

I could have done otherwise: availability of counterfactual comparisons informs the sense of agency

Eugenia Kulakova, Nima Khalighinejad, Patrick Haggard
Institute of Cognitive Neuroscience, University College London, London, UK

Corresponding author:
Eugenia Kulakova
Institute of Cognitive Neuroscience
University College London
Alexandra House, 17 Queen Square
WC1N 3AZ London, United Kingdom
E-mail: e.kulakova@ucl.ac.uk
Phone: 020 7679 25570

Abstract

Personal control and agency are closely associated with the counterfactual notion that a person *could have done otherwise* (CDO). In both philosophy and law, this counterfactual evaluation determines responsibility and punishment, yet little is known about its influence on agents' experience during action. We used a risky decision-making task to study how counterfactual evaluations influenced participants' sense of agency. Two factors were manipulated independently: the presence/absence of counterfactual comparisons between actions and the presence/absence of counterfactual comparisons between outcomes of these actions. Perceived agency was highest when both counterfactual comparisons were available. Interestingly, this pattern persisted even when counterfactual information was only revealed after action, suggesting a purely reconstructive evaluation effect. These findings allow a more precise phrasing of the CDO element of personal agency: a person feels most control when she *could have performed another action, thereby obtaining another outcome*.

1. Introduction

The sense of agency (SoA) refers to the feeling of controlling events in the outside world by one's own action (Haggard & Tsakiris, 2009). This notion of personal control is associated with the feeling of being a causal agent who actively shapes the world (Frith, 2014). One particularly important cue to such notion of personal causality is counterfactual thinking – the mental evaluation of possible but unrealised alternatives (Lewis, 1986). According to the counterfactual theory of causation, an action causes an outcome, if both action and outcome occur, and the outcome would not have occurred without the action (Mackie, 1974). Hence, establishing a causal link between two events requires a counterfactual evaluation. This suggests that counterfactual comparisons might play an important role for agency attributions, as sense of agency potentially reduces to ascription of causation to oneself (de Biran, 1841). Despite this close link between agency, causality and counterfactual thought, the direct influence of counterfactual thinking on the subjective experience of agency during action remains poorly studied (but see Dannenberg, Förster & Jostmann, 2012 on the influence of counterfactual thinking on illusory authorship and Desantis, Roussel & Waszak on a link between causation and sense of agency). If sense of agency relies on attributions of personal causation derived from counterfactual thinking, then the availability of counterfactual comparisons should directly increase SoA, even if such comparisons have no effect on objective control, that is the likelihood to obtain the desired outcome. The present study used a risky decision-making task to investigate how comparisons between factual and counterfactual action and outcome alternatives influenced participants' sense of agency during active decision-making.

Counterfactual thinking refers to the mental evaluation of the way things might have happened, as opposed to how they actually did happen (Byrne, 2005). Counterfactuals simulate how changes to the past would have affected the subsequent course of events, potentially leading to a different present. Such thought experiments help to identify causal relations, especially when events have already occurred and will not re-occur in exactly the same way (Byrne, 2016). With regard to agency and personal control, the counterfactual thought "*I could have done otherwise*" (CDO) holds particular theoretical importance (Frankfurt, 1988). Influential models of legal and moral reasoning consider the CDO evaluation as an important test whether somebody personally and deliberately caused an outcome, and consequently should be held responsible for it (Pereboom, 2011). The evaluation of CDO type counterfactuals therefore seems particularly important for the ascription of agency, as a satisfied CDO condition might directly increase the sense of agency. However, no previous study has investigated whether actions made under CDO conditions indeed produce an altered sense of agency compared to actions not so made.

One obstacle to the empirical investigation of CDO type comparisons is a rather vague meaning of "*could have done otherwise*". *Doing otherwise* can either refer to performing a different action, or obtaining a different outcome, or both. The CDO requirement can thus be interpreted as having counterfactual between-action alternatives ("*I could have performed another action*"), or refer to available between-outcome alternatives ("*I could have obtained another outcome*"). In the present study we developed a paradigm which allowed us to manipulate these two aspects independently. This allowed us to study how available action and/or outcome alternatives informed the sense of agency via a counterfactual evaluation of personal causality.

In an active choice task participants made risky decisions between two button presses which resulted in positive or negative outcomes, respectively. The availability of counterfactual between-action and between-outcome alternatives was introduced as two independent factors. Participants could either make free choices between two buttons, or were instructed to select one specified button. This contrast defined the factor of counterfactual action alternatives (*"I could have performed another action"*) at the point of action selection. Counterfactual outcome alternatives were introduced by informing participants whether the two actions would result in identical or opposite outcomes. The consideration *"I could have obtained another outcome"* was true for opposite outcomes, but false if the outcomes associated with both actions were identical. Counterfactual outcome alternatives were potentially available for free choices. However, the counterfactual outcome was not truly obtainable when the action was instructed.

In two experiments we further manipulated the time within the course of action at which counterfactual outcome information was provided. This allowed us to investigate whether purely retrospective (*"I could have obtained a different outcome (but now I no longer can)"*) or potentially prospective (*"I can obtain different outcomes"*) evaluations of outcome alternatives influenced sense of agency in a different way.

Importantly, participants had no objective control over the outcomes they obtained during the task because outcome assignment was random. In the absence of learnable action-outcome contingencies counterfactual comparisons could not be used to direct action choices, thereby preventing any adaptive preparatory function of counterfactual evaluations which could increase participant's task success (Epstude & Roese, 2008). Furthermore, our task did not explicitly instruct participants to generate or evaluate counterfactuals, unlike most previous studies of counterfactual thinking and reasoning (e.g., Kahneman & Miller, 1986). Instead, we reasoned that simply providing information about un-selected or un-obtained alternatives would trigger implicit counterfactual comparisons. Previous findings suggest that such implicit tracking of counterfactual alternatives does indeed occur. For example, the valence of un-obtained gamble outcomes strongly influences satisfaction with the outcomes that were actually obtained (Camille et al., 2004; Zeelenberg et al., 1998). However, it remains unclear whether similar comparisons operate with respect to sense of agency evaluations.

The sense of agency was assessed with explicit and implicit measures. Explicit SoA was measured by explicit ratings of degree of control (Chambon et al., 2013), while implicit SoA was assessed with an intentional binding measure (Engbert, Wohlschläger, & Haggard, 2008). Intentional binding refers to the effect that stronger sense of agency shifts action and outcome closer together in subjective time (Haggard, Clark, & Kalogeras, 2002). This leