

Memory distortions resulting from a choice blindness task

Philip Pärnamets (philip.parnamets@lucs.lu.se)

Lars Hall (lars.hall@lucs.lu.se)

Lund University Cognitive Science, Lund University,
Box 192, S-221 00, Lund, Sweden.

Petter Johansson (petter.johansson@lucs.lu.se)

Swedish Collegium for Advanced Study, Linneanum,
Thunbergsvägen 2, S-752 38 Uppsala, Sweden.
Lund University Cognitive Science, Lund University,
Box 192, S-221 00, Lund, Sweden.

Abstract

Using a choice blindness paradigm, it is possible to switch decisions and outcomes in simple choice tasks. Such switches have been found to carry over into later choices, hypothesized to be mediated by beliefs about earlier decisions. Here we investigated participants' memories for stimuli in a simple choice blindness task involving preferential choices between pairs of faces. We probed participants' recognition and source memory following a round of choices where on some trials participants were presented with the opposite face to the one they actually selected. We found no effect on recognition memory accuracy. Source memory was impaired such that participants failing to detect the manipulation later misremembered recognized non-chosen faces as being previously chosen. The findings are discussed in the light of self-perception theory and previous work on how beliefs affect memories for choices.

Keywords: choice blindness; memory; decision making; preference

Choice blindness

Choice blindness is the finding that people can be blind to mismatches between decisions and outcomes during simple choice situations. For example, in Johansson, Hall, Sikström & Olsson (2005) participants made preferential choices based on attractiveness between pairs of faces printed on playing cards. Following their choice, the experimenter presented the chosen card and asked the participants to motivate their choice. However, on some trials, using a surreptitious technique from close-up magic, the participants ended up with the card opposite of their choice. Participants were not only blind to this switch in the vast majority of trials, but also proceeded to confabulate reasons for choices they had never made.

Choice blindness is a robust effect and the basic findings have been replicated in a wide range of domains, from gustatory choices between the taste of jams and smell of tea, to attitude judgments on divisive political issues on the cusp of a national election. Together, the literature on choice blindness shows that humans are capable of a surprisingly high degree of cognitive flexibility in light of changing environmental feedback. Choice blindness, additionally, puts pressure on any cognitive architecture positing strong intention monitoring mechanisms for preferential choices (cf. Hall & Johansson, 2008). Instead, these findings

indicate that intention and agency are actively constructed in tandem with the feedback that the agent receives during and immediately following choice.

One important extant question concerning choice blindness is to understand why participants fail to detect the false feedback about their choices. One proposed possible reason has been that social demand effects in the participant-experimenter interaction lead to participants refraining from reporting when detecting the false feedback. Recently, using a computerized choice blindness task, psychophysiological correlates, such as pupil dilation and eye-movement patterns, were found to differ markedly between detected and non-detected trials. This indicates that detection registers as a differentiable event in the cognitive system, and consequently that participants' acceptance of the manipulation can be taken at face value (Pärnamets, Hall, Strandberg, Balkenius & Johansson, 2015a).

However, another possible factor in participants failing to detect manipulations might be that they fail to encode the choice options properly. In the present study, following a first block of choice trials, participants were given a recognition memory task involving one of the original faces and a foil option. This allowed us to compare recognition rates for manipulated and non-manipulated trials, as well as for detected and non-detected trials following an earlier choice blindness manipulation. If participants indeed fail to encode the options prior to accepting the false feedback, we can expect to see much lower recognition rates for non-detected trials, while observing similar recognition rates for detected and non-manipulated trials. If, however, failure to encode the choice options is not a factor determining detection, there should be no difference in recognition rate between non-detected and non-manipulated trials.

