between behavioral histories one week post-encoding. Moreover, a recent EEG study found no difference in item memory for cheaters and cooperators, nor any statistically significant difference in ERP activation during retrieval (Li & Nie, 2021), although activity during encoding was not measured.

These results dispute the existence of an explicit item-memory enhancement for cheaters but are not directly incompatible with the existence of a cheating-detection module as proposed by Cosmides and Tooby (1992). Recognition, i.e. distinguishing an individual as previously seen or not, is not enough to protect against future cheating behavior. Rather, it is important to be able to categorize a previously encountered individual as a cheater, not only to remember them, and this source memory enhancement has been repeatedly observed within the literature (Buchner et al., 2009; Bell & Buchner, 2012; Schaper et al, 2019). For example, Buchner et al (2009) found that while behavioral history appeared to have no impact on item memory, there was a higher accuracy in source memory for cheaters. If a cheater was recognized, participants were more likely to correctly identify their behavioral history, than for a cooperator (Buchner et al., 2009).

An alternate explanation which has been proposed as the Occam's razor account for the memory enhancement for cheaters is a general negativity bias in memory. In a study comparing the negative content of behavioral descriptors of cheating, Bell and Buchner (2011) found that when the cheating behavior was described less negative, there was no clear memory enhancement for cheaters. However, a general negativity bias has been labeled as inconsistent with much of the available research on memory, and as too simplistic. Most, non-clinical studies report a tendency of enhanced memory for positive, rather than negative events (Kroneisen et al, 2014). Moreover, a general negativity bias would likely be maladaptive, as it would promote avoidance over cooperation regardless of the other trust-relevant conditions of a social interaction (Rothermund et al, 2008) as well as incur potential costs to wellbeing (Grant et al, 2020). In addition to being maladaptive to the individual, a general negativity bias does not explain the pervasive tendency toward cooperation (Axelrod & Hamilton, 1981).

Thus, there is likely a more nuanced mechanism which stimulates avoidance, which either flexibly develops in line with environmental requirements or is situationally dynamic. One hypothesized condition for an enhanced memory of cheaters is threat-level. Individuals tend to attend to survival relevant material (Nairne et al, 2007; Bell & Buchner, 2012), thus memory of such events tends to be enhanced (Tay et al, 2019; Tse & Altarriba, 2010). Some researchers have highlighted that cheaters are a form of threat, and that the survival-relevance of their behavior is more important than the social-contract violation itself (Bell & Buchner, 2012). This is consistent with more general, parsimonious, flexible Survival Processing Effects compared to a specialized cheating module.

However, any potential ontogeny of cheating sensitivity remains relatively unexplored. For example, while there is a growing integration of Life History Theory into evolutionary psychology, which emphasizes that there may be important cognitive trade-offs throughout an individual's life (Del Giudice et al, 2015), there has been little examination of potential age differences in cheating sensitivity. Moreover, while participants' own behavioral history appears to have limited impact on memories of cheaters (Oda, 1997), personal history of exposure to cheating experiences, which likely affects cheating sensitivity (Fehr & Schurtenberger, 2018), has yet to be studied in the realm of a memory enhancement for cheaters.

In the same vein, there are limited studies on the impact of dispositional traits on cheating sensitivity. An exception for this is research conducted on clinical populations; for example, Niedtfeld and Kroneisen (2020) found that individuals with borderline personality disorder had impaired memory for cooperators compared to non-clinical controls, and treated cheaters as confirmation of negative social expectations. This was attributed to maladaptive cognitive schema and suggests that dispositional interindividual expectations influence memory of cheating behavior. Likewise, while some literature has addressed the influence of general gender stereotypes on memory enhancements for cheaters, such as the expectation of cooperation in women stimuli (Rule et al, 2012; Kroneisen & Bell, 2013), there tends to be little focus on the potential gender differences in cheating sensitivity itself (Oda, 1997).

Beyond the mere existence of a memory enhancement for cheaters, a key facet of social contract theory remains virtually untested, namely the inferential role of this enhancement in cheating-detection. Prior social interactions are likely to form expectations and influence perception of others (Fehr & Schurtenberger, 2018). Specifically, social contract theory emphasizes that previous experience with cheaters influences social decision-making in novel cooperative encounters (Cosmides & Tooby, 1997; Cosmides et al, 2005), i.e. social-contract interactions with individuals of which there is no trust-relevant information available. Most previous research on the memory enhancement for cheaters has focused on direct exposure to cheaters, either through behavioral experiments such as the Prisoner's Dilemma paradigm (e.g. Pärnamets et al, 2020 and Li & Nie, 2021), or longitudinal studies employing old-new recognition tasks (Buchner et al, 2009). Few studies have addressed how memory of cheaters may impact the memory of others; a specialized cognitive module must allow generalization to new encounters, or there is limited adaptive value (Nairne & Pandereida, 2016). While perceptions of trustworthiness have been widely studied and their influence on social decision-making (van't Wout & Sanfey, 2008), the link between cheating-detection, memory, and inference remains understudied.

Generalization and inference

A potential mnemonic mechanism which may serve the inferential role hypothesized by Cosmides and Tooby (1992) is generalization. Memory formation is constructive and reconstructive, meaning memories are not formed in isolation but rather overlapping and similar experiences are encoded together (Shohamy & Wagner, 2008). Integrative encoding of memory traces allows reactivation during inferential reasoning to provide information and guide behavior in a novel situation (Zeithamova et al, 2012; de Araujo Sanchez & Zeithamova, 2023). This allows for values, preferences, or perceptions to be generalized across memories. Retrospective generalization was observed by Wimmer and Shohamy (2012) in a seminal fMRI study; when associate stimuli and exposure stimuli are associated, and exposure stimuli are subsequently associated with a reward, both associate stimuli and exposure stimuli are preferred over unrewarded stimuli. This occurs despite associate stimuli never being explicitly associated with a reward, and without any conscious awareness of decision bias or logic strategy (Wimmer & Shohamy, 2012). Generalization of value through associated memories thus appears to guide decision-making, despite no explicit memory of the association or awareness (Wimmer & Shohamy, 2012; Klein et al, 2002). This process relies on pattern completion, localized in the hippocampus (Zeithamova et al, 2012; Wimmer & Shohamy, 2012). Interestingly, while the researchers did not find a general tendency to generalization, it was observed in participants whose fMRI data showed heightened activation in the hippocampus and striatum.

There is evidence to suggest that integrative encoding allows preparation for novel experiences by extracting inferences from previous knowledge, much like the theorized cheating detection module. In an EEG study conducted by Jiang et al (2020), researchers found that memory of past cognitive control demands appeared to guide cognitive control. Temporally, this cognitive control was deployed before task demands were evident, implying previous experiences shape behavioral cognition. Furthermore, this effect was present across similar, but not identical tasks; integrative encoding allowed effective, flexible retrieval of relevant information to guide behavior in novel situations. This reflects the adaptivity of the mechanism, including its automaticity and problem-solving function (Klein et al, 2002), and is consistent with the decision rules outlined in social contract theory (Cosmides & Tooby, 1992)

While integrative encoding and generalization appears to guide behavior in terms of rewards and problem-solving strategies, this retrospective mechanism has yet to be applied to social interactions. In fact, there is limited research on the role of hippocampal mnemonic processes in cooperative decision-making (Zeithamova et al, 2012; Pärnamets et al, 2020), despite the growing body of research suggesting memory impacts cooperative norms and conditional defection (Fehr & Shurtenberger, 2018). In conjecture, while there is ample research on the cheating detection module, few social contract studies have tried to identify the specialized neural mechanism itself, which would underlie a memory enhancement and its inferential qualities (Cosmides & Tooby, 1992). Finally, there remains a gap in the literature on why there are interindividual differences in hippocampal activation during incidental encoding and the accompanied variation in generalization. Thus, this study aimed to explore the role of integrative encoding in trust-relevant decision-making, specifically to investigate if this process is potentially more sensitive to cheaters and what interindividual differences impact generalization.

Pilot study

A pilot experiment using the preference-by-association paradigm (Wimmer & Shohamy, 2012) in a social decision-making context was conducted prior to this study (Tucker, 2023). The results showed that participants had a negative decision bias against individuals described as cheaters, supporting the memory enhancement for cheaters; participants were less likely to consider cheaters trustworthy. Moreover, there was no discernible difference between cooperators and neutral individuals, contrary to recent literature findings (Li & Nie, 2021). However, this study did not measure memory directly, thus these results could not be conclusively attributed to mnemonic processes. Additionally, this memory enhancement did not generalize to influence perceptions of associated individuals, nor did it consistently impact how willing participants were to cooperate with the individuals. However, participants were exposed to 75 pairs of faces and likely suffered from memory fatigue, which may have caused the lack of results for associated individuals (Tucker, 2023).

Present study

This study aimed to expand on the research in trust-relevant memory, interindividual differences in generalization, and social contract theory by employing Wimmer and Shohamy's (2012) preference by association paradigm in a remote study on the hypothesized cheating memory enhancement. This is one of the first studies testing if generalization via integrative encoding is especially sensitive to cheaters over cooperators and neutral individuals, thus also among the first studies to use the preference-by-association paradigm in a social context. Second, this study measured old-new discrimination to see if explicit recognition was crucial for generalization to occur or if familiarity with a lack of conscious awareness of associative memory was enough (Wimmer & Shohamy, 2012; de Araujo Sanchez & Zeithamova, 2023). Third, this study recorded individuals' previous experiences of cheating behavior and dispositional trust to explore how this may influence sensitivity to cheaters as well as act as an interindividual variable which may explain why generalization may or may not occur (Wimmer & Shohamy, 2012). Fourth, this study aimed to explore whether the alternate explanation of a negativity bias was a better predictor for memory performance and subsequent generalization (Bell & Buchner, 2010). Fifth, this is the first preference-by-association paradigm study conducted fully remote.

Based on previous research and the theoretical framework, behavioral descriptor type was predicted to (1a) influence old-new recognition, with the hypothesis that there (1b) would be a memory enhancement for cheaters (Mealey et al, 1996). Furthermore, it was predicted that behavioral descriptor type (2) would influence decision bias (Buchner et al, 2009; Bell et al, 2010; Li & Nie, 2021) for (2a) exposure stimuli (Tucker, 2023; Wimmer & Shohamy, 2012) and (3a) associate stimuli (Wimmer & Shohamy, 2012). Furthermore, following the expectations of social contract theory (Cosmides & Tooby. 1992), it was predicted that participants (2b) would consider exposure stimuli paired with cooperative descriptors more trustworthy than exposure stimuli paired with neutral descriptors (Tucker, 2023; Li & Nie, 2021), but (2c) consider exposure stimuli paired with cheating descriptors much less trustworthy than neutral exposure stimuli, i.e. there would be a negative decision bias against exposure stimuli paired with cheating descriptors (Tucker, 2023).

This effect was hypothesized to generalize to associate stimuli, i.e. (3b) associate stimuli paired with cooperative exposure stimuli would be considered more trustworthy than associate stimuli paired with neutral exposure stimuli. Furthermore, (3c) associate stimuli paired with cheating exposure stimuli were predicted to be considered much less trustworthy than those paired with neutral exposure stimuli, i.e. there would be a negative decision bias against associate stimuli paired with cheating exposure stimuli. Additionally, (4) dispositional trust in the form of General Trust Scale scores and (5) previous experience with instances of cheating (Memory of Cheating Scale) are hypothesized to influence sensitivity to cooperative and cheating behavior (Niedtfeld & Kroneisen, 2020; Kroneisen, 2018) and thus impact decision bias. In addition, (6) age and (7) gender was hypothesized to influence cheating sensitivity in decision bias. Finally, valence was hypothesized to (8a) influence old-new recognition, specifically that valence (8b) would be negatively associated with recognition. Moreover, valence was predicted to influence decision bias in (9a) exposure and (9b) associate stimuli (Bell & Buchner, 2010).

Method

Participants

Participants were recruited online via opportunity and snowball sampling. The study was advertised on social media platforms, such as Facebook and Reddit, and participant recruitment platforms, such as SurveySwap.io and SurveyCircle. Only English-proficient adults (18+ years of age) with no neurological or visual impairments were eligible, due to the rapid presentation of image stimuli being a potential irritant. Responses were collected between the 1st of July and the 14th of August 2023. Of the 116 participants who started the experiment, 114 participants completed all experimental phases. Of these 114, 50 were men, 58 were women, 5 were nonbinary and 1 preferred not to say. Ages ranged between 18 and 65 years of age, with a mean age of 30 (SD = 11.1) years.

