# Correction of Manipulated Responses in the Choice Blindness Paradigm: What are the Predictors?

**Thomas Strandberg (thomas.strandberg@lucs.lu.se)**
**Lars Hall (lars.hall@lucs.lu.se)**
**Petter Johansson (petter.johansson@lucs.lu.se)**
Lund University Cognitive Science, Lund University, Box 192, S-221 00 Lund, Sweden
**Fredrik Björklund (fredrik.bjorklund@psy.lu.se)**
Department of Psychology, Lund University, Box 213, S-221 00 Lund, Sweden
**Philip Pärnamets (philip.parnamets@ki.se)**
Department of Clinical Neuroscience, Div. of Psychology, Karolinska Institute
Solnavägen 1, S-17177 Solna, Sweden

**Abstract**

Choice blindness is a cognitive phenomenon describing that when people receive false feedback about a choice they just made, they often accept the outcome as their own. Little is known about what predisposes people to correct manipulations they are subjected to in choice blindness studies. In this study, 118 participants answered a political attitude survey and were then asked to explain some of their responses out of which three had been manipulated to indicate an opposite position. Just over half (58.4%) of the manipulations were corrected. We measured extremity, centrality and commitment for each attitude, and one week prior to the experiment we assessed participants' preference for consistency, need for cognition and political awareness. Only extremity was able to predict correction. The results highlight the elusiveness of choice blindness and speak against dissonance and lack of motivation to engage in cognitively demanding tasks as explanations why the effect occurs.

**Keywords:** choice blindness; attitude change; attitude strength; need for cognition; preference for consistency; political awareness.

## Introduction

Choice blindness (CB) is a cognitive phenomenon indicating a dissociation between making a choice and its later justification. It highlights the limitations of our introspective capacity when reasoning about past choices. CB occurs when people receive false feedback about a choice they just made accepting the outcome as their own and reporting seemingly introspective (albeit confabulated) reasons for having made that choice (see Johansson et al., 2005 for details). CB has been reported for many domains and modalities, ranging from taste and smell preferences (Hall, Johansson, Tärning, Sikström & Deutgen, 2010) to eye-witness testimony (Cochran, Greenspan, Bogart & Loftus, 2018), and has been shown to affect both later memories and preferences (e.g. Strandberg, Sivén, Hall, Johansson & Pärnamets, 2018; Pärnamets, Hall & Johansson, 2015; Johansson, Hall, Tärning, Sikström & Chater, 2014). CB has also been applied to the study of attitudes and attitude change, an area of research where deliberation and introspection are often seen as important ingredients. In Hall, Johansson and Strandberg (2012) about 60% of manipulations to a survey on moral dilemmas were accepted by the participants' as being their own attitudes. Hall et al., (2013) reported similar findings for salient political issues in the run up for a Swedish general election. In that study participants not only changed their attitudes on political issues, but their actual voting intention was also affected in the direction of the false feedback. Notably, Strandberg and colleagues (2018) found that when participants accepted the manipulations to political attitudes, these shifted congruently with the false feedback when re-elicited one week later. Although CB is ubiquitous, and undeniably relevant for the study of attitudes and decisions, little is known about what factors that predisposes people to correct the manipulated responses. So far, only a few studies have attempted to establish CB mediators, and thereby link the effect to other psychological constructs (e.g. Strandberg et al., 2018). However, no studies have focused purely on why people correct the false feedback. In this study, we aim to explore several factors that we have identified as meaningful for understanding why correction in the CB paradigm occurs, particularly in the domain of attitudes.