Choice blindness, preference change and source memory

Experiencing the choice blindness manipulation has been shown to carry downstream effects on participants' later choices. For example, one study found that participants' ratings of initially non-chosen faces, following false feedback, increased when rated a second time (Johansson, Hall, Tärning, Sikström & Chater, 2014). Similarly, when given a second round of choices, participants were more

likely to switch their preference to the option they had been manipulated to believe they originally preferred. This effect was primarily driven by preference change in the non-detected manipulated trials (Johansson et al., 2014). This preference change effect was recently expanded to group choices. Participants formed dyads and made mutual choices about future flatmates, presented as faces on a screen. On some trials, dyads were given false feedback about their mutual choices, with detection rates as low as 35%. When given a second round of choices, dyads were more likely to switch preference for previously manipulated trials, this despite there being two members available to monitor and remember the original preference (Pärnamets et al., 2015b). It has been hypothesised that beliefs about choices, formed during the post-choice feedback, drive the effects on preferences described.

People's memory of past choices tends to be systematically distorted in favor of previously selected items (Mather, Shafir & Johnson, 2000). For example, after first choosing between two options, such as apartment rentals, and later prompted to recall features of the options, people tend to attribute more positive features to chosen options and more negative features to non-chosen options. This bias in source monitoring might occur due to differential processing of options (for example via attentional mechanisms) during both encoding and retrieval of memories. By inducing false beliefs about which option had been chosen, one week after the original choice had been made, Henkel and Mather (2007) were able to demonstrate that belief in having made a previous choice might function as a mediating factor for this bias. In the study, participants exhibited the same attribution errors for originally non-chosen options as non-manipulated participants normally exhibit for chosen options. This result indicates that it is primarily the process of retrieval rather than the encoding that is responsible for the bias observed.

In the context of the choice blindness paradigm, there is mixed evidence concerning possible effects on participants' memories as a result of the false feedback concerning choices. One recent study indicated that memories for norm-violating behaviors might be influenced (Sauerland et al., 2013). There participants filled in a questionnaire concerning the frequency of past behavior involving transgressions such as cheating on tests or stealing kitchen utensils. While participants exhibited high detection rates for subsequent manipulations of their reports (~85%), undetected manipulations were later integrated into participants reports during a follow-up test. Norm-violating behaviors represent events with considerable affective force that can be expected to be strongly encoded in the first place. However, in a later study involving numerical ratings of female faces found that later recall accuracy of those ratings was unaffected by false feedback (Sagana, Sauerland & Merckelbach, 2014). This suggests that acceptance of the choice blindness manipulation need not impair later memories concerning the original choice.

In the present study, to begin to disentangle the mechanisms involved in post-choice preferential change, we examined the relationship between detection in a choice blindness task and later source memory accuracy for the options presented. Following a first recognition memory task, participants were given a second source memory task on the selected option from the recognition task. Participants were asked to identify if the recognized face was their original choice or not. We hypothesized that participants would misattribute their original preference to the believed-to-be chosen option for manipulated trials. Such a result would support the notion that *beliefs* about previous choices, rather than preferential adjustment through the act of choice, is what drives preference change over time (cf. Bem, 1967).

Method

Participants

We recruited 37 participants (26 female, 11 male) from the student population at Lund University, with an average age of 25.0 ($SD = 8.2$).

Procedure

The experiment had two phases, both consisting of 36 trials (see Fig. 1). In the first phase participants made preferential choices between pairwise presented face pairs. The pairs were presented in random order. At the start of each trial, two playing card were presented on-screen lying face down. After 0.5s the cards rotated so as to flip face up. The faces were displayed for 2s before flipping back. Once the animation was complete participants selected their preferred face by clicking on the corresponding card. The selection was marked by a colored rectangle surrounding the card of their choice (see Fig. 1). Participants were then asked to indicate their confidence in their choice on a 1-7 scale. This task lasted for at least 7s regardless of how fast participants responded. Once the occlusion time had passed the chosen, highlighted, card would flip back face-side up. Participants were then asked during this feedback screen to provide additional information about their choice: "You chose the face above. Why did you choose this face?". Six facial features were provided as responses, "mouth", "eyes", "nose", "proportion", "skin", and "shape". In addition, a seventh response was provided: "I actually prefer the other face". Participants selected one face by clicking on it. Following a 2s pause the next trial started.

After the completion of the first phase, participants did an unrelated filler task which lasted approximately 15 minutes. Following the filler task, the second phase commenced.

During each trial, participants were shown two faces: one face from one of the original face pairs, and one face not presented previously. The original face was randomly selected from the two previous options, meaning participants saw either the originally chosen or non-chosen face. The new face, which was a morph between one of the faces previously presented and a different non-presented

face (details described under Stimuli below). Participants were asked to select which face of the two they recognized from the first phase of the experiment. This was the recognition memory task. Once participants selected one face, the other disappeared, and they were asked if this face was the one they had originally chosen. This was the source memory task. There was a 2s pause between each trial.

Choice blindness manipulation

In 8 of the 36 trials during the first phase, the chosen face was not displayed during the feedback screen. Instead the non-chosen face was displayed. We refer to such trials with false feedback about the participants' choice as manipulated trials and the other trials with veridical feedback as non-manipulated trials.

We operationalized detection of the false feedback as when participants clicked the "I actually prefer the other face" button ('other preference' button). Such trials are referred to as detected trials and manipulated trials where

participants clicked any of the facial features are referred to as non-detected trials. A pilot study was conducted prior to running the present study testing this operationalization. We found that participants rarely clicked 'other preference' apart from on manipulated trials, and post-test interviews confirmed that participants use of the 'other preference' button coincided with them consciously being aware that the presented face was not their original choice.

The first six trials were always non-manipulated. Following those first trials, manipulated and non-manipulated trials were presented in random order with the condition that two manipulated trials never immediately followed one another.

Stimuli

Both male and female faces were used, but in same gender configurations.

To construct the stimuli, face quadruples were constructed in the following manner (see also Fig. 2). First face pairs