Design

The study was comprised of an experiment using a self-paced preference-by-association paradigm and an old-new discrimination task, as well as a questionnaire regarding previous experience with cheating behavior and dispositional trust. The study was conducted online via Qualtrics (2023) and Pavlovia (2023). The experiment features a within-subjects design with behavioral descriptor (i.e. cheating, cooperative or neutral behavior) as the primary independent variable. Previous experience of cheating, dispositional trust, age and gender were also exploratory subject variables. The two dependent variables were decision bias in the preference-by-association paradigm and explicit memory performance on the old-new discrimination task. Decision bias was operationalized as the tendency to choose a certain category over another category (Wimmer & Shohamy, 2012), i.e. that an individual deemed a stimulus associated a certain behavioral descriptor as more trustworthy than a stimulus associated with another behavioral descriptor. Explicit memory performance was operationalized as the percentage accuracy in identifying true associated stimuli pairs un the old-new discrimination task (Mayes et al, 2007). Valence of behavioral descriptors was measured in a prior norming study and used as an alternate, continuous predictor for explicit memory performance and decision bias (Bell & Buchner, 2010).

Materials

The experiment was initially programmed using PsychoPy v2022.2.4 (Peirce et al, 2022) in Python and subsequently translated to JavaScript to host the study online. Image stimuli of 150 faces (79 female, 71 male) were sourced from the Chicago Face Database (CFD; Ma et al, 2015) and divided into two stimulus batteries, exposure stimuli and associate stimuli, each containing 75 images total. Each participant only viewed 24 images, 18 associate stimuli and 6 exposure stimuli. The CFD was selected for this study as it has extensive parametric data available; the 150 images were all rated between 3 and 4 on a 7-point Likert scale of perceived trustworthiness, to avoid systematic influence of facial trustworthiness on the results (Todorov et al, 2015). The stimuli only featured self-identified Caucasians, to avoid racial biases confounding the results. As this study was conducted remotely, the images sizes were .75 × .57 relative height units to scale the image consistently across different screen sizes. Sample images can be found in Appendix A.

The primary independent variable stimuli were 75 descriptors of behavior, divided equally into cheating, cooperative, or neutral behavior. Each descriptor contained 7 words for consistency (Rule et al, 2012). The cooperative and cheating descriptors featured compliances and violations of social contracts, respectively, while the neutral descriptors featured behavior irrelevant to social contracts. Sample behavioral descriptors can be seen in Table 1. A norming survey was conducted to obtain values for the valence of each statement. Respondents of this survey were asked to rate how positively or negatively they perceived each statement, ranging from -3 (“Very negative”) to +3 (“Very positive”). Respondents’ data was included if the respondent had completed more than 75% of the survey. Of these 67 respondents, 44 were female, 20 were male, 1 was non-binary/third gender, and 1 preferred not to say. The mean age was 25.39 years (SD = 6.94). See Appendix B for a full list of behavioral descriptors and their valence rating obtained from the norming study.

Table 1

Examples of statements from the three behavioral descriptor types and their valence.

Behavioral descriptor type	*Description*	*Mean valence*
Cooperative	Recycled bottles and cans after a party	1.785
	Returned a stray wallet to the police	1.985
	Cooperated with the company's energy-saving initiative	1.369
Cheating	Faked being sick to receive others sympathy	-1.554
	Used a fake ID to buy alcohol	-1.092
	Took silverware from a fine-dining restaurant	-.769
Neutral	Bleached their hair after their last breakup	-.2
	Wrote a book about well-known grandfather	1.185
	Ordered surf and turf at a restaurant	.293

The General Trust Scale (GTS; Yamagishi & Yamagishi, 1994; see Appendix C) was added as a post-experimental assessment of general tendencies to trust others and perceive others as trustworthy. The GTS consists of 6 statements, such as "Most people are honest" and "Most people will respond in kind to others", to which participants respond via a 5-point Likert scale where 1 means "Strongly disagree" and 5 means "Strongly agree".

In addition to the GTS, which assesses trust as a dispositional trait, a post-experiment questionnaire was created to address specific personal memories of the cheating behavior described in the experiment. This scale was created with ChatGPT (OpenAI, 2023), which condensed the 25 cheating descriptors into 10 binary questions, to measure an individual's previous exposure to cheating behavior. This provided an experiential counterpoint to the GTS, and featured questions such as "Have you ever been cheated on or betrayed in a romantic relationship?" and "Have you ever come across false advertising or misleading claims while purchasing products or services?". A question regarding whether an individual had been a victim of bullying, as a child or as an adult, was also included, as experiences with bullying are likely to influence trust, in turn affecting social decision-making and other social outcomes (Wolke et al, 2013; Matthews et al, 2022). The questions provided by ChatGPT were edited for clarity, and the full 11 question scale will henceforth be referred to as the Memory of Cheating Scale (MoCS; see Appendix D).

Procedure

Participants were given an outline of the studies purpose, without the mention of memory or cheating to avoid priming, as well as informed of their right to withdraw and their data protection rights according to GDPR. Participants were assigned a randomly generated ID which was automatically carried across the different parts of the study. Upon providing consent, participants were asked to state their gender and age. Participants were then redirected to Pavlovia for the experimental part of the study. The experiment was comprised of four phases: the association phase, the description phase, the decision phase and the explicit memory test phase. The experiment was self-paced and was comprised of 136 trials which took ca. 15-17 minutes to complete.

Association phase

In the association phase, each trial consisted of two images presented side-by-side (at -.4 and .4 horizontal relative height units) for 3s. After the image presentation, participants were then asked to indicate if the individuals shown were siblings, romantic partners, or co-workers using the left, upward, and right arrow-button, respectively. This was an indirect encoding task and intended to promote more elaborate processing of the stimuli through unitization, in which two components are encoded as a single whole (Diana et al, 2008) and allowing for more effective partial retrieval (Bastin et al, 2010, 2013). The task presented some difficulty as participants needed to evaluate both faces and the potential relationship between the two. Moreover, by only allowing participants to respond after the stimuli disappeared, participants were required to hold the images in memory. These aspects of the task promoted effortful encoding and engagement.

Each pair consisted of an exposure stimulus image and an associate stimulus image, randomly assigned. Participants were exposed to 18 pairs three times, with a total of 54 trials. Associate stimuli were presented on the left and exposure stimuli were presented on the right; this was kept consistent throughout the repetitions to promote encoding. Each exposure stimuli was paired with 3 associate stimuli, thus participants were exposed to 24 unique images. This pair configuration was selected in order to limit the number of faces a participant encoded, to avoid the exhaustion and memory fatigue observed in the pilot study (Tucker, 2023).

Description phase

In the next phase, exposure stimuli from the association phase were presented together with a behavioral description. After 2.5s, participants were prompted to read the descriptor carefully and rate the attractiveness of the individual on a 7-point Likert scale, with 1 meaning “Not attractive at all” and 7 meaning “Very attractive”. This task-delay was to promote association between the exposure stimuli and the description. The task itself was to ensure engagement. Once participants had made their rating, they pressed spacebar to move onto the next trial. This phase consisted of 6 trials, with 2 stimuli paired with a cheating descriptor, 2 stimuli paired with a cooperative descriptor, and 2 stimuli paired with a neutral descriptor. The stimuli orders and image-descriptor pairs were randomized to avoid appearance-behavior congruity (Cassidy & Gutchess, 2015) or incongruity (Suzuki & Suga, 2010). The descriptive phase provided the substrate for subsequent test phases, where exposure stimuli were assorted into cooperative, cheating and neutral categories, based on the description they had been paired with.

Decision phase

In this phase, participants were presented with pairs of within-battery images and asked to indicate who seemed more trustworthy, by pressing the left or the right key to indicate the left or the right stimulus, respectively (Wimmer & Shohamy, 2012). Participants were encouraged to respond with their immediate judgement, to promote speedy generalizations over accuracy (Klein et al, 2002). First, associate stimulus images were presented in pairs; cooperator associated stimuli and cheater associated stimuli images, neutral associated stimuli and cooperator associated stimuli images, cheater associated stimuli and neutral associated stimuli images, depending on which exposure stimuli they had been paired with in the initial association phase (associate stimuli paired with cooperator exposure stimuli were cooperator associate stimuli, and so forth). Second, exposure stimuli were presented in the same way. Decision bias for associate stimuli were tested first to avoid decision bias toward exposure stimuli influencing bias toward associate stimuli, or any potential priming and demand characteristics. This test phase consisted of 48 trials: 36 associate stimulus trials and 12 exposure trials.

Explicit memory test phase

The final experimental phase consisted of an old-new discrimination test. Participants were shown pairs of associate and exposure stimuli pairs and asked to indicate this pair was old, i.e. previously seen in this exact configuration, or recombinant, i.e. previously seen faces but in a novel pairing. Participants responded using the left and right arrow-button. Recombinant pairs were used rather than fully new pairs to disentangle recognition from mere familiarity (Mayes et al, 2007), i.e. that participants remember the associations rather than the stimuli in general. This phase comprised of 28 trials, 18 old pairs and 10 recombinant pairs, and was counterbalanced so recombinant pairs were presented before old pairs 50% of the trials, to ensure participants could not rely on inferring if a pair was old or new.

Post-experiment

Upon completing the final experimental phase, participant were redirected to a Qualtrics survey containing the GTS and the MoCS. Upon answering the 17 items, participants were debriefed on the purpose of the study and given resources to read more regarding social contract theory and generalization.

Statistical software and analysis

All statistical analyses were performed in RStudio (RStudio Team, 2021), using the *tidyverse* and *dplyr* packages (Wickham et al, 2019; 2021), the *zoo* package (Zeileis & Grothendieck, 2005), *ggpubr* and *rstatix* (Kassambara, 2020, 2021) packages. Linear mixed models (LMMs) were created using the *lme4* package (Bates et al, 2015), with *p*-values obtained using the *lmerTest* package (Kuznetsova et al, 2017). Binomial models were created using the base R package and the multinomial model was created using the *nnet* package (Venables & Ripley, 2002). Plots for all models were created using *ggplot2* (Wickham, 2016).

LMMs were selected as they are robust against non-normal distributions and can account for inter- and intraindividual variability (Baayen et al, 2008). A LMM was initially created with explicit memory performance as the outcome variable and behavioral descriptor type as the categorical predictor. Two general LMMs were then created with decision bias, for associate stimuli and exposure stimuli, as the outcome variable and behavioral descriptor type as the categorical predictor. Each LMM tested if decision bias was impacted differentially by descriptions of cooperative, cheating and neutral behavior, within low memory performance trials and high memory performance trials. Low memory performance trials were defined as decision bias trials where at least one individual was recognized whereas high memory performance trials were defined as decision bias trials where both individuals were recognized, based on the old-new discrimination test.

Moreover, LMMs were created to see if dispositional trust and previous experience with cheating behavior impacted the influence of behavioral descriptors on decision bias, by using the GTS and MoCS scores as predictors. Exploratory LMMs were also created to see if gender and age influenced decision bias. The neutral behavioral descriptor was computed as the intercept, or baseline, in all LMMs. This entails that both unstandardized estimates and standardized β for other levels of the behavioral descriptor and predictors are comparisons to the neutral behavioral descriptor. Due to the non-normality of the data, which violates the assumption of parametric statistical tests, and the uncertainty introduced by the incorporation of random and fixed effects in LMMs, no straight-out pre-hoc or ad-hoc power-analysis was made. Rather, the study aimed to collect around 15% more responses than the norm of the field, which tends to range between 70 (e.g. Oda, 1997, and Kroneisen & Bell, 2013) and 130 participants (e.g. Mealey et al, 1996 and Bell & Buchner, 2011), thus 115 responses was the goal for data collection.

As valence was connected to each descriptor item, binomial models were created using valence as a predictor and hit rate as the binary outcome in the old-new discrimination task and left-right choice as the binary outcome in the decision-bias tasks. Nagelkerke's R^2 was obtained from the *fmsb* package (Nakazawa, 2023). Throughout these analyses, participant IDs were used as a random effect.

Ethical considerations

This study was conducted in line with the Swedish Ethical Review Authority (2023) guidelines and the Act concerning the Ethical Review of Research Involving Humans (2003:460) within Swedish law. This study was anonymous and confidential, and no traceable, sensitive personal information was collected. Participants were informed on their right to withdraw from the study, and how to do so. This study did not intend to affect participants physically or mentally, nor did the study present any harm. The risks presented in this study were no greater than the risks of regular computer use, and individuals with visual or neurological impairments were excluded from participation.

Results

Behavioral descriptors were treated as three levels of one factor in the memory enhancement and decision bias LMMs. Model coefficient *b* is the unstandardized estimate of the dependent variable, i.e. the difference in the dependent variable with reference to the neutral intercept. Model coefficient *β* is the standardized estimate of the dependent variable, which represents the change in the dependent variable for a one-unit change in the predictor if all other variables is kept constant, i.e. the effect size. Alpha-level for all analyses is $p = .05$. Decision bias at 50% implies that there is no positive or negative decision bias, as the tendency to choose one stimulus over another is equal in a binary choice. Decision bias below 50% implies there is a negative decision bias for a stimulus. Decision bias above 50% implies there is a positive decision bias for a stimulus.

Is there a general memory enhancement for cheaters?

Table 2 contains coefficients for the LMM with behavioral descriptor as a categorical predictor and explicit memory performance as the outcome. Overall, the mean explicit memory performance was high for exposure stimuli, and no significant enhanced memory performance was found for cheaters (*M* = 93.6, SD = 20.3, *β* = 0) or cooperators (*M* = 94.4 SD = 20.5, *β* = .02) compared to neutral individuals in exposure trials (*M* = 93.6, SD = 21.3). There were minor, non-significant memory enhancements for cheater associates (*M* = 71.9, SD = 23.2, *β* = .04), or cooperator associates (*M* = 73.1, SD = 23.9, *β* = .06), compared to neutral associates (*M* = 69.8, SD = 25.9). Memory performance was lower in associate trials than exposure trials. These models accounted for 59.9% (conditional $R^2 = .5987$) and 43.8% (conditional $R^2 = .4383$) of the variance in explicit memory performance for exposure stimuli and associate stimuli, respectively, suggesting large interindividual variability.

Table 2

Summary of the General Memory Performance Model with Behavioral Descriptors as Predictors in both Exposure and Associate Trials.

		Coefficients		*95 % confidence intervals*				
Stimulus battery	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value*	*df*	*p-value*
Exposure trials	Intercept	93.59	0	89.84	97.34	48.879	203	<.001*
	Cheating	0	0	-3.36	3.36	0	232	1

	Cooperative	.85	.02	-2.51	4.22	.498	232	.619
Associate trials	Intercept	69.8	0	65.39	74.21	31.01	252	<.001*
	Cheating	2.14	.04	-2.55	6.82	.894	232	.373
	Cooperative	3.28	.06	-1.41	7.96	1.371	232	.172

Note * denotes significance at alpha-level .05. Significance for intercepts indicate deviation from 0. Unstandardized *b* is the mean difference in memory performance (%), in reference to the neutral baseline. Effect sizes are denoted by *β*.

Is there a decision bias against cheaters?

Table 3 presents model coefficients and statistical analyses for decision bias in both exposure and associate stimulus trials, in both low and high memory performance datasets. Overall, there was no general significant decision bias observed for either exposure or associate stimuli. In all trials, there was a similar mean decision bias for all three behavioral descriptors: mean decision bias tended to center around 50%, as denoted by *b*-values in Table 3, i.e. participants did not tend to consider one behavioral descriptor more trustworthy than others. In exposure trials, there was a minor, negligible difference in decision bias toward cooperators (b = 5.1, SD= 26.3, β = .09) and cheaters (b = 3.52, SD = 27.2, β = .06) in low memory performance trials but there was no difference in high memory performance trials (b = 1.66, SD = 27.1, and b = 1.74, SD = 28.4, respectively; β = .03).

A similar difference was found in low memory performance associate trials, with a small bias toward cooperator associate stimuli (b = 3.26, SD = 17.1, β = .09) while no bias for cheater associate stimuli (b = -.174, SD = 16.3, β = 0). In high memory performance trials, there was a minor negative decision bias against cheaters (b = -1.99, SD = 23, β = -.04). The exposure decision bias model explained .9% (conditional R^2 = .009) of the variance in the low memory performance dataset and .8% (conditional R^2 = .008) in the high memory performance dataset. The associate decision bias model accounted for .9% (conditional R^2 = .0089 of the variance in the low memory performance dataset and .23% (conditional R^2 = .0023) of the high memory performance dataset.

Table 3

Summary of the General Decision Bias Model with Behavioral Descriptors as Predictors for Exposure and Associate Stimuli in High and Low Memory Performance Trials.

			Coefficients		95% confidence intervals			
Stimulus battery	*Memory perf.*	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value (df=345)*	*p-value*
Exposure	Low	Intercept	47.2	0	42.38	52.02	19.162	<.001*
		Cheating	3.52	.06	-3.3	10.34	1.011	.313
		Cooperative	5.1	.09	-1.72	11.92	1.464	.144
							(df=343)	
	High	Intercept	48.91	0	43.82	54.01	18.791	<.001*
		Cheating	1.66	.03	-5.53	8.85	.452	.651
		Cooperative	1.74	.03	-5.46	8.94	.472	.637
							(df=345)	
Associate	Low	Intercept	48.93	0	45.91	51.95	31.672	<.001*
		Cheating	-0.17	0	4.45	4.1	-.08	.936
		Cooperative	3.26	.09	-1.01	3.26	1.493	.136
							(df=341)	
	High	Intercept	50.56	0	46.27	54.85	23.07	<.001*
		Cheating	-1.99	-.04	-8.05	4.08	-.642	.522
		Cooperative	.69	.01	-5.38	6.77	.223	.823

Note * denotes significance at alpha-level .05. Significance for intercepts indicate deviation from 0. Unstandardized *b* refers to the mean decision bias in percent, in reference to the neutral intercept. Effect sizes are denoted by *β*.

Dispositional trust

There was no significant decision bias observed in exposure stimuli with the inclusion of GTS score in the LMM, for either the low or high dataset. See Appendix E for full model coefficients and statistical analyses, including confidence intervals and *p*-values. There was a non-significant, general decision bias against cheaters (*b* = -7, SD = 16.4, *β* = -.13) and cooperators (*b* = -12.8, SD = 16.4, *β* =-.23) compared to the neutral baseline (*b* = 55.7, SD = 11.6) in the low memory performance dataset. A similar, non-significant trend was found in the high memory performance data, with a general decision bias against cheaters (*b* = -14.1, SD = 17.3, *β* = -.26) and cooperators (*b* = -19, SD = 17.3, *β* = -.32) compared to neutral (*b* = 60.6, SD = 12.2). With each unit increase on GTS score, there was a non-significant mean increase of 3% positive decision bias toward cheaters (SD = 4.8 , *β* =.19) and 5% toward cooperators (SD = 4.8, β =.31) in the low memory performance, while there was a non-significant 4.9% and 5.9% increase in positive decision bias toward cheaters (SD = 5.1 , β = .28) and cooperators (SD = 5.1, β = .34), respectively, in the high memory performance dataset. The GTS model for exposure stimuli explained .74% (conditional R^2 = .0074) of the variance in the low memory performance dataset and .49% (conditional R^2 = .0049) in the high memory performance dataset.

GTS score model coefficients and statistical analyses, including confidence intervals and *p*-values, for associate stimuli in both low and high memory performance trials can be seen in Table 4. There was a non-significant decision bias against cheater associated stimuli (*b* = -10.39, SD = 10.2, β = -.29) and in favor of cooperator associated stimuli (*b* = 5.77, SD = 10.3, β = .16) compared to neutral baseline (*b* = 50.94, SD = 7.3) in the low memory performance trials. GTS score had a negligible main effect (*b* = -.44, SD = 2.1, β = -.02) in this dataset, and no interaction effect with cheater associated stimuli (*b* = 2.7, SD = 3, β = .26) or cooperator associated stimuli (*b* = -.87, SD = 3, β = -.09). This model explained 1.6% of the variance in the low memory dataset (conditional R^2 = .0159).

However, by controlling for GTS score in the high memory performance dataset, a significant negative decision bias against cheater associates (*b* = -33.52, SD = 14.4, β = -.68, *p* = .02) compared to neutral baseline (*b* = 68.52, SD = 10.2) was observed. Decision bias toward cooperator associates was non-significant (*b* = -21, SD = 14.4, β = -.42). Furthermore, while there was no significant interaction effect between decision bias towards cooperator associates and GTS score (*b* = 6.39, SD = 4.2, β = .44), there was an interaction between decision bias towards cheater associates and GTS score (*b* = 8.77, SD = 14.4, β = .61, *p* = .038). For each unit of increase in GTS score, there was an 8.77% increase in decision bias toward cheater associates, suggesting that while there was an overall negative decision bias against cheater associates, this was weaker in participants with higher dispositional trust. However, this model only explained 2.19% of the variance in the dataset (conditional R^2 = .0219). See Figure 1 for a linear model graph indicating the relationship between decision bias, behavioral descriptor and GTS score. See Table 4 for GTS score model coefficients for all associate stimuli LMMs.

Table 4

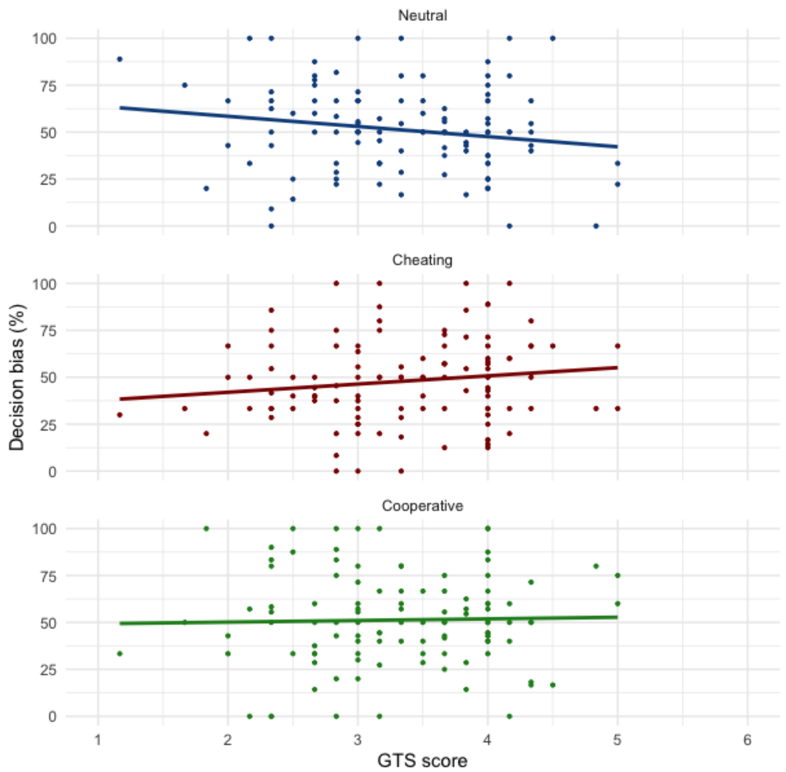
Summary of the Decision Bias Model with Behavioral Descriptors and GTS Score as Predictors for Associate Stimuli in High Memory Performance Trials.

	Coefficients		*95% confidence intervals*			
Behavioral descriptor	*b*	*β*	*lb*	*ub*	*t-value (df=323)*	*p-value*
Intercept (Neutral)	68.52	0	48.74	88.29	6.73	<.001*
Cheating	-33.52	-.68	-61.48	-5.56	-2.38	.02*
Cooperative	-21.02	-.42	-49.03	6.99	-1.41	.145
GTS score	-5.12	-.16	-10.92	.68	1.71	.087
Cheating × GTS score	8.77	.61	0.56	16.98	2.15	.038*
Cooperative × GTS score	6.39	.44	-1.84	14.62	1.43	.132

Note * denotes significance at alpha-level .05. Significance for intercepts indicate deviation from 0. Unstandardized *b* refers to the mean decision bias in percent, in reference to the neutral intercept. Effect sizes are denoted by *β*.

Figure 1

Linear Models Representing the Interaction Effect of GTS Score and Behavioral Descriptor on Decision Bias in High Memory Performance Associate Stimuli Trials.



Previous cheating experiences

No significant decision bias toward cheaters (*b* = 1.93, SD = 8.18, β = .03) or cooperators (*b* = 11.53, SD = 8.18, β = .21) emerged with the of inclusion MoCS score in the model in the low memory performance exposure trials. Non-significance persisted in the high memory performance exposure trials (*b* = 2.06, SD = 8.63, β = .03 for cheaters, and *b* = 7.4, SD = 8.64, β = .13 for cooperators). There was no main effect of MoCS score for stimulus exposure trials, in either the low (*b* = .39, SD = .97, β = .04) or high memory performance (*b* = .56, SD = 1.02, β = .05) dataset. There was no significant main effect of MoCS score, and no significant interactions; a fully detailed MoCS model summary for both low and high memory performance exposure trials can be found in Appendix F. The MoCS model explained .89% (conditional R^2 = .0089) of the variance in the low memory performance dataset and .31% (conditional R^2 = .0031) in the high memory performance dataset.

Likewise, no significant decision bias was found for cheater associates (*b* = 1.84, SD = 5.11, β = .03) or cooperator associates (*b* = -1.35, SD = 5.11, β = -.04) in the low memory performance trials. This pattern was also observed in the high memory performance associate trials (*b* = -4.68, SD = 7.21, β = -.09 and *b* = -1.15, SD = 7.23, β = -.02, for cheater associates and cooperator associates, respectively). As with the exposure trial models, MoCS did not affect decision bias in either associate trial model. The MoCS score models explained 1.93% of the variance in the low memory performance dataset and .91% of the high memory performance dataset of associate trials. See Appendix F for all model coefficients and statistical analyses.

Age

Age was the only variable which appeared to impact decision bias for specifically exposure trials. Table 5 contains full model coefficients and statistical analyses for the exposure models with age and behavioral descriptor as a predictor. A general decision bias emerged with inclusion of age as a predictor but remained barely non-significant in both the low (*b* = -17.71, SD = 10.12, β = -.32, *p* = .083) and high memory performance (*b* = -18.35, SD =10.75, β = -.31, *p* = .089) dataset. There were significant interactions between age and behavioral descriptor in the low memory performance dataset. With each year of age increased decision bias toward cheaters with .69% (SD = .32, β = .41, *p* = .03) and toward cooperators with .67% (SD = .32, β = .39, *p* = .036), and consequently decision bias toward neutral individuals reduced (*b* = -.44, SD =.32, β = -.19). This effect persisted in the high memory performance dataset but was only statistically significant for cooperators (*b* = .77, SD =.33, *β* = .43, *p* = .024). The age model explained 2.21% (conditional R^2 = .0221) of the low memory performance dataset's variance and 1.81% (conditional R^2 = .0181) of the variance in the high memory performance. Figure 2 and Figure 3 feature linear model graphs depicting the change in exposure stimuli decision bias with years of age according to behavioral descriptors for the low memory performance dataset and the high memory performance dataset, respectively.

Table 5

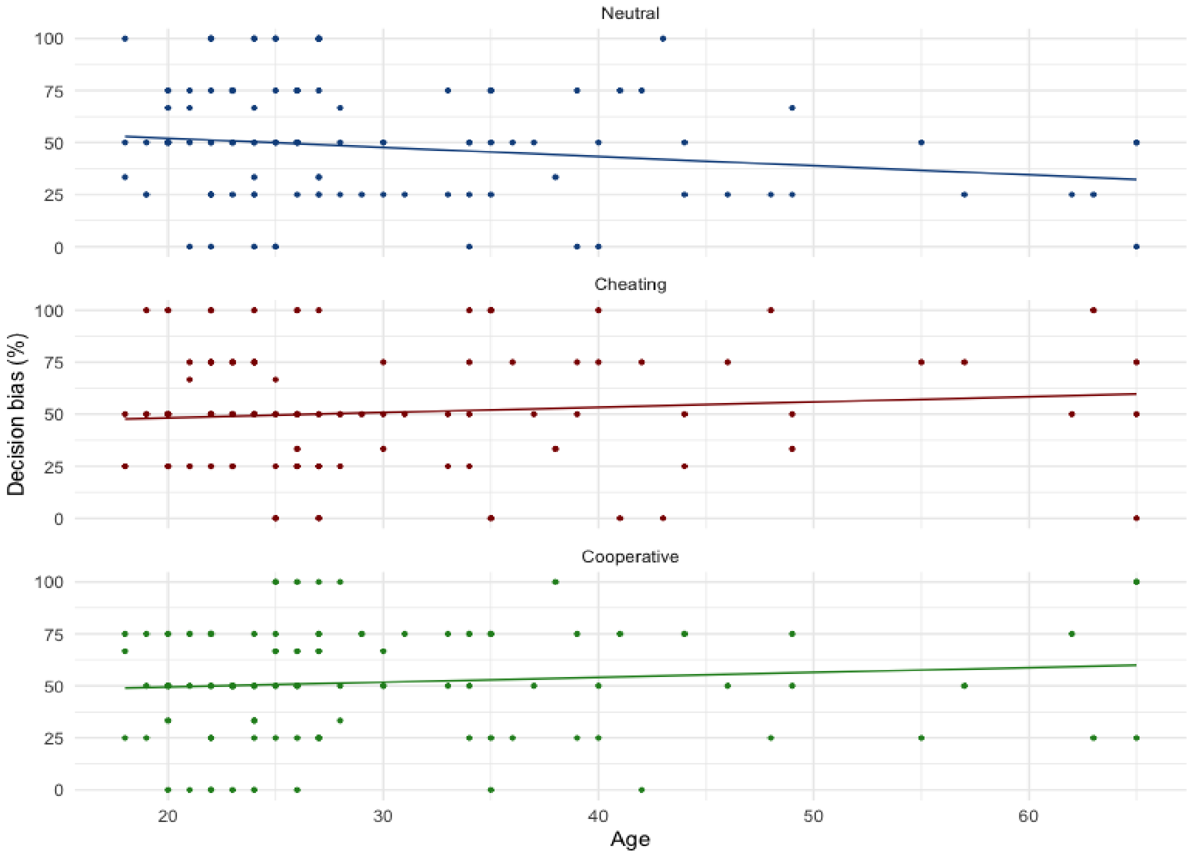
Summary of the Decision Bias Model with Behavioral Descriptors and Age as Predictors for Exposure Stimuli in Low and High Memory Performance Trials.

		Coefficients		*Confidence intervals*			
Memory perf.	*Behavioral descriptor*	*b*	*β*	*95% CI lb*	*95% CI ub*	*t-value (df=327)*	*p-value*
Low	Intercept (neutral)	60.87	0	46.86	74.87	8.465	<.001*
	Cheating	-17.71	-.32	-61.48	-5.56	-1.741	.083
	Cooperative	-16.1	-.29	-35.91	3.71	-1.583	.114
	Age	-.44	-.18	-.88	0	-1.947	.052
	Cheating × Age	.69	.41	.07	1.32	2.173	.03*
	Cooperative × Age	.67	.39	.05	1.3	2.107	.036*
						t-value (df=325)	
High	Intercept (neutral)	63.56	0	48.7	78.41	8.335	<.001*
	Cheating	-18.35	-.31	-39.29	-2.58	-1.707	.089
	Cooperative	-22.52	-0.38	-43.49	-1.56	-2.092	.037*
	Age	-.47	-.19	-.93	0	-1.96	.051
	Cheating × Age	.65	0.36	-0.01	1.31	1.925	.055
	Cooperative × Age	.77	0.43	0.11	1.43	2.276	.024*

Note * denotes significance at alpha-level .05. Significance for intercept indicate deviation from 0. Unstandardized *b* refers to the mean decision bias in percent, in reference to the neutral intercept. Age-coefficients are the change in decision bias estimate per unit increase in years of age. Effect sizes are denoted by *β*.

Figure 2

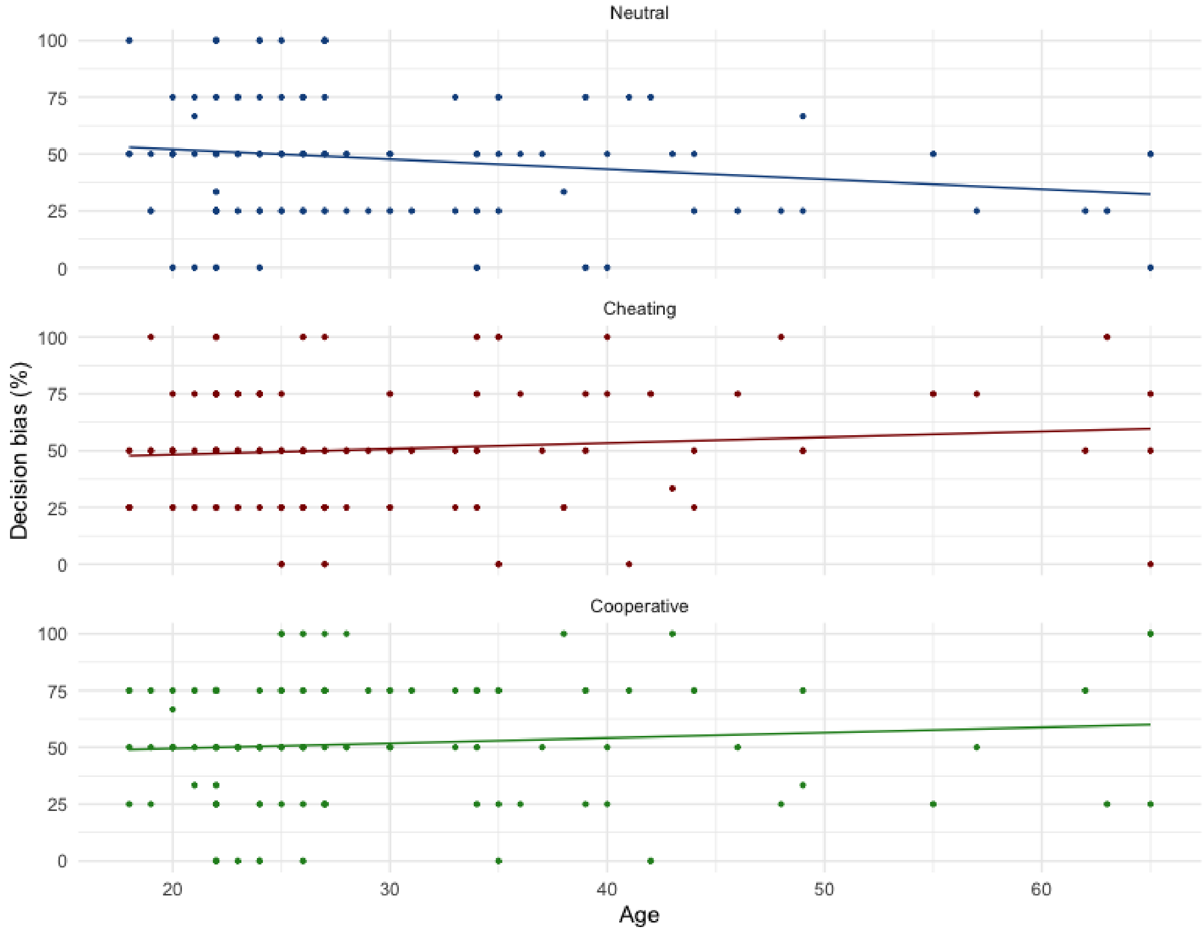
Linear Models Depicting Interaction Between Age and Behavioral Descriptor on Decision Bias in Low Memory Performance Exposure Trials.



Contrary to exposure stimulus trials, controlling for variance in age did not reveal a significant, general decision bias for associate cheaters (*b* = .03, SD =.2, β = 0) or associate cooperators (*b* = 5.41, SD =.2 , β = .15) in the low memory performance data. This was echoed in the high memory performance dataset (*b* = -1.44, SD = 9.09, β = -.03 and *b* = 2.92, SD =9.12 β = .06, for cheater associates and cooperator associates respectively). There were also no interaction effects between age and behavioral descriptors, in the low or high memory performance trials. See Appendix G for all model coefficients and statistical analyses for the age model on associate trials, including confidence intervals and *p*-values. These models only explained 1.21% (conditional R^2 = .0121) and .86% (conditional R^2 = .0086) of the low and high memory performance datasets, respectively.

Figure 3

Linear Models Depicting Interaction Between Age and Behavioral Descriptor on Decision Bias in High Memory Performance Exposure Trials.



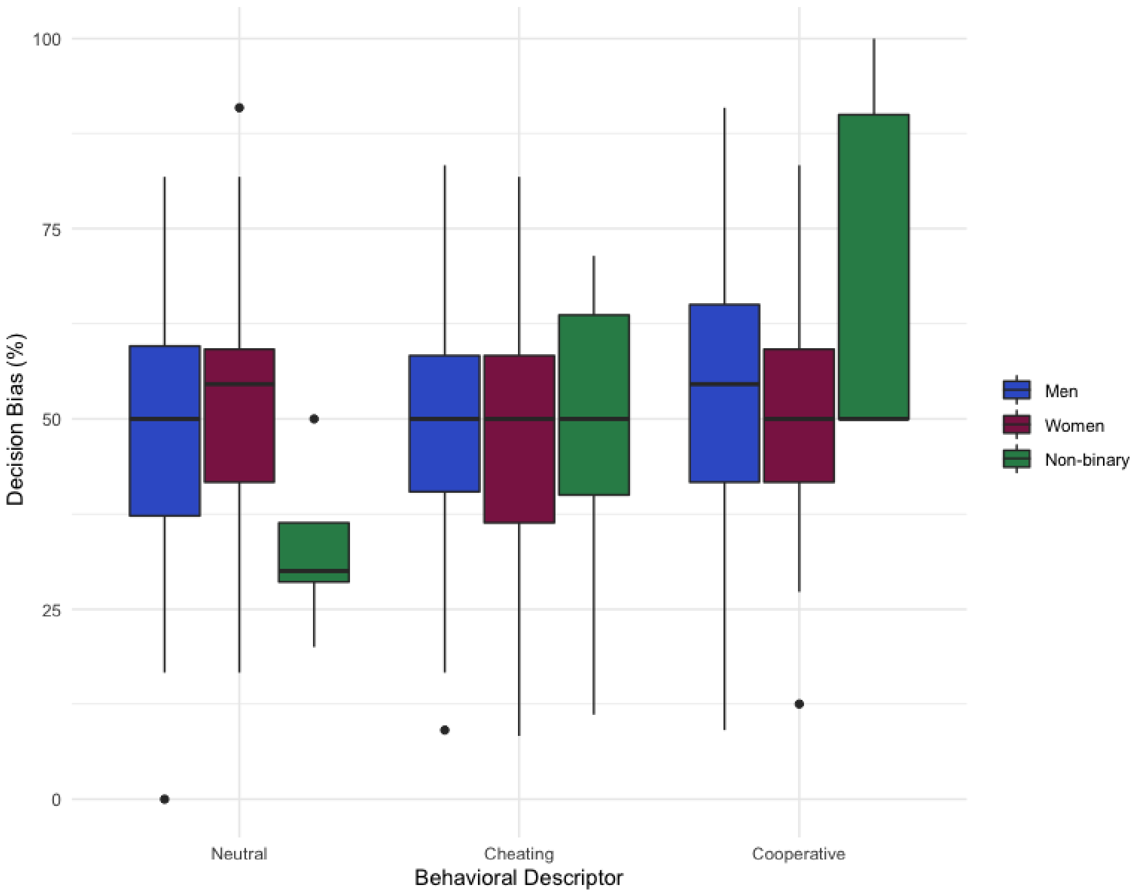
Note Only decision bias toward cooperators maintained significance at *p*<.05 in high memory performance trials.

Gender

The LMM with gender and behavioral descriptor as predictors did not show any significant decision bias in exposure trials, in either low or high memory performance trials. There were non-significant differences in decision toward cheaters between men (*b* = -.83, SD = 5.35, β = -.01), women (*b* = 7.2, SD =7.39, β = .1) and non-binary participants (*b* = .83, SD =17.75, β = 0) in the low memory performance exposure trial data. There was also no significant difference in decision bias towards cooperators between men (*b* = 1.33, SD = 5.35, β = .02), women (*b* = 3.1, SD =7.39, β = .05) and non-binary participants (*b* = 13.67, SD = 17.75, β = .06). Similarly, any differences in decision bias toward cheaters between men (*b* = -2.36, SD = 5.66, β = -.04), women (*b* = 5.85, SD = 28.45, β = .08), and non-binary participants (b = 5.69, SD = 18.69, β = .02) were non-significant in the high-memory performance trials. Men (b = -1.69, SD = 5.66, β = -.03), women (b = 1.21, SD = 7.8, β = .02), and non-binary participants (b = 23.36, SD = 18.69, β = .1) did not significantly differ in decision bias toward cooperators either, although there was greater variation. Full coefficient table and statistical analyses, including confidence intervals and *p*-values, can be found in Appendix H. These models explained 1.37% % (conditional R^2 = .0137) and 1.23% % (conditional R^2 = .0123) of the variance in the low memory performance data and the high memory performance data, respectively.

Figure 4

Boxplot depicting the interaction between gender and behavioral descriptor on decision bias in low memory performance associate stimulus trials.



Note As only one person preferred not to disclose their gender, they were omitted from the graph but included in the analysis. Single dots illustrate influential outliers. A legend for gender identities can be seen within the graph.

Figure 4 shows a box-and-whiskers plot illustrating decision bias by gender in low memory performance associate trials. Like exposure trials, gender differences in decision bias toward cheater associates between men (*b* = -.29, SD = 3.31, *β* = -.01) and women (*b* = -3.21, SD = 4.57, *β* = -.07) were negligible and non-significant in the low performance associate trials. Decision bias toward cooperator associates also did not differ significantly between men (*b* = 3.58, SD = 3.31, *β* = .1) and women (*b* = -3.89, SD = 4.57, *β* = -.09) in these trials. However, non-binary participants exhibited a positive bias toward cooperator associates (*b* = 14.54, SD = 10.96, *β* = .11) which was significant (*p* = .004); in conjunction with this, non-binary participants, also exhibited a significant, negative decision bias toward neutral individuals (*b* = -15.79, SD = 3.23, *β* = -.2, *p* = .042). It is important to note that this sample only comprised of 5 non-binary participants, thus significance is likely due to influential outliers.

Moreover, this significance did not persist in high memory performance trials, where decision bias toward cooperator associates was not significantly different from neutral descriptors for non-binary participants (*b* = 24.26, SD = 15.51, *β* = .13), men (*b* = 2.76, SD = 4.74, *β* = .56) and women (*b* = -6.9, SD = 6.5, *β* = -.11). There was also no significant decision bias toward cheater associates for either men (*b* = -.54, SD = 4.72, *β* = -.01), women (*b* = -6.26, SD = 6.49, *β* = -.11), or non-binary participants (*b* = -7.47, SD = 15.5, *β* = -.04). See Appendix H for a table of model coefficients and full statistical analyses.

Table 6 shows the model coefficients and full statistical analyses, including confidence levels, for decision bias according to gender and behavioral descriptor in low performance associate trials. The gender model accounted for 4.49% (conditional $R^2 = .045$) of the variance in the low memory performance associate trials and 3.25% of the variance in high memory performance associate trials (conditional $R^2 = .033$).

Table 6

Confidence intervals and *p*-values for age and the low memory performance associate stimulus and gender model coefficients.

		Coefficients		*Confidence intervals*			
Gender	*Behavioral descriptor*	*b*	*β*	*95%CI lb*	*95%CI ub*	*t-value (df=321)*	*p-value*
Men	Intercept	48.78	0	44.27	53.29	20.868	<.001*
	Cheating	-.29	-.01	-6.67	6.09	-.087	.93
	Cooperative	3.58	.1	-2.8	9.96	1.084	.279
Women	Neutral	2.57	.08	-3.66	8.8	.796	.427
	Cheating	-3.21	-.07	-12.02	5.61	-.702	.483

	Cooperative	-3.89	-.09	-12.7	4.93	-.851	.395
Non-binary	Neutral	-15.79	-.02	-30.75	-.83	-2.37	.042*
	Cheating	14.54	.11	-6.62	35.7	1.326	.186
	Cooperative	31.43	.23	10.27	52.59	2.867	.004*

Note * denotes significance at alpha-level .05. Significance for intercept indicate deviation from 0. Unstandardized *b* refers to the mean decision bias in percent, in reference to the neutral intercept. Effect sizes are denoted by *β*.

Is there a negativity bias?

A binomial regression model was conducted on the correct trials in the old-new discrimination test and mean valence of associated descriptors. The results indicate that mean valence had a negative influence on the probability of a correct trial (β = -.13, p <.001; see Table 7 and Figure 5), suggesting descriptors with a higher, positive valence had a reduced probability of a correct trial. A multinomial regression model was conducted to further explore the influence of valence on signal detection and found that valence had a positive effect on hits (β =1.059, p < .001), i.e. pairs of individuals associated with higher statements were recognized more, a negative effect on false alarms (β = -.173, p < .001), i.e. reduced interference from recombined pairs. However, there also appeared to be a positive effect on misses (β = .587, p < .001), suggesting pairs associated with higher valence statements were also falsely rejected more often. However, valence only explained 1.1% of the variation in hit rate (Nagelkerke's R^2 = .011). There was a weak, negative effect of valence on decision bias for both exposure stimulus trials (β = -.0136, p < .001) and associate stimulus trials (β = -.0395, p < .001). However, this appeared to only explain .001% (Nagelkerke's R^2 = .001) of the variance for exposure stimulus trials and .01% of the variance for associate stimulus trials (Nagelkerke's R^2 = .001).

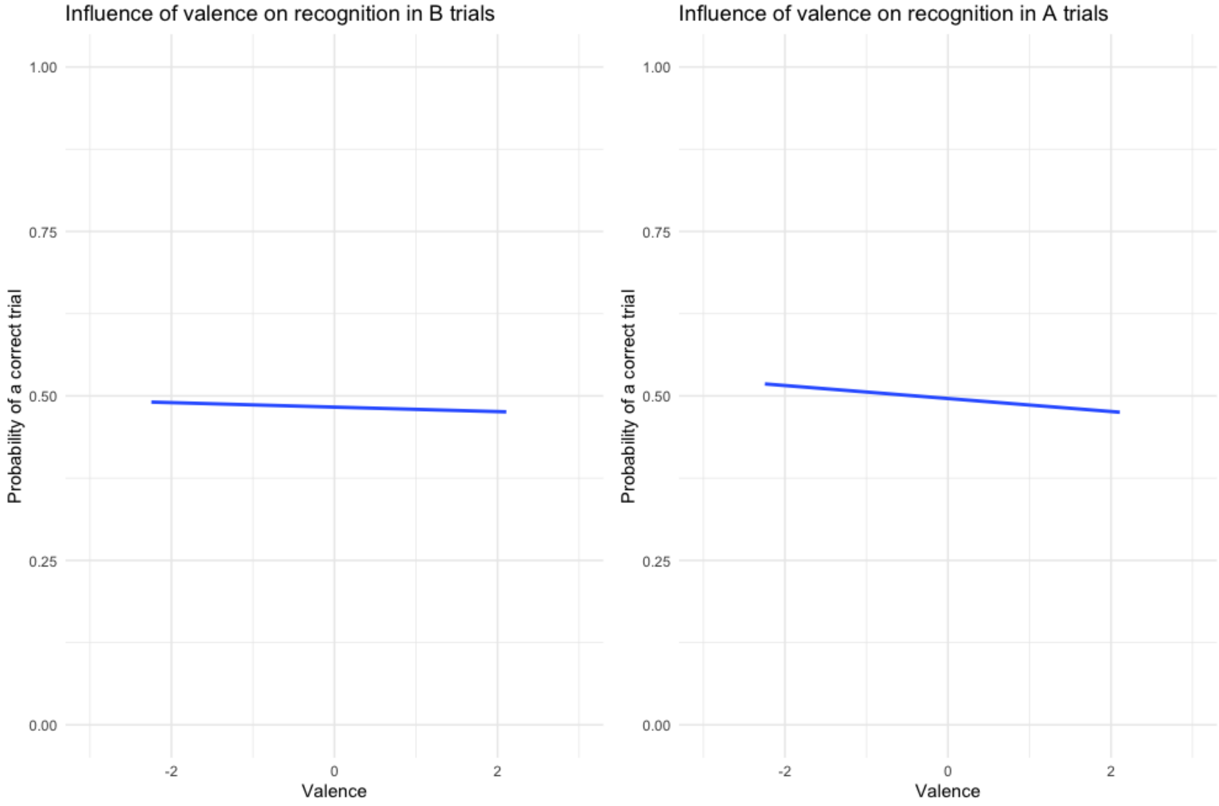
Table 7

Confidence intervals and *p*-values for the model coefficients for influence of valence on old-new discrimination.

	b	*SE*	*z-value*	*95% CI lb*	*95% CI lb*	*p-value*
Intercept	.43	.02	18.69	.38	.47	<.001*
Valence	-.13	.02	-8.22	-.16	-.1	<.001*

Figure 5

Models depicting the probability of a correct trial based on valence for exposure and associate stimulus trials, as predicted by binomial models.