### Subjective experience of attitude strength

One possible key to CB susceptibility could be in the relationship between the individual and the attitude itself. This is supported by the literature describing strong attitudes as "resistant to change, persuasion, and contextual influence" and weak attitudes as "unpredictable, malleable, and created in the moment" (Krosnick & Petty, 1995). Given this definition, it seems reasonable that correction of manipulations to attitudes should correlate with attitude strength. Here we tested three self-report measures adopted from Bassili's (1996) seminal work on attitude strength: extremity, centrality and commitment. Extremity directly estimates how strongly a person agrees with an issue on a bipolar scale. Extremity, which is basically just the response to the survey item, is what Bassili calls an operative measure based on first order cognitive processing. Extremity is operative because, for example, the

experienced valence of the extremity could be directly retrieved from memory and not the product of inference. Centrality and commitment, on the other hand, are so called meta-attitudes. These are second order impressions of attitudes that rely on people to report on psychological properties not necessarily represented in long-term memory. As such, meta-attitudes are often inferred from sources more or less relevant to the strength of which the attitude is held. Centrality is described as tapping into the importance of an attitude and how it relates to personal values. Studies show that central attitudes are often more memorable and resistant to persuasion and contextual influence compared to peripheral attitudes (Holland, 2003; Pomerantz et al., 1995). Commitment is described as tapping into the confidence in an attitude: the conviction that the attitude is correct and valid. Commitment has been shown to moderate self-perception and contextual influence in attitudes (Holland, 2003; Pomerantz et al., 1995). Since these measures are meant to capture attitude strength – with strong attitudes being defined by their "resistance to change, persuasion, and contextual influence" – they should also correlate with correction of CB manipulations.

### Variation in cognitive style

Another possibility is that aspects of the CB task might be experienced as rather cognitively demanding, such that some individuals may be more susceptible to CB than others due to being less motivated to perform them. Previous studies have shown that individuals with a larger set of general analytic skill are more prone to correct the manipulations (Strandberg et al., 2018). Hence, measures capturing peoples' motivation to engage in cognitively demanding task, such as the *Need for Cognition* (NC; Cacioppo, Petty & Kao, 1984; Cacioppo, Petty, Feinstein & Jarvis, 1996) might also correlate with correction. NC is commonly used in attitude change research, where studies have shown that people with high NC tend to form attitudes that are more resistant to persuasion compared to people with low NC (Haugtvedt & Petty, 1992). CB could also be affected by a consistency motive, which is the case for dissonance phenomena such as cognitive dissonance, cognitive balance, foot-in-the-door etc. These phenomena show that people often change either their behavior or their attitudes to appear consistent (cf. Festinger, 1957). One measure for estimating peoples' need to have consistent cognitions is *Preference for Consistency* (PFC; Cialdini, Trost & Newsom, 1995). Further, PFC has also been shown to predict if people change their attitudes due to social pressure or external demand (Bator & Cialdini, 2006). Thus, if CB share properties with cognitive dissonance phenomena; or if participants accept manipulations due to demand from the experimental situation, correction may correlate with the PFC score.

### Variation in political awareness

We would also like to consider variation in political awareness, since much research in political science highlights political awareness as one of the most important factors when forming strong and resilient political attitudes (Zaller, 1992). Interestingly, recent CB studies involving political attitudes have yielded mixed results. In Hall et al. (2012) politically involved participants were more likely to correct the manipulations, and this was not found in Strandberg et al. (2018). However, since political awareness is supposed to determine how people select, interpret and internalize political information (Sidanius, 1988; Lusk & Judd, 1988) we continue to explore the relationship between various measures of political awareness and participants' behavior in a CB study involving political issues.

Thus, we set out to test if susceptibility to correct manipulated responses in CB could be predicted by any of the attitude strength measures, variation in cognitive style, or political awareness described above.

## Method

### Participants

A total of 128 (70 female) participants, with ages ranging from 18 to 64 years ($M$ = 23.5, $SD$ = 16.8), were recruited to answer a political survey. Sample size was predetermined based on previous CB studies (e.g. Johansson et al. 2005). Ten participants were excluded due to malfunctions with the experimental equipment. Thus, 118 participants remained for the final analysis. The participants were recruited through posters and flyers distributed at the university campuses of Lund and Malmö and compensated with a cinema voucher. At the start of the experiment, we described the general purpose of the study, but without telling the participants that some of their answers would be manipulated. Participants were informed that they could quit the experiment at any time, request their data to be erased, and still receive the cinema voucher. Participants were fully debriefed at the end of the experiment, before consenting to their anonymized data to be used by signing a consent form. All but six participants allowed their interviews to be recorded (leaving a total of 112 verbal recordings to be analyzed). The study was approved by the Lund University Ethics board, D.nr. 2008–2435.

### Materials and design

**Pre-test** One week before the main experiment, participants completed an online questionnaire assessing their demographics, political awareness, PFC and NC. PFC was assessed using the abbreviated 9-item version (Cialdini et al., 1995) with scales ranging from 1 (low consistency) to 9 (high consistency). The PFC questionnaire assessed the participants' internal and external consistency and included items such as: "It is important to me that my actions are consistent with my beliefs". For NC, we used the 18-item version (Cacioppo, Petty & Kao, 1984) with scales ranging from 1 to 9 where a nine gave four points and a one subtracted four points (five gave zero points, and so on).

The NC questionnaire assessed the participants' attitudes towards effortful thinking, and contained items such as: "I usually end up deliberating about issues even when they do not affect me personally". Further, political awareness was established by assessing the participants' political interest with a scale ranging from extremely uninterested (1) to extremely interested (9), and whether they were involved in any political party or organization (yes/no). Visit *https://osf.io/zsy47/* for a list of all measures and items.

**Main experiment** After the pre-test, participants scheduled to partake in the main experiment being held one week later. It consisted of a questionnaire running on a tablet with a touch-based interface that the participants interacted with using a tablet pen. The experiment consisted of two parts: (1) responding to political issues, (2) explaining the responses, and ended with a full debriefing.

## Procedure

**Part 1 – responding to political issues** During the first part, participants responded to 12 sets of political issues with each set containing a political statement and corresponding six meta-attitudes; three centrality, such as "how important is this issue to you?", and three commitment, such as "how confident are you about your attitude towards this issue?" (visit *https://osf.io/zsy47/* to see all centrality and commitment items). The political issues were selected together with leading political scientists, and represented 12 of the most salient and important issues in Sweden at the time of the study (Table 1). As such, we believe that the vast majority of our participants were familiar with them. This was also confirmed by the verbal reports: most participants were able to intelligibly and knowingly discuss the various issues. Below each item were visual analog scales with endpoints at 0 and 100 (completely disagree to completely agree for the political statements and for example extremely unimportant to extremely important for the centrality item "importance"). The participants were instructed respond to each item by drawing a mark using the pen. They could change their responses as many times as they wanted by clicking a change icon located to the left of each scale, as well as toggle freely between the 12 sets of issues. The participants were left to complete the questionnaire at own, and told to inform the experimenter when finished.

**False feedback and correction** When going over and explaining the responses, participants had received false feedback on three of the six trials. Trials 2, 4 and 6 had been manipulated by the tablet application to indicate a position opposite to the original (Figure 1). Trials 1, 3 and 5 were non-manipulated controls. The manipulation had two rules: move the participants' rating across the midline of the scale (with a minimum of 5 mm from the middle, i.e. ratings 45 or 55), and then randomly positioned on the opposite axis. If participants in any way indicated that their responses did not correspond with their views, or indicated that something was wrong, the experimenter would tell them that they could change their response if they wanted to, after which they could base their explanation on that response instead. Correction was operationalized when change was clicked and a new response drawn.

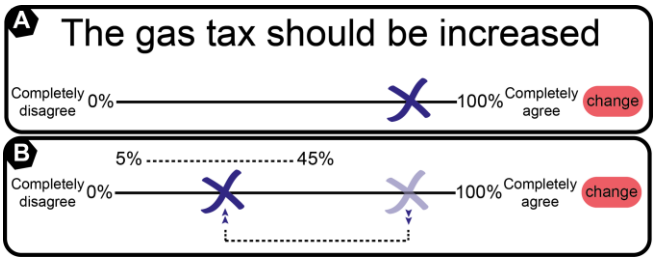


Figure 1: To respond, participants drew an X on a scale going from completely disagree to completely agree (A). On manipulated trials, participant's X was surreptitiously moved from one side of the scale and then randomly placed on the other side (B). Participants could change their X as many times as they wanted by clicking 'change' (A-B).

Table 1: The political issue statements.

| | |
|---|---|
| 1. | The gas tax should be increased |
| 2. | A wealth tax should be reinstated |
| 3. | The labor taxes should be lowered |
| 4. | The monarchy should be abolished |
| 5. | The government should run all elementary schools |
| 6. | The punishment for violent crimes should be stricter |
| 7. | The subsidized service for homework assistance should be abolished |
| 8. | High schools should offer more applied and fewer theoretical courses |
| 9. | Women should be recruited to company boards through affirmative action |
| 10. | Private health care companies should be allowed to make profits in the welfare sector |
| 11. | Copyright protected material from internet should be free to download for personal use |
| 12. | The government should be allowed to monitor telephone conversations and internet traffic |

## Analysis

Consistent with Bassili (1996) extremity was calculated by taking the absolute value of the deviation between a rating on the 100 point scale and the midpoint. All other variables are reported using their averages. Since attitude extremity, and the difference between the original rating and the manipulated rating, labeled 'manipulation length', are core features in CB studies using rating scales; we first tested how well these would predict correction. In our dataset, extremity and manipulation length were highly correlated, $r$ = .73, $t_{(333)}$ = 19.6, $p$ = 2.2*10$^{-16}$. To address this we performed our analyses using decorrelated variables by transforming manipulation length to be the distance on the scale the manipulated attitude was moved *beyond* the midpoint. The resulting variables were independent, $r$ = -.028, $t_{(333)}$ = -0.52, $p$ = .61. We then used these two variables to fit a baseline for the other predictor variables (i.e. meta-attitudes and cognitive style). We analyzed our data using mixed regression models including by participant varying intercepts and slopes. Models were estimated in a Bayesian framework using the brms package in R (Bürkner, 2016). Weakly regularizing priors were used for all parameters.

# Results

On average participants were moderately interested in politics ($M$ = 6.0, $SD$ = 2.1) and about one fifth identified as politically involved ($M$ = 22.9, $SD$ = 42.2). As we can see in Table 2, extremity, centrality and commitment was rated fairly strong, averaging between 60 to 65 points of 100. The PFC score in our sample was similar to the 48.9 ($SD$ = 10.7) that Cialdini et al. (1995) reported, and the NC score was similar to that reported in a recent meta-analysis of the NC scale ($M$ = 33.2, $SD$ = 10.2 (de Holanda & Wolf, 2018)).

Table 2: Means and SD for the main predictor variables.

| Predictor | Mean | SD |
|---|---|---|
| Extremity | 29.2 | 13.9 |
| Centrality | 63.9 | 18.2 |
| Commitment | 65.0 | 20.1 |
| NC | 29.1 | 17.8 |
| PFC | 44.8 | 12.6 |

## False feedback correction

Participants corrected 58.4% of the total 347 manipulations. Each participant was exposed to three manipulations and the average correction rate was 1.66 ($SD$ = 0.98), with 15 participants accepting all manipulations and 27 participants correcting all. After correcting a manipulation participants were instructed to replace it with a new response. This corrected rating was on average placed within 9.43 points ($SD$ = 11.7) of their original rating; or -4.45 points ($SD$ = 14.4) when taking the direction of the corrected rating into account (defining a weakened new rating as a negative quantity and a strengthened new rating as a positive quantity). As in previous CB studies, correction did not vary as a function of sex, gender, age, or political party.