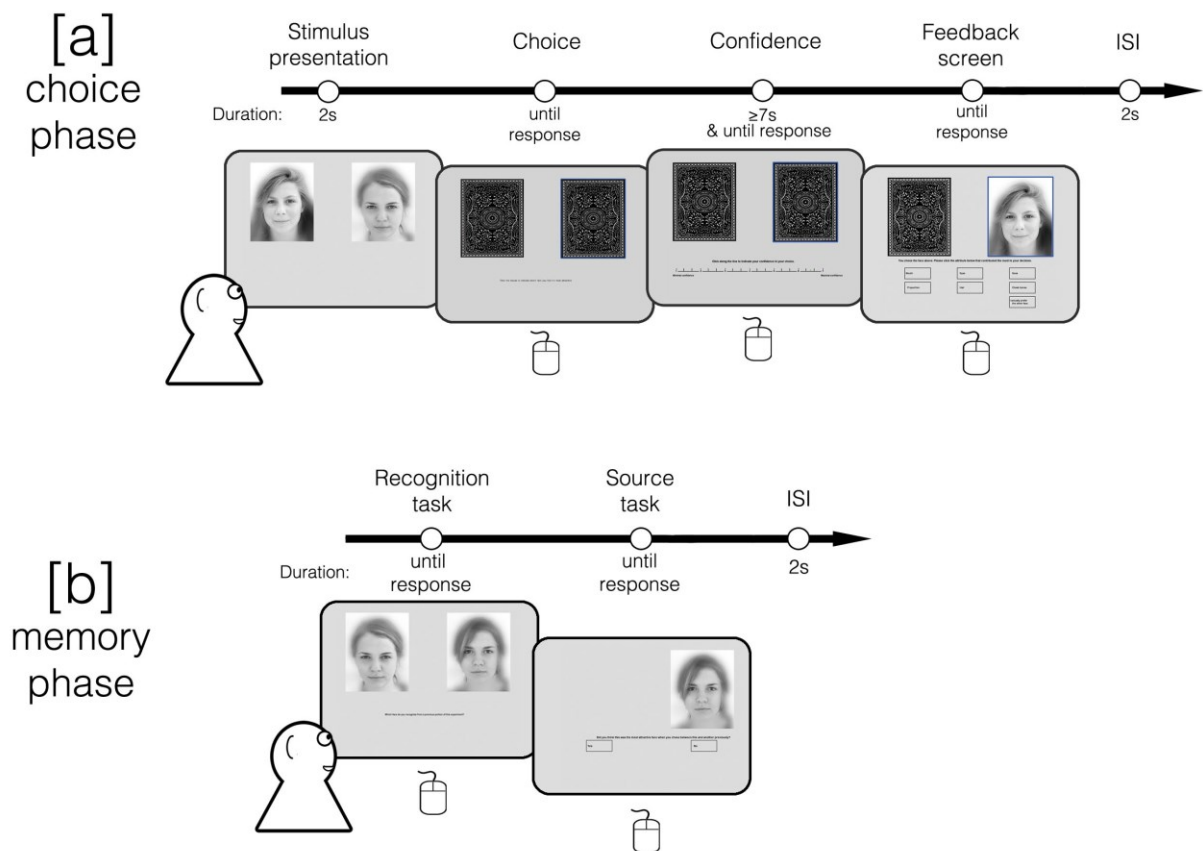


Figure 1. Overview of experiment [a] **Choice phase.** Participants are presented with two faces for 2s and then select their preferred option. The chosen face is presented during a feedback screen where participants are asked to indicate which facial feature of the chosen face contributed most to their decision. On manipulated trials (8/36; pictured) the non-chosen face was presented as chosen. [b] **Memory phase.** Participants were presented with a recognition task where one face from the original pair was shown together with a foil. Participants were asked to indicate which face they had seen during the choice phase. Following their selection participants were asked to indicate if the recognized face was their originally non-chosen or not [source memory task].



Figure 2. Example of stimuli. AA and BB are faces from the original database while AC and BC are morphs created to function as foils.

were selected from a larger database (Johansson et al., 2005). From these pairs of primary faces (labelled AA and BB in Fig. 2), two more faces were constructed by morphing the AA and BB face with a third face (CC, not shown) unique to each quadruple, resulting in two new faces, AC and BC. Morphs were created using Fantamorph software (Abrosoft Co., 2010).

Face pairs for the first phase of the experiment were determined at the start of the experiment from each quadruple. Participants always choose between one of the A and one of the B faces. Since the morphed faces, due to the invariable smoothing that occurs during their construction, might stand out compared to the original photos (i.e. AA and BB), the stimulus during the choice phase could be between any of the following pairings: AA & BB, AA & BC or AC & BB. In half the trials the first phase choice involved one of the morphed faces. For the recognition memory task in the second phase participants chose between either both the A* faces or both the B* faces. This design was chosen to make the recall task more difficult for participants, allowing us to probe if participants' failure to detect might be related to a failure to encode the original choice options.

Face pairs were presented in a randomized order during each experiment, ensuring that any face pair was eligible for choice blindness manipulation. This was to achieve roughly equal amounts of detected and non-detected trials for subsequent analysis.

Detection rates

Participants detected the manipulation in 62.8% of trials. Average by participant detection rate was 5.0 ($SD = 2.7$). This matched our expectations from piloting, for this specific set of stimulus items, and allowed us to have similar amounts of both detected and non-detected trials for further analysis.

Results

Recognition memory

Overall recognition rate was high, with participants correctly recognizing the target face in 89.4% of trials.

There was no difference in recognition memory accuracy comparing manipulated (91.2%) with non-manipulated trials (88.9%; Fisher's Exact Test, $OR = 1.30$, $p = .285$). Similarly, examining only the manipulated trials, there was no difference in accuracy comparing detected trials (92.5%)