Discussion

Key findings

This study aimed to investigate the hypothesized memory enhancement for cheaters and test its proposed inferential role. By using the preference-by-association paradigm (Wimmer & Shohamy, 2012), this study attempted to trigger retroactive spread of social value across associated images to see if this mechanism may be particularly sensitive to cheaters. This study did not find any general explicit memory enhancement for cheaters (1), which is contrary to theoretical predictions but consistent with the general research consensus (Buchner et al, 2009; Li & Nie, 2021). Moreover, there was initially no discernable decision bias found in either high or low memory performance datasets, for either exposure stimuli (2a) or associate stimuli(2a) trials, which suggests that there was no retroactive spread of value (2b, 3b), and thus no general sensitivity towards cheating (2c, 3c), contrary to predictions. It is key to note that these general decision bias models explain the most variance in the data, despite the non-significance of predictors. This may be due to a lack of power in the data, but also likely due to interindividual differences.

While there were non-significant effects of dispositional trust (4) and previous experience (5) of cheating in exposure trials, there was a strong negative cheating bias in high memory performance associate stimuli trials in the GTS score model (4). This suggests that once the interindividual variation in dispositional trust is accounted for, there is a general negative decision bias against cheaters (3c). Additionally, there was an interaction between GTS score and decision bias toward cheaters; for each unit of increase in GTS score, there was an increase in positive decision bias toward individuals associated with cheaters. This effect was not present in the low performance associate stimulus trials, suggesting the effect required explicit associative memory and available knowledge for both individuals presented.

Moreover, by including age as a predictor, a strong, negative decision bias against cheaters emerged in the low memory performance exposure stimulus trials (6). This suggests when participants remembered a cheater, but not the other person, they were less likely to prefer the cheater as trustworthy. This effect was nonetheless non-significant. There also appeared to be moderate interaction effects with age; older individuals exhibited a positive decision bias towards both cheaters and cooperators compared to neutral individuals (2b, 2c). This suggests that older participants were more likely to trust an individual they had information regarding their trustworthiness more than an individual of whom they were not in possession of trust-relevant information. In the high memory performance dataset, i.e. where participants remembered both individuals in the trial, this indiscrimination largely persisted, with the positive cheating bias not quite reaching significance. This interaction effect was not present in the associate stimulus trials, however, suggesting there was either no generalization overall or that this indiscrimination did not generalize (3b, 3c).

While there were only weak, non-significant gender differences between men and women in decision bias in both exposure (2a, 2c) and associate stimuli trials (3a, 3c), non-binary participants (7) showed a strong positive decision bias towards cooperators in associate stimuli trials (3b) but not in exposure stimuli trials (2b). This effect was only significant in low memory performance trials, indicating that non-binary individuals were more likely to trust both cooperators and cooperators associations compared to individuals they did not recognize, but showed no significant discernment between cheaters and neutral individuals, nor their associations. However, this particular result ought to be cautiously interpreted, given that the sample consisted of only 5 non-binary participants and thus likely heavily influenced statistically by outliers.

A negativity bias was observed in the old-new discrimination task (8a); individuals associated with negative descriptors were more likely to be remembered than individuals associated with positive descriptors (8b). This is conducive with the proposition that the memory enhancement for cheaters is a simple negativity bias, rather than a highly specialized cognitive module. Valence was also a significant predictor for decision bias for both stimulus batteries (9a, 9b), but this effect appeared to be minor and negligible, explaining very little of the variance in the data. This suggests that while negative valence enhances recognition, negativity itself has only a small impact on decision bias and generalization of negative value.

Implications

Theoretical

The failure to find a memory enhancement for cheaters adds to the growing body of literature questioning the existence of a specialized cheating detection module (Mehl & Buchner, 2008; Lie & Nie, 2021; Bell & Buchner, 2010; Bell & Buchner, 2011). The results obtained in this study are mixed and vary with interindividual characteristics, which is in opposition of the universal mechanism proposed by Cosmides and Tooby (1992). This study did observe a negativity bias (Bell & Buchner, 2012), but this bias accounted for little variation in social decision-making, which challenges the proposition that this general mechanism explains the processing and memory of socially relevant information (Kroneisen et al, 2014; Rothermund et al, 2008). Moreover, as generalization of negative value between exposure and associate stimulus images was negligible, this may suggest that the formation of decision bias is largely dissociated from negative information processing.

The main scientific contribution of this study is the exploration of intricate interactions between memory, decision-making and trust. Specifically, these processes are complex and subject to interindividual differences in dispositional characteristics. The models presented in this study accounted for a limited amount of variance, which suggests that are many other, unexplored components influencing social decision-making. The interaction between dispositional trust and behavioral descriptor suggests there may be a conditional cheating sensitivity. This interaction occurred despite no decision bias against the cheaters themselves, and only in the trials where both individuals were later recognized. This suggests that generalization was contingent on explicit memory traces and that individuals who have lower dispositional trust are potentially more likely to attend to information about associations with cheaters compared to with cooperators. This sensitivity is consistent with previous findings that suggest personal relevance of cheaters (Kroneisen, 2018) and survival relevance of their behavior (Bell & Buchner, 2012; Tse & Altarriba, 2010) are key conditions for any memory enhancement. These factors, in turn, contribute to a more specialized adaptivity of memory which emphasizes that different environment and exposures may propose different fitness challenges (Nairne & Pandereida, 2016), leading to interindividual differences.

Practical

Trust is crucial for many social situations, and the formation of trust has not been studied in great depth within memory research. This study found that participants with less dispositional trust exhibited a greater decision bias against individuals who had been associated with cheaters, despite having no direct knowledge of this individual's trustworthiness. This may influence approach-avoidance behavior (Rothermund et al, 2008; Cosmides & Tooby, 1992), which in turn affects many social outcomes such as loneliness (Matthews et al, 2022). As there was no significant influence of previous experience with cheating behavior, the formation of trust, or rather distrust, may be influenced by indirect effects, such as generalization. This study highlights the need for more research on the formation of trust.

In addition to dispositional trust, age appeared to influence decision bias against both cheaters and cooperators. The models on decision bias and age found that in the decision bias trials where participants remembered both individuals, there was a positive decision bias towards cooperators which was positively impacted by age. However, if only one individual in the trial was remembered, older participants exhibited a positive decision bias towards both cooperators and cheaters. This suggests that older participants are more likely to trust a cheater, despite memory of this cheater, if no trust-relevant information is available on the other individual. This highlights how there may be age-related differences in memory biases, but also that these may have detrimental effects on decision-making. This in turn has important practical implications for protection against exploitation, such as developing strategies to circumvent these memory biases by actively attending to socially relevant information.

Finally, the limited gender differences in this study oppose the preconceived notion that women are more trusting than men. There was no noteworthy difference between men and women in trials where participants had been directly exposed to socially relevant information. Additionally, in the low memory performance dataset, women were more likely to trust a neutral individual or a cooperator than a cheater, while men had no clear bias. The more striking difference was the generalization of cooperation observed in non-binary individuals. This may have implications for a multitude of social interactions by stimulating approach behavior, something which may function as a protection mechanism against, or be a symptom of, ostracization or other challenges faces by the non-binary community. While this effect is likely due to the limited number of non-binary participants, it highlights the importance of gender-inclusion, especially in research testing general versus specialized mechanisms and evolutionary theories.

Limitations

There are some limitations to this study which entail that results ought to be interpreted with some caution. Firstly, there are some methodological limitations due to pragmatic trade-offs required in the design of a memory experiment. For example, this study employed quite complex tasks which involved several higher levels of cognitive processing to promote deeper encoding (Diana et al, 2008). However, more complex tasks may risk participant performance varying due to differences in cognitive abilities such as attention and working memory, rather than the intricate effects this study aimed to measure, thus impacting the validity of the models. While previous research on logic tasks has shown that performance on social contract tasks tends to be less affected my differences in cognitive abilities (van Lier et al, 2013), no such comparisons have been made with the level of complexity in this study. Conversely, the self-paced nature of the tasks may have led to limited memory engagement, as participants can adopt different strategies, leading to both inter- and intraindividual variation.

Moreover, the ecological validity of this study is low. The tasks employed by this study do not represent real-world situations, as they are designed to maintain internal control, thus the results have limited ecological validity. Furthermore, the dependent variable in this study was decision bias in response to which of the two individuals presented seemed more trustworthy. This does not measure cooperative actions; the pilot experiment for this study found that while there was a decision bias against cheaters, there did not appear to be any significant difference in cooperation likelihood between cheaters, cooperators, and neutral individuals. Fundamentally, cheating sensitivity is only an evolutionary stable strategy if it impacts approach-avoidance behavior and influences social rewards and punishments (Ibrahim, 2022). Future studies will need to further dissect the influence of generalization in social interactions on cooperative decision-making, in both perception and action.

Sample characteristics and sample size also present some limitations to the generalizability of the results. For example, the online recruitment of participants presents a sampling bias; the sample represents individuals with an online presence who self-selected participation, which may present different motivations and characteristics to the general population. Individuals who choose to spend 15 minutes on an online experiment may be more conscientious and cooperative than the general norm. The sample size was also relatively small, which adds to the likelihood that there are certain influential response biases. Moreover, the small sample limited the statistical power of the analysis, especially with the low number of trials per condition.

Future research directions

Further research should be conducted on personal experiences with cheaters. While there was no significant impact of previous experiences of cheating in this study, this may be due to the measurement used; while the MoCS was constructed to encapsulate relevant cheating memories, this measurement may have been too blunt to capture how experiences with cheating affected participants, e.g. an event which altered dispositional trust. Most importantly, some items of the scale may be more influential than others on social decision-making, such as experiencing bullying compared to local corruption, which was not accounted for in this pilot-use of the scale. However, there are validated questionnaires for specific types of trust, such as the Trust in Close Relationships Scale (Rempel et al, 1985) which measures trust in romantic partners. Using a preference-by-association paradigm to explore how this type of trust, as well as previous experiences of romantic cheating, may influence social perception and decision-making in infidelity-specific scenarios may shed light on other socially adaptive qualities of generalization (Ein-dor et al, 2015).

Moreover, the interaction between decision bias and age suggests cheating sensitivity may vary with ontogeny. Most research conducted on the hypothesized memory enhancement for cheaters has been conducted on young, student populations, while thus study had a quite wide age range. However, it is unclear where this originates from. The tendency to trust others may vary between age-groups, or there may be cognitive differences in attention or memory of cheating behavior. Alternatively, this may reflect generational differences; as this study was cross-sectional, ontogeny of and any progressive changes in cheating sensitivity were not captured. Future studies ought to diversify sample populations to further investigate this phenomenon.

Finally, several design choices in this study can be altered in future studies to study other aspects of this mnemonic process. For example, while this study tested if explicit memory of association between the associate stimuli and exposure stimuli was necessary for generalization, this study did not try to measure if generalization of value was contingent on awareness of the three-way association with the behavioral descriptor. Future studies ought to account for this, which may also be a key part of the puzzle; perhaps individuals who exhibited limited decision bias but high memory performance simply did not care about the behavioral descriptors, or that this generalization is contingent on some awareness. In addition, as this study wanted to avoid priming effects, participants were not asked to memorize or recall associations and thus no such data is available. Incidental encoding may differ from intentional encoding (Tse & Altarriba, 2010), in both memory enhancement and generalization. In sum, future research ought to explore the role of conscious reasoning and awareness of bidirectional associations in generalization and decision bias.

Conclusion

This study contributed to the continued debate on the evolutionary origins of socially relevant memory functions. While this study did not find a general cheating sensitivity in generalization, negative biases against cheaters and cheater associates did emerge when controlling for certain interindividual variables. Generalization of cheating appeared to be influenced by dispositional trust, but only if there was explicit memory of associations. Moreover, age appeared to influence social decision-making toward both cooperators and cheaters. Finally, while a negativity bias was observed, this general mechanism failed to explain more of the variance in the data than behavioral descriptors. Investigating integrative encoding of social information contributes to both evolutionary psychology and cognitive neuroscience, by dissecting the role of interindividual differences and the potential adaptive performance of memory mechanisms. With this interdisciplinary approach and statistical modelling, this study was able to show the intricate relationship between cheating sensitivity, inference, and trust.