## Predictors of correction

To test for predictors of correction we conducted mixed-effects logistic regression analyses using standardized variables. We first fit a baseline model consisting of extremity and manipulation length. This model (LOO = 402.77, $SE$ = 14.86) indicated a large effect of extremity on correction (β = 1.77, $SD$ = 0.32, *95% CI* = [1.17, 2.43], $BF_{10}$ > 1.0*10^5), but only a smaller, uncertain effect of manipulation length (β = 0.53, $SD$ = 0.28, *95% CI* = [-0.0043, 1.09], $BF_{10}$ = 1.67), with the intercept estimated as β = 0.46 ($SD$ = 0.19, *95% CI* = [0.10, 0.84]). See Figure 2 for the marginal posterior predictions of the attitude extremity and manipulation length.

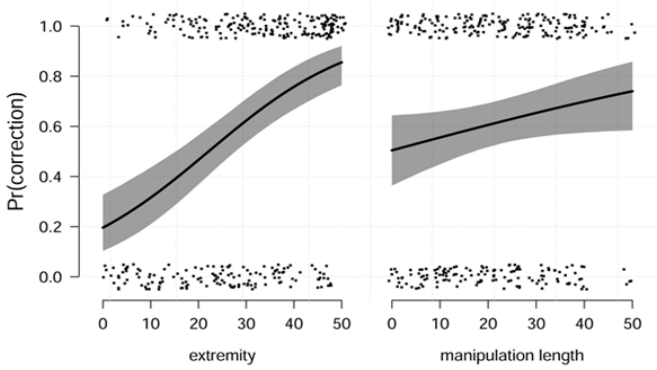


Figure 2: Marginal posterior predictions from the baseline model. Predictions assume other variable held at its average value (0 for standardized predictors). X-axes renormalized to increase interpretability. Shaded regions indicate 95% posterior intervals.

We next fit a full model with all our candidate predictors: extremity, centrality, commitment, preference for consistency (PFC), need for cognition (NC), political involvement, political interest and manipulation length (LOO = 401.65, SE = 17.03). The estimated coefficients, their credible intervals and associated Bayes Factors can be found in Table 3. Marginal posterior predictions are depicted in Figure 3. Notably, when comparing the baseline and full model using LOO we found that the baseline model and the full model did not differ, with a difference of 1.12 (SE = 6.23), this is also mirrored in the estimates where there is little evidence that any of the added predictors are particularly successful at estimating correction.

## Predictors of correction types

On an exploratory note, we tried to better capture participants' subjective experience of correcting a manipulation. We conducted a simple classification of the reasons participants reported for wanting to correct.

Table 3: Estimates and Bayes Factors from the full model.

| Predictor | Est (β) | SD | 95% CI | $BF_{10}$ |
|---|---|---|---|---|
| (Intercept) | 0.51 | 0.21 | [0.12, 0,94] | - |
| Extremity | 1.34 | 0.38 | [0.06, 2.10] | 333.69 |
| Centrality | 0.44 | 0.39 | [-0.31, 1.23] | 0.71 |
| Commitment | 0.69 | 0.39 | [-0.06, 1.46] | 1.78 |
| NC | -0.07 | 0.40 | [-0.86, 0.72] | 0.40 |
| PFC | -0.12 | 0.37 | [-0.85, 0.59] | 0.38 |
| Pol.Involvement | 0.71 | 0.47 | [-0.21, 1.62] | 1.45 |
| Pol.Interest | -0.22 | 0.43 | [-1.06, 0.65] | 0.49 |
| Manip.Length | 0.59 | 0.31 | [-0.0013, 1.20] | 1.94 |