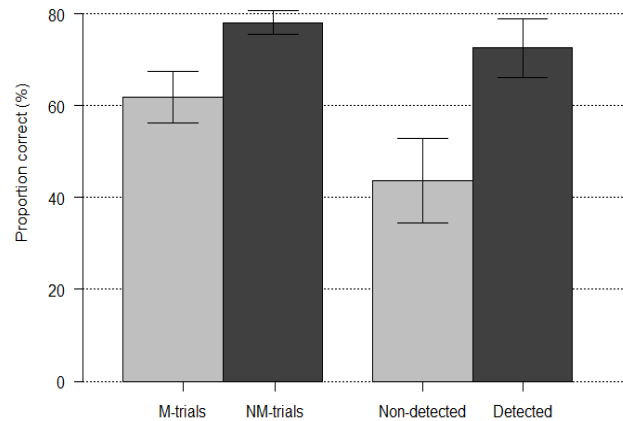


Figure 3 Source memory accuracy for manipulated and non-manipulated trials (left) and for non-detected and detected [manipulated] trials (right). Error bars depict 95% confidence intervals.

with non-detected trials (89.1%; Fisher's Exact Test, $OR = 1.50$, $p = .396$).

Preference and accuracy Overall accuracy improved if the recognition pair included the originally chosen face, with an accuracy rate of 93.1% when the originally chosen face was present and 85.1% when it was not (Fisher's Exact Test, $OR = 2.36$, $p < 10^{-5}$).

We compared the effect of the presence of the preferred face between manipulated and non-manipulated trials using a logistic regression with Presence and Trial Type as factors including an interaction term. Non-manipulated trials and absence of chosen face were taken as reference levels. The model was significant compared to a null model ($\chi^2_{(3)} = 53.04$, $p < 10^{-10}$), and a model without an interaction term ($\chi^2_{(1)} = 29.27$, $p < 10^{-7}$). The analysis indicated a significant interaction between Presence and Trial Type ($b = 2.69$, $OR = 14.7$, $p < 10^{-5}$), as well as significant effects of Presence ($b = -1.30$, $OR = 0.27$, $p = .011$), and Trial Type ($b = -1.71$, $OR = 0.18$, $p < .001$). Participants were more accurate in recognizing the face provided in the feedback, both on trials in which this feedback was veridical (94.8%) as well as when false feedback was given (96.2%), than they were in recognizing the face only shown once during the initial choice (veridical feedback: 82%, false feedback: 87.3%).

We also compared the detected with non-detected trials using a logistic regression with Presence and Detection as factors. The model had better fit than the null model ($\chi^2_{(2)} = 8.89$, $p = .012$), and adding an interaction term did not improve fit ($\chi^2_{(1)} = 0.14$, $p = .706$). We found a significant effect of Presence ($b = -1.31$, $OR = 0.22$, $p = .011$) but not of Detection ($b = 0.42$, $OR = 1.53$, $p = .31$). Participants' memory performance was highest when the originally preferred face was not present both for detected (96.3%) and non-detected trials (95.9%). When the originally preferred face was present accuracy was 89.4% for detected trials and 83.6% for non-detected trials.

Source memory

Overall accuracy for the source memory query was 74.5%. Source memory accuracy was higher following an accurate recognition memory response (76.5%) compared to when not (57.4%; Fisher's Exact Test, $OR = 2.41$, $p < 10^{-5}$).

Source memory accuracy was higher for non-manipulated trials (78.1%) compared to manipulated trials (61.8%; Fisher's Exact Test, $OR = 2.20$, $p < 10^{-7}$, see Fig. 3). Within manipulated trials, source accuracy was higher during detected trials (72.6%) compared to non-detected trials (43.6%; Fisher's Exact Test, $OR = 3.40$, $p < 10^{-6}$, see Fig. 3).

Response types To better understand how source memory might be affected by the prior choice blindness task we analyzed the distribution of responses available to the participants borrowing terminology from a signal detection framework. We classified all responses as being a True Positive, True Negative, False Positive, or False Negative. We then restricted our analyses to those trials where participants had made an accurate recognition memory response, as in the other case participants were already holding false beliefs about the queried face. Finally, since the relative distribution of all types of responses will be affected by the overall accuracy (proportion True responses), we compared the relative proportion of responses within each category, i.e. relative amount of True Positives compared with True Negatives and relative amount of False Positives with False Negatives.