References

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390–1396. https://doi.org/10.5433/2178-8189.2021v25n2p208

Barclay, P., & Lalumière, M. L. (2006). Do people differentially remember cheaters? *Human Nature*, *17*(1), 98–113. https://doi.org/10.1007/s12110-006-1022-y

Bastin, C., Diana, R. A., Simon, J., Collette, F., Yonelinas, A. P., & Salmon, E. (2013). Associative memory in aging: The effect of unitization on source memory. *Psychology of Aging*, *28*(1), 275–283. https://doi.org/10.1037/a0031566.Associative

Bastin, C., Van Der Linden, M., Schnakers, C., Montaldi, D., & Mayes, A. R. (2010). The contribution of familiarity to within- and between-domain associative recognition memory: Use of a modified remember/know procedure. *European Journal of Cognitive Psychology, 22*(6), 922–943. https://doi.org/10.1080/09541440903007834

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bell, R., & Buchner, A. (2010). Valence modulates source memory for faces. *Memory & Cognition*, *38*(1), 29–41. https://doi.org/10.3758/MC.38.1.29

Bell, R., & Buchner, A. (2011). Source memory for faces is determined by their emotional evaluation. *Emotion*, *11*(2), 249–261. https://doi.org/10.1037/a0022597

Bell, R., & Buchner, A. (2012). How adaptive is memory for cheaters? *Current Directions in Psychological Science*, *21*(6), 403–408. https://doi.org/10.1177/0963721412458525

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*(3), 261–275. https://doi.org/10.1016/j.cognition.2010.08.020

Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, *30*(3), 212–224. https://doi.org/10.1016/j.evolhumbehav.2009.01.004

Cassidy, B. S., & Gutchess, A. H. (2015). Influences of appearance-behaviour congruity on memory and social judgements. *Memory*, *23*(7), 1039–1055. https://doi.org/10.1080/09658211.2014.951364

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of culture*. Oxford University Press.

Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. *CIBA Foundation Symposia*, *208*, 132–161. https://doi.org/10.1002/9780470515372.ch8

Cosmides, L., Tooby, J., Fiddick, L., & Bryant, G. A. (2005). Detecting cheaters. *Trends in Cognitive Sciences*, *9*(11), 503–505. https://doi.org/10.1016/j.tics.2005.09.001

de Araujo Sanchez, M. A., & Zeithamova, D. (2023). Generalization and source memory in acquired equivalence. *Cognition*, *234*(105385), 1–13. https://doi.org/10.1016/j.cognition.2023.105385

Del Giudice, M., Gangestad, S. W., & Kaplan, H. S. (2015). Life history theory and evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology – Vol 1: Foundations* (2nd ed., pp. 88-114). Wiley.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730.The

Ein-Dor, T., Perry-Paldi, A., Hirschberger, G., Birnbaum, G. E., & Deutsch, D. (2015). Coping with mate poaching: Gender differences in detection of infidelity-related threats. *Evolution and Human Behavior*, *36*(1), 17–24. https://doi.org/10.1016/j.evolhumbehav.2014.08.002

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation review-article. *Nature Human Behaviour*, *2*(7), 458–468. https://doi.org/10.1038/s41562-018-0385-5

Grant, S. S., Huskey, A. M., Faunce, J. A., & Friedman, B. H. (2020). Chapter 8 - Negative memory biases in health and psychiatric disorders. In T. Aue & H. Okon-Singer (Eds.), *Cognitive Biases in Health and Psychiatric Disorders* (pp. 173–191). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-816660-4.00008-8

Henrich, N., & Henrich, J. P. (2007). Why humans cooperate: A cultural and evolutionary explanation. Oxford University Press.

Ibrahim, A. M. (2022). The conditional defector strategies can violate the most crucial supporting mechanisms of cooperation. *Scientific Reports*, *12*(1), 1–10. https://doi.org/10.1038/s41598-022-18797-2

Jiang, J., Bramão, I., Khazenzon, A., Wang, S. F., Johansson, M., & Wagner, A. D. (2020). Temporal dynamics of memory-guided cognitive control and generalization of control via overlapping associative memories. *Journal of Neuroscience*, *40*(10), 2343–2356. https://doi.org/10.1523/JNEUROSCI.1869-19.2020

Kassambara, A. (2020). ggpubr: ’ggplot2’ based publication ready plots. R package version 0.4.0. https://CRAN.R-prokect.org/package=ggpubr

Kassambara, A. (2021). rstatix: Pipe-friendly framework for basic statistical tests. R package version 0.7.0. https://CRAN.R-project.org/package=rstatix

Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychological Review*, *109*(2), 306–329. https://doi.org/10.1037/0033-295X.109.2.306

Kroneisen, M. (2018). Is he important to me? Source memory advantage for personally relevant cheaters. *Psychonomic Bulletin and Review*, *25*(3), 1129–1137. https://doi.org/10.3758/s13423-017-1345-1

Kroneisen, M., & Bell, R. (2013). Sex, cheating, and disgust: Enhanced source memory for trait information that violates gender stereotypes. *Memory*, *21*(2), 167–181. https://doi.org/10.1080/09658211.2012.713971

Kroneisen, M., Woehe, L., & Rausch, L. S. (2014). Expectancy effects in source memory: how moving to a bad neighborhood can change your memory. *Psychonomic Bulletin and Review*, *22*(1), 179–189. https://doi.org/10.3758/s13423-014-0655-9

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi:10.18637/jss.v082.i13.

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*(3), 247–271. https://doi.org/10.1080/20445911.2021.1894157

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Matthews, T., Caspi, A., Danese, A., Fisher, H. L., Moffitt, T. E., & Arseneault, L. (2022). A longitudinal twin study of victimization and loneliness from childhood to young adulthood. *Development and Psychopathology*, *34*(1), 367–377. https://doi.org/10.1017/S0954579420001005

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*, *11*(3), 126–135. https://doi.org/10.1016/j.tics.2006.12.003

Mealey, L., Daood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology*, *17*(2), 119–128. https://doi.org/10.1016/0162-3095(95)00131-X

Mehl, B., & Buchner, A. (2008). No enhanced memory for faces of cheaters. *Evolution and Human Behavior*, *29*(1), 35–41. https://doi.org/10.1016/j.evolhumbehav.2007.08.001

Nairne, J. S., & Pandeirada, J. N. S. (2016). Adaptive memory: The evolutionary significance of survival processing. *Perspectives on Psychological Science, 11*, 496–511. https://doi.org/10.1177/1745691616635613.

Nairne, J.S., Thompson, S.R., & Pandeirada, J.N. (2007). Adaptive memory: survival processing enhances retention. *Journal of experimental psychology. Learning, memory, and cognition, 33 2*, 263-73. https://doi.org/10.1037/0278-7393.33.2.263

Nakazawa, M. (2023). fmsb: Functions for medical statistics book with some demographic data. R package version 0.7.5. https://CRAN.R-project.org/package=fmsb

Niedtfeld, I., & Kroneisen, M. (2020). Impaired memory for cooperative interaction partners in borderline personality disorder. *Borderline Personality Disorder and Emotion Dysregulation*, *7*(1), 1–9. https://doi.org/10.1186/s40479-020-00137-3

Oda, R. (1997). Biased face recognition in the Prisoner's Dilemma game. *Evolution and Human Behavior, 18*, 309–315. https://doi.org/10.1016/S1090-5138(97)00014-7

OpenAI. (2023). *ChatGPT* (GPT-3.5) [Large language model]. https://chat.openai.com/share/7f51346c-0818-4f91-9a05-80a82d7c47ba.

Pavlovia. (2023). Online Research Platform. https://pavlovia.org/

Peirce, J. W., Hirst, R. J., & MacAskill, M. R. (2022). *Building experiments in PsychoPy.* (2nd ed.). Sage.

Pärnamets, P., Shuster, A., Reinero, D. A., & Van Bavel, J. J. (2020). A value-based framework for understanding cooperation. *Current Directions in Psychological Science, 29*(3), 227–234. https://doi.org/10.1177/0963721420906200

Qualtrics. (2023). Online survey platform. https://www.qualtrics.com

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*(1), 95–112. https://doi.org/10.1037/0022-3514.49.1.95

Rothermund, K., Voss, A., & Wentura, D. (2008). Counter-regulation in affective attentional biases: A basic mechanism that warrants flexibility in emotion and motivation. *Emotion, 8,* 34–46. https://doi.org/10.1027/1864-9335/a000043

RStudio Team (2021). RStudio: Integrated development environment for R. RStudio, PBC, Boston, MA.

Rule, N. O., Slepian, M. L., & Ambady, N. (2012). A memory advantage for untrustworthy faces. *Cognition*, *125*(2), 207–218. https://doi.org/10.1016/j.cognition.2012.06.017

Schaper, M. L., Mieth, L., & Bell, R. (2019). Adaptive memory: Source memory is positively associated with adaptive social decision making. *Cognition*, *186*, 7–14. https://doi.org/10.1016/j.cognition.2019.01.014

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*(2), 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Suzuki, A., & Suga, S. (2010). Enhanced memory for the wolf in sheep's clothing: Facial trustworthiness modulates face-trait associative memory. *Cognition*, *117*(2), 224–229. https://doi.org/10.1016/j.cognition.2010.08.004

Swedish Ethical Review Authority. (2023). *What the act says*. https://etikprovningsmyndigheten.se/en/what-the-act-says/

Tay, P.K., Jonason, P.K., Li, N.P., & Cheng, G.H. (2019). Is memory enhanced by the context or survival threats? A quantitative and qualitative review on the Survival Processing Paradigm. *Evolutionary Behavioral Sciences, 13*, 31–54. https://doi.org/10.1037/ebs0000138

The Act concerning the Ethical Review of Research Involving Humans (2003:460). https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/lag-2003460-om-etikprovning-av-forskning-som_sfs-2003-460/

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(August 2014), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755

Tse, C. S., & Altarriba, J. (2010). Does survival processing enhance implicit memory? *Memory and Cognition*, *38*(8), 1110–1121. https://doi.org/10.3758/MC.38.8.1110

Tucker, H. L. (2023). Generalization of behavior: A pilot study on the role of integrative encoding in cheating-detection [A preliminary investigation on the impact of cheating on decision bias]. Unpublished report.

van Lier, J., Revlin, R., & de Neys, W. (2013). Detecting cheaters without thinking: Testing the automaticity of the cheater detection module. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0053827

van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, *108*(3), 796–803. https://doi.org/10.1016/j.cognition.2008.07.002

Venables, W. N., & Ripley, B. D. (2002). Modern Applied Statistics with S (Fourth Edition). Springer, New York. ISBN 0-387-95457-0.

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2021). dplyr: A grammar of gata ganipulation. R package version 1.0.5. https://CRAN.R-project.org/package=dplyr

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*, *338*, 270–273. https://doi.org/10.1126/science.1223252

Wolke, D., Copeland, W. E., Angold, A., & Costello, E. J. (2013). Impact of bullying in childhood on adult health, wealth, crime and social outcomes. *Psychological Science*, *24*(10), 1–9. https://doi.org/10.1177/0956797613481608.Impact

Yamagishi, T. & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, *18*, 129-166. https://doi.org/10.1007/BF02249397

Zeileis, A., & Grothendieck, G. (2005). zoo: S3 Infrastructure for regular and irregular time series. *Journal of Statistical Software, 14*(6), 1-27. doi:10.18637/jss.v014.i06

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: Building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*, 1–14. https://doi.org/10.3389/fnhum.2012.00070

Appendix A

Figure A1

Sample Images used in the Experiment, Obtained from the CFD.

Appendix B

Table B1

Full List of Behavioral Descriptors and their Accompanying Valence, by type.