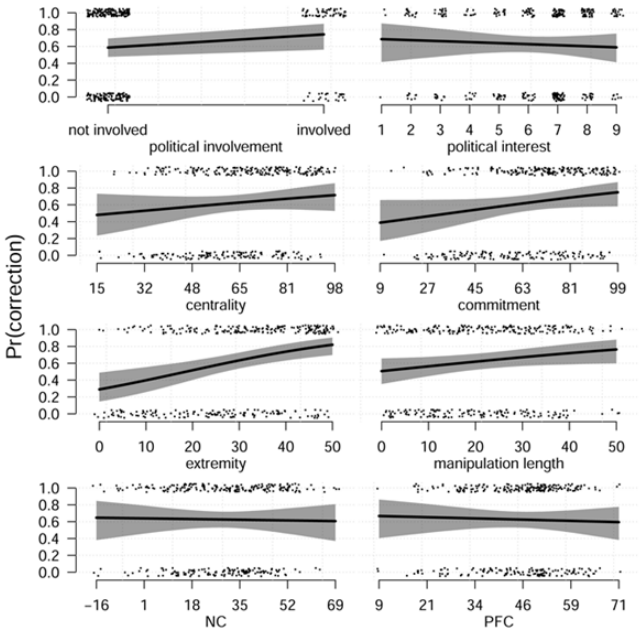


Figure 3: Marginal posterior predictions from the full model presented with the same properties as Figure 2.

One independent rater listened to all the 112 recorded interviews and coded the different reasons participants gave when correcting a manipulation. We identified three distinct types distributed evenly among the corrections: *internal attribution* (36.8%), when participants claimed to have misinterpreted the question, the scale, or something in the task; *external attribution* (33.9%), when participants blamed the experimental equipment; and *change* (29.2%), when participants felt they had spontaneously changed their minds about the issue. Only for a few trials did participants report suspicion that their responses had been manipulated; these were categorized as external attribution. A second rater then classified a subset of 40 interviews; the raters agreed on 90% of the classifications. To test for determinants of the correction types we conducted a hierarchical multinomial logistic regression analysis using correction type as dependent variable and the predictors used in previous analyses. Since we were mainly interested in whether people attributed the wish to correct internally or externally, the change category was used as the reference level in the analysis. Consistent with previous findings, most variables were unable to predict whether participants would attribute correction internally (e.g. feeling that they had made a mistake) or externally (e.g. blaming the experimental equipment). However, we did find that the larger absolute difference between the original response and the manipulated response the more likely participants were to attribute correction internally (β = 2.40, *SD* = 0.57, *95% CI* = [1.03, 3.57], $BF_{10}$ = 1017.52). We also found a small negative effect of political involvement, meaning that participants that were uninvolved politically were more likely to attribute correction externally (β = -1.15, *SD* = 0.77, *95% CI* = [-2.63, 0.37], $BF_{10}$ = 2.36). However, the effect size of this latter finding was very small, but could potentially be a subject for future research.

## Discussion

To summarize, we first assessed participants' preference for consistency, need for cognition, and political awareness; and one week later measured attitude extremity, centrality and commitment on a questionnaire containing 12 political issues. Participants were then asked to explain their responses to six of these issues out of which three had been manipulated to indicate the opposite position using the Choice Blindness Paradigm. Just over half of the manipulations were corrected by the participants, meaning that the remaining was accepted by the participants as being their own attitudes. This is similar to previous CB studies on political attitudes (Strandberg et al. 2018; Hall et al. 2012).

### Attitude strength

In this study we were particularly interested in testing potential underlying factors that predisposes participants to correct the manipulations. We found that correction was mainly predicted by attitude extremity; meaning that the stronger participants agreed with an issue on the bipolar scale, the more likely they were to correct it. That attitude extremity correlates with correction is also in line with previous CB research (Strandberg et al. 2018; Hall et al. 2012; 2013) and corresponds with for example Bassili's (1996) findings on the relationship between extremity and attitude stability. However, surprisingly, the two meta-attitudes centrality and commitment did not contribute to the correction prediction. One possible explanation to this could be that operative measures of attitude strength, such as extremity, are more relevant to the task compared to second order impressions such as centrality and commitment. Bassili (1996) suggested that extremity is closely associated with the cognitive processing involved in attitude formation and retrieval which is two main components in a CB task. Centrality and commitment on the other hand rather tap into more abstract concepts of the attitude structure (Holland, 2003) not necessarily relevant for scrutinizing one's own survey responses. It could also be that higher extremity is the product of deeper and more involved elaboration,

making those responses more salient and memorable (Petty & Cacioppo, 1986). These results highlight the difficulties in assuming an attitude's strength and stability based on seemingly relevant self-report measures.