For incorrect source memory responses, we found no difference between manipulated trials (49.5% False Positive) and non-manipulated trials (50.2% False Positive; Fisher's Exact Test, $OR = 1.03$, $p = 1$). Similarly, there was no difference for the correct responses comparing manipulated trials (54.9% True Positive) and non-manipulated trials (59.3% True Positive; Fisher's Exact Test, $OR = 1.20$, $p = .304$).

Comparing within the manipulated trials, participants had a higher rate of False Positive responses for non-detected trials (58.2%) compared to detected trials (38.1%; Fisher's Exact Test, $OR = 2.24$, $p = .066$). There was no difference between the rates of True Positives comparing non-detected trials (65.1%) with detected trials (51.5%; Fisher's Exact Test, $OR = 1.75$, $p = .157$). However, both these analyses suggest a higher rate of Positive response during non-detected trials. We tested this directly and found a higher Positive response rate for non-detected trials (61.2%) compared to detected trials (48.3%; Fisher's Exact Test, $OR = 1.69$, $p = .04$).

Discussion

Using both recognition and source memory tasks we investigated participants' memories for previously encountered choice options in the context of a choice blindness manipulation.

Overall, participants performed the recognition memory task with a high degree of accuracy. Importantly,

participants' accuracy did not differ between manipulated and non-manipulated trials or between detected and non-detected trials. This suggests that acceptance of the choice blindness manipulation is not due to a general failure of participants to attend to or encode the choice options. Our results are in line with previous findings indicating that recall accuracy for ratings is not affected by choice blindness manipulations (Sagana, Sauerland & Merckelbach, 2014).

Further analysis showed that participants were better at recognizing the choice option which had been presented to them during the feedback portion of the choice trials. This suggests that the prolonged exposure during the feedback portion of the trial has a predominant effect on accuracy (cf. Reynolds & Pezdek, 1992), rather than any preferential encoding of the chosen face. Merely, mentally refreshing a visually presented scene has found to have effects on memory similar to actual representation (Yi, Turk-Browne, Chun & Johnson, 2008). One interpretation of the findings might be that presenting one option a second time overrides, in manipulated trials, attempts to recall the actual chosen option, which participants arguably successfully do when detecting the false feedback. To address this issue fully and more extensively probe how recognition memory interacts with the choice feedback, future work should modify the false feedback portion, to include a condition where both faces are presented to participants. By combining with a measure of visual attention, exposure to the options could be precisely quantified and the recognition probe would become more sensitive to any differential encoding between detected and non-detected manipulated trials.

Source memory accuracy, the ability to classify a recognized face as previously having been chosen or not, was found to be adversely affected by the choice blindness manipulation. This lower accuracy during manipulated trials is primarily attributable to participants' performance during non-detected trials. In fact, strikingly, for these trials, participants performed worse than they would have, had they been randomly guessing! This provides evidence that, not only are later choices affected by choice blindness manipulations, but also explicit memories about those choices. This is significant because on alternative accounts of what drives preference change through choice, those effects arise through post-choice preferential adjustment which are not linked to beliefs or later memories.

Considering the types of responses participants made, revealed that the differences in source memory accuracy were due to misattributing the non-chosen face as chosen. We found that participants responded with a positive response to a much higher degree during non-detected trials compared to detected trials. This means that participants often thought the recognized face also was their chosen, which in turn translates into the higher False Positive rate for non-detected trials. Participants' source memory is thus selectively distorted as a result of accepting the false feedback, even though their overall capacity to recognize previous encountered stimuli is largely unaffected. One

objection might be that participants do not process the faces enough for veridical source monitoring to occur. However, we observed high source accuracy in the non-manipulated trials suggesting that the experimental task is sufficient in the regard. Nevertheless, future studies could test this further by varying the stimulus and task, beyond facial preference, to generalize and expand on the findings presented here