Behavioral type	*Descriptor*	*Valence*
Cooperative	Returned a stray wallet to the police	1.98
	Picked up trash after a picnic in the park	1.74
	Donated blood at a local blood drive	1.98
	Participated in a charity ice hockey game	1.58
	Watched father's dog over the Christmas holidays	1.62
	Shovelled grandfather's driveway after a bad snowstorm	2.00
	Put up the tent for a fundraiser	1.39
	Brought groceries to a quarantined neighbour	1.98
	Organized a toy-drive for impoverished children	2.09
	Gave used clothes to nearby homeless shelter	1.91
	Contributed to a gift for retiring colleague	1.40
	Supported a friend through a mourning period	2.09
	Coordinated a street clean-up in their neighbourhood	1.94
	Tutored underprivelieged children in math for free	2.11
	Volunteered at the town's soup kitchen	1.89
	Participated in a medical study for diabetes	1.48
	Signed up to the national donor register	1.69
	Cooperated with the company's energy-saving initiative	1.37
	Tipped waiter at a fine-dining restaurant	1.18
	Recycled bottles and cans after a party	1.78
	Repainted grandmother's old fence over the weekend	1.75
	Repaired mother's engagement ring when it broke	1.74
	Helped an elderly man cross the street	1.89
	Paid back small loan on their car	1.17
	Reimbursed employee for a work lunch	1.22
Cheating	Had an affair while in a relationship	-2.11
	Ran an internet-scam on young people	-1.94
	Used to be a corrupt local politician	-2.06
	Hit a parked car and drove away	-1.78
	Regularly takes money from their mother's purse	-1.73
	Cheated on a very difficult entrance exam	-1.17
	Used a fake ID to buy alcohol	-1.09
	Betrayed their best friend of five years	-2.15
	Misappropriated money intended for an environmental charity	-2.05

	Lied on resume to secure a job	-0.95
	Sold expired eggs labelled as fresh ones	-2.09
	Swindled a small business out of money	-2.20
	Tricked a vegetarian into eating chicken nuggets	-2.08
	Has not paid taxes for two years	-1.72
	Pretended to be younger to receive benefits	-1.09
	Used dating-apps while in a relationship	-1.70
	Plagiarized a friend's book and became famous	-2.25
	Photoshopped advertisement images to sell makeup products	-0.94
	Faked being sick to receive others sympathy	-1.55
	Failed to return a rented popcorn machine	-1.13
	Borrowed grandfather's watch without asking for permission	-1.06
	Took silverware from a fine-dining restaurant	-0.77
	Broke promise made to their ailing grandmother	-1.72
	Rigged card game to beat their father	-0.80
	Withheld employees wages to pay own bills	-2.23
Neutral	Received an invitation to a destination wedding	1.12
	Gave a lecture on economics over Zoom	0.62
	Went to a pumpkin patch on Halloween	0.97
	Sent a long letter in the mail	0.62
	Ordered surf and turf at a restaurant	0.29
	Attended a local stand-up comedy show	1.08
	Watched an exciting football game on TV	0.69
	Bought a scented candle for the bedroom	1.11
	Rented a tuxedo for a fancy event	0.69
	Took a cooking-class with their father	1.57
	Vacationed in a warm country with their mother	1.25
	Built a tool-shed over the summer	1.03
	Completed a difficult online Python programming course	1.23
	Listened to a podcast about art history	0.72
	Learned a new language while studying abroad	1.65
	Barbequed burgers for the family reunion dinner	1.11
	Sold their lawnmower after only two years	-0.02
	Ran two marathons in a year	1.06
	Used a chair to reach the lightbulb	0.34
	Watered the houseplants with a flower fertilizer	0.82
	Dropped a heavy pan on their toe	-1.23
	Started playing chess with their grandmother weekly	1.58
	Wrote a book about well-known grandfather	1.18
	Hosted a house-warming party last Saturday	1.13
	Bleached their hair after their last breakup	-0.20

Appendix C

General Trust Scale, from Yamagishi & Yamagishi (1994).

Scale:

Using the following scale, please indicate how much you agree or disagree with the following statements:

1	2	3	4	5
Strongly Disagree	*Disagree*	*Neutral*	*Agree*	*Strongly Agree*

1.) Most people are basically honest.
2.) Most people are trustworthy.
3.) Most people are basically good and kind.
4.) Most people are trustful of others.
5.) I am trustful.
6.) Most people will respond in kind when they are trusted by others.

Scoring:

The score for each item is averaged together to form a continuous measure of generalized trust.

Appendix D

The Memory of Cheating Scale, Developed with the Aid of OpenAI (2023).

Please indicate with a simple "Yes" or "No" whether you have experienced the following:

1. Have you ever been cheated on or betrayed in a romantic relationship?
2. Have you ever been a victim of fraud or scams, either online or offline?
3. Have you ever encountered or been affected by corruption in your local community or government?
4. Have you ever experienced theft or had someone misappropriate your belongings or funds?
5. Have you ever witnessed or suspected academic dishonesty, such as cheating or plagiarism?
6. Have you ever been a victim of identity fraud or had someone impersonate you?
7. Have you ever been deceived or manipulated by someone for personal gain?
8. Have you ever encountered plagiarism, where someone took credit for your work without permission?
9. Have you ever come across false advertising or misleading claims while purchasing products or services?
10. Have you ever had a negative experience with rental agreements, such as someone not returning borrowed items or violating rental terms?
11. Have you ever been bullied, where someone repeatedly subjected you to physical, verbal, or emotional harm?

Appendix E

Table E1

Summary of the Exposure Stimuli Decision Bias Model with Behavioral Descriptors and GTS Score as Predictors in High and Low Memory Performance Trials.

		Coefficients		*Confidence intervals*			
Memory perf.	*Behavioral descriptor*	*b*	*β*	*95% CI lb*	*95% CI ub*	*t-value (df=327)*	*p-value*
Low	Neutral	55.7	0	33.13	78.26	4.807	<.001*
	Cheating	-7.04	-.13	-38.96	-24.88	-.43	.668
	Cooperative	-12.78	-.23	-44.7	19.13	-.78	.436
	GTS score	-2.39	-.07	-9.02	4.24	-.78	.483
	Cheating × GTS score	3.02	.19	-6.35	12.4	.628	.53
	Cooperative × GTS score	5.05	0.31	-4.33	14.43	1.048	.295
						(df=325)	
High	Neutral	60.56	0	36.75	84.37	4.954	<.001*
	Cheating	-15.06	-.26	-48.7	18.58	-.872	.384
	Cooperative	-18.99	-.32	-52.72	14.75	-1.096	.274
	GTS score	-3.31	-.09	-10.3	3.68	-.923	.357
	Cheating × GTS score	4.85	.28	-5.03	14.73	.956	.34
	Cooperative × GTS score	5.86	0.34	-4.06	15.78	1.150	.251

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.

Appendix F

Table F1

Summary of the Decision Bias Model with Behavioral Descriptors and MoCS Score as Predictors in High and Low Memory Performance Exposure Stimuli Trials.

		Coefficients		*Confidence intervals*			
Memory perf.	*Behavioral descriptor*	*b*	*β*	*95% CI lb*	*95% CI ub*	*t-value (df=327)*	*p-value*
Low	Intercept (Neutral)	45.69	0	34.42	56.95	7.031	<.001*
	Cheating	1.93	.03	-14	17.86	-1.014	.813
	Cooperative	11.54	.21	-4.39	27.46	.563	.159
	MoCS score	.39	.04	-1.51	2.28	-.207	.692
	Cheating × MoCS score	.2	.02	-2.48	2.88	.897	.884
	Cooperative × MoCS score	-1.41	-.16	- 4.09	1.27	-.289	.305
						t-value (df=323)	
High	Intercept (Neutral)	46.56	0	34.68	58.44	6.75	<.001*
	Cheating	2.06	.03	-14.74	-18.86	-2.335	.811
	Cooperative	7.4	.13	-9.44	24.23	-1.461	.393
	MoCS score	.56	.05	-1.44	2.56	-1.718	.586
	Cheating × MoCS score	-.19	-.02	-3.01	-3.01	2.081	.897
	Cooperative × MoCS score	-1.3	-.14	-4.12	1.53	1.511	.372

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.

Table F2

Summary of the Decision Bias Model with Behavioral Descriptors and MoCS Score as Predictors in High and Low Memory Performance Associate Stimuli Trials.

		Coefficients		*95% confidence intervals*			
Memory perf.	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value (df=327)*	*p-value*
Low	Intercept (Neutral)	49.88	0	42.84	56.91	13.808	<.001*
	Cheating	1.84	0.05	-8.11	11.79	.36	.719
	Cooperative	-1.35	-0.04	-11.3	8.6	-.264	.792
	MoCS score	-0.08	-0.01	-1.26	1.11	-.124	.901
	Cheating × MoCS score	-0.61	-0.11	-2.28	1.06	-.709	.479
	Cooperative × MoCS score	0.79	-14	-0.88	2.46	.919	.359
						t-value (df=323)	
High	Intercept (Neutral)	.66	0	40.74	60.59	9.943	<.001*
	Cheating	-4.68	-.09	-18.72	9.35	-.65	.516
	Cooperative	-1.15	-.02	-15.23	12.94	-.159	.874
	MoCS score	.16	.02	-1.51	1.82	.182	.855
	Cheating × MoCS score	.06	.01	-2.3	2.42	.048	.962
	Cooperative × MoCS score	.25	.03	-2.11	2.62	.207	.836

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.

Appendix G

Table G1

Summary of the Decision Bias Model with Behavioral Descriptors and Age as Predictors in High and Low Memory Performance Associate Stimuli Trials.

		Coefficients		*95% confidence intervals*			
Memory perf.	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value (df=327)*	*p-value*
Low	Neutral	47.64	0	38.8	56.49	10.496	<.001*
	Cheating	.03	0	-12.48	12.53	.4	.516
	Cooperative	5.41	.15	-7.09	17.91	.843	.874
	Age	.06	.04	-.22	.34	.43	.855
	Cheating × Age	-.05	-.04	-.44	.34	-.24	.962
	Cooperative × Age	-.08	-.08	-.48	.31	-.421	.836
						t-value (df=323)	
High	Intercept (Neutral)	50.28	0	37.78	62.79	7.833	<.001*
	Cheating	-1.44	-.03	-19.12	16.24	-.159	.874
	Cooperative	2.92	.06	-14.84	20.67	.32	.749
	Age	.04	.02	-0.35	0.43	.202	.84
	Cheating × Age	-.1	-.07	-0.65	0.46	-.345	.731
	Cooperative × Age	-.09	-.06	-0.65	.47	-.317	.751

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.

Appendix H

Table H1

Model Summary Illustrating the Difference in Decision Bias According to Behavioral Descriptor by Gender, in both Low and High Memory Performance Exposure Trials.

			Coefficients		*95% confidence intervals*			
Memory perf.	*Gender*	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value (df=321)*	*p-value*
Low	Men	Intercept	49.83	0	42.53	57.14	13.17	<.001*
		Cheating	-.83	-.01	-11.16	9.49	-.156	.876
		Cooperative	1.33	.02	-8.99	11.66	.249	.803
	Women	Neutral	-3.32	-.06	-13.41	6.77	-.635	.526
		Cheating	7.2	.1	-7.07	21.47	.973	.331
		Cooperative	3.21	.05	-11.06	17.48	.434	.664
	Non-binary	Neutral	-4.83	-.04	-29.05	19.39	-.385	.7
		Cheating	.83	0	-33.42	35.09	.05	.963
		Cooperative	13.67	.06	-20.59	47.92	13.67	.442
							t-value (df=319)	
High	Men	Intercept	51.36	0	43.6	59.12	12.78	<.001*
		Cheating	-2.36	-.04	-13.28	8.56	-.42	.677
		Cooperative	-1.69	-.03	-12.62	9.23	-.299	.803
	Women	Neutral	-2.27	-.04	-12.94	8.41	-.41	.682
		Cheating	5.85	.08	-9.21	20.9	.749	.454
		Cooperative	1.21	.02	-13.88	16.3	.155	.877
	Non-binary	Neutral	-9.69	-.07	-35.21	15.82	-.733	.464
		Cheating	0.83	0	-33.42	35.09	.305	.963
		Cooperative	13.67	.06	-20.59	47.92	.68	.442

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.

Table H2

Model Summary Illustrating the Difference in Decision Bias According to Behavioral Descriptor by Gender, in High Memory Performance Associate Trials.

		Coefficients		*95% confidence intervals*			
Gender	*Behavioral descriptor*	*b*	*β*	*lb*	*ub*	*t-value (df=317)*	*p-value*
Men	Intercept	49.19	0	42.76	55.62	14.75	<.001*
	Cheating	-0.54	-.01	-9.64	8.56	-.115	.908
	Cooperative	2.76	.06	-6.38	11.91	.583	.56
Women	Neutral	5.05	.11	-3.8	13.9	-1.1	.272
	Cheating	-6.26	-.1	-18.78	6.25	-,966	.335
	Cooperative	-6.9	-.11	-19.45	5.6	-1.06	.289
Non-binary	Neutral	-8.91	-.08	-30.06	12.23	-.813	.417
	Cheating	-7.47	-.04	-37.38	22.43	-.482	.963
	Cooperative	24.26	.13	-5.66	54.18	1.57	.442

Note * denotes significance at alpha-level .05. Significance for neutral descriptors indicate deviation from 0. Unstandardized *b* is the difference in mean decision bias (%), in reference to the neutral intercept. Effect sizes are denoted by *β*.