## Individual difference and cognitive style

The two measures of cognitive style, preference for consistency and need for cognition, were also not able to predict correction.

**Preference for consistency** In the case of PFC (Cialdini, Trost & Newsom, 1995), we interpret this as an indicator that the correction of CB manipulations is not based on consistency motives or social influence. Further, PFC is mainly about people self-monitoring and being aware about their own consistency; whereas CB corrections tend to occur outside of the participants' awareness. This could be seen in the reasons people reported when wanting to correct: they were almost exclusively about having made a mistake, detected a glitch in the survey application, or having spontaneously changed their minds. Importantly this result also distinguishes CB from cognitive dissonance (Festinger, 1957) and other consistency phenomena that are typically highly correlated with PFC. This is useful when discussing CB and its consequences in a larger theoretical context.

**Need for cognition** NC (Cacioppo, Petty & Kao, 1984) is often used in social psychology research for its supposed implications to people's attitudes, judgments and decisions. In this literature, NC is described as associated to peoples' tendency to process information and form elaborated and coherent attitudes. Because of this, attitudes of individuals high in NC should be more resilient to change, persuasion, and context effects (e.g. Haugtvedt & Petty, 1992). This is not what we found in this study. However, while individuals high in NC tend to be more resistant to various biases, previous research argue that even these individuals can be influenced if the bias is very subtle (Cacioppo, Petty, Feinstein & Jarvis, 1996). The subtlety factor might help explain why NC and CB correction did not correlate. Further, people with low NC can perform at a comparable level to those with high NC given enough external motivators. One such motivator could be the perception of what participants believe to be their own survey response.

**Political awareness** The two political awareness measures (political interest and involvement) also did not correlate with correction. While there is nothing uniquely special to *political* awareness per se, the awareness part addresses a domain specific aspect that could determine the participants' understanding, knowledge, and vested interest about the current CB theme (Zaller, 1992). For example, one previous CB study did find that political involvement correlated with correction (Hall, Johansson & Strandberg, 2012), and in this study we found a tendency (albeit small) that politically involved participants were more likely to attribute the correction externally (e.g. believing that there was some error with the equipment). This tendency at least indicates that politically involved participants experienced the false feedback differently from the uninvolved. It could simply be that politically involved individuals have stronger convictions in the politically attitudes; so when they notice a discrepancy between their original and present response, their main explanation is that software application malfunctioned.

## Limitations and future studies

The main limitation of this study was the small number of participants. While we only found a relationship between correction and attitude extremity, the lack of relationship between the other variables might at least be partially explained by the small sample size. Thus, one interesting avenue of future research would be to more systematically, and with more participants, test how a variety of attitude, personality, and performance measures affect correction rates and correction types. This would also allow us to examine subgroups within our sample; for example: what is it that makes some participants correct all the manipulations and some accept all? Importantly, while we found no relationship between correction and any of the two motivated cognition measures (NC and PFC), other more performance based variables might be relevant to CB and worth exploring. For example, in Strandberg et al. (2018) the Cognitive Reflection Test (CRT) correlated with correction, with participants having higher CRT score also being more likely to correct the manipulations. CRT is a performance based cognitive processing measure that captures peoples' ability to use reflective and deliberative thinking instead of gut feelings (Frederick, 2005). Thus, future research could try to link CB to performance based measures that taps into working memory, attention, or perhaps factual knowledge. Another potential shortcoming of this study was that the majority of the participants were students. Although we have no reason to believe, given previous studies, that a phenomenon such as CB would drastically differ between different demographics, it is always important to establish whether the experimental findings generalize across the public. However, similar levels of correction have been found in experiments with a more diverse and representative sample (Strandberg, Olsson, Hall, Woods & Johansson, in preparation).

## Conclusion

Choice blindness is a cognitive phenomenon powerful enough to influence peoples' opinions and reasoning in important political issues. Still, it is difficult to pinpoint what disposes people to accept or correct the manipulations. It seems that the CB manipulation is so surreptitious that it sometimes flies under the radar even for people with strong convictions and motivations to engage in political reasoning. This study contributes to the understanding of CB, serving as both a backdrop for future research, and an important piece of a broader theoretical puzzle.

## Acknowledgments

We would like to thank Anders Lindén at AndVision for implementing the experimental design to a tablet interface, Henrik Ekengren Oscarsson at the University of Gothenburg for helping to select the political issues, and Uno Otterstedts Fund (RFh2014-0390) for financing the movie tickets.

## References

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

Bator, R. J., & Cialdini, R. B. (2006). The nature of consistency motivation: Consistency, aconsistency, and anticonsistency in a dissonance paradigm. *Social Influence, 1*, 208–233.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology, 71*(4), 637-653.

Bürkner, P. C. (2016). brms: Bayesian regression models using Stan. *R package version 0.10. 0*.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306-307.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*, 197-253.

Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology, 69*(2), 318-328.

Cochran, K. J., Greenspan, R. L., Bogart, D. F., & Loftus, E. F. (2018). (Choice) blind justice: Legal implications of the choice blindness phenomenon. *University of California, Irvine Law Review 8*, 85.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. California: Stanford University Press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

Hall, L., Johansson, P., Tärning, B., Sikström, S., & Deutgen, T. (2010) Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition, 117*, 54–61.

Hall, L., Johansson, P., & Strandberg, T. (2012). Lifting the veil of morality: Choice blindness and attitude reversals on a self-transforming survey. *PLoS One, 7*(9), e45457.

Hall, L., Strandberg, T., Pärnamets, P., Lind, A., Tärning, B., & Johansson, P. (2013). How the polls can be both spot on and dead wrong: Using choice blindness to shift political attitudes and voter intentions. *PLoS One, 8*(4), e60554.

Haugtvedt, C. P., & Petty, R. E. (1992). Personality and persuasion: need for cognition moderates the persistence and resistance of attitude change. *Journal of Personality and Social Psychology, 63*(2), 308-319.

de Holanda, G. L., P., P. H., & Wolf, L. J. (2018). The Very Efficient Assessment of Need for Cognition: Developing a Six-Item Version. *Assessment*.

Holland, R. W. (2003). *On the structure and consequences of attitude strength.* Dissertation Thesis, Radboud University, Nijmegen.

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

Johansson, P., Hall, L., Tärning, B., Sikström, S., & Chater, N. (2014). Choice blindness and preference change. *Journal of Behavioral Decision Making, 27*(3), 281–289.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Ohio State University series on attitudes and persuasion, Vol. 4*. Attitude strength: Antecedents and consequences (pp. 1–24). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lusk, C. M., & Judd, C. M. (1988). Political expertise and the structural mediators of candidate evaluations. *Journal of Experimental Social Psychology, 24*(2), 105-126.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205.

Pomerantz, E. M., Chaiken, S., & Tordesillas, R. S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology, 69*(3), 408–419.

Pärnamets, P., Hall, L., & Johansson, P. (2015). Memory distortions resulting from a choice blindness task. In: *Proceedings of the 37th Annual Conference of the Cognitive Science Society*.

Sidanius, J. (1988). Political sophistication and political deviance: A structural equation examination of context theory. *Journal of Personality and Social Psychology, 55*(1), 37-51.

Strandberg, T., Sivén, D., Hall., L., Johansson, P., & Pärnamets, P. (2018) False beliefs and confabulation can lead to lasting changes in political attitudes. *Journal of Experimental Psychology: General, 147*(9), 1382-1399.

Strandberg, T., Olson, J.A., Hall, L., Woods, A.T., Johansson, P. (manuscript in preparation).

Zaller, J. R. (1992). *The nature and origins of mass opinion*. New York: Cambridge University Press.