Human visual long-term memory is highly detailed and capable of retaining a large number of details for thousands of newly acquired objects (Brady, Konkle, Alvarez & Oliva, 2008). Given this, the memory task in the present study was relatively simple and the high recognition task accuracy reflects this. On the other hand, source monitoring has been hypothesized to be underpinned by evaluative and reconstructive processes distinct from those involved in recall and recognition (Johnson, Hashtroudi & Lindsay, 1993; Yonelinas, 1999). One component which has been implicated in studies of memory misattribution for choices is self-perception, that is, the general finding that attitudes and preferences are constructed partly on the basis of external cues (Mather, Shafir & Johnson, 2000; Bem, 1967). Similarly, in the recent literature on preference change through choice, different experimental paradigms involving blind choices and choice blindness have suggested that beliefs mediate preference shifts (Egan, Bloom & Santos, 2010; Sharot, Velasquez & Dolan, 2010; Johansson et al., 2014; Pärnamets et al., 2015b). Together with these findings, this suggests that common mechanisms might be involved in how both preferences and memories are affected post-choice.

Better understanding how beliefs about past choices shape memories for past options is an important step towards understanding the mechanisms of preference change as well as how interactions and feedback from the environment shape these processes. The current results indicate high degrees of flexibility for source attributions coexisting with stable memories of previously encountered options.

Acknowledgments

We thank S. Sikström and C. Balkenius for helpful discussions. PJ was supported by Bank of Sweden Tercentenary Foundation and The Swedish Research Council (2014-1371). LH was supported by Bank of Sweden Tercentenary Foundation (P13-1059:1) and The Swedish Research Council (2011-1795).

References

Bem, D.J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183-200.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38), 14325-14329.

Egan, L. C., Bloom, P., & Santos, L. R. (2010). Choice-induced preferences in the absence of choice: Evidence

from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology*, 46(1), 204-207.

Hall, L., & Johansson, P. (2008). Using choice blindness to study decision making and introspection. In P. Gärdenfors, & A. Wallin (Eds.), *A Smorgasbord of Cognitive Science*. Nora, Sweden: Nya Doga.

Henkel, L. A., & Mather, M. (2007). Memory attributions for choices: How beliefs shape our memories. *Journal of Memory and Language*, 57(2), 163-176.

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, 310(5745), 116-119.

Johansson, P., Hall, L., Tärning, B., Sikström, S., & Chater, N. (2014). Choice blindness and preference change: You will like this paper better if you (believe you) chose to read it!. *Journal of Behavioral Decision Making*, 27(3), 281-289.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3-28.

Reynolds, J. K., & Pezdek, K. (1992). Face recognition memory: The effects of exposure duration and encoding instruction. *Applied Cognitive Psychology*, 6(4), 279-292.

Mather, M., Shafir, E., & Johnson, M. K. (2000). Misremembrance of options past: Source monitoring and choice. *Psychological Science*, 11(2), 132-138.

Pärnamets, P., Hall, L., Strandberg, T., Balkenius, C., & Johansson, P., (2015a). *Looking at choice blindness: Evidence from eye-tracking and pupil dilation*. Manuscript submitted for publication.

Pärnamets, P., von Zimmermann, J., Hall, L., Raafat, R., Chater, N., & Johansson, P. (2015b). *Collective choice blindness and preference change through choice*. Manuscript submitted for publication.

Sagana, A., Sauerland, M., & Merckelbach, H. (2014). Memory impairment is not sufficient for choice blindness to occur. *Frontiers in Psychology*, 5(449). doi: 10.3389/fpsyg.2014.00449

Sauerland, M., Schell, J. M., Collaris, J., Reimer, N. K., Schneider, M., & Merckelbach, H. (2013). "Yes, I Have Sometimes Stolen Bikes": Blindness for Norm-Violating Behaviors and Implications for Suspect Interrogations. *Behavioral Sciences & the Law*, 31(2), 239-255.

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science*, 21(9), 1231-1235.

Yi, D. J., Turk-Browne, N. B., Chun, M. M., & Johnson, M. K. (2008). When a thought equals a look: Refreshing enhances perceptual memory. *Journal of Cognitive Neuroscience*, 20(8), 1371-1380.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1415-1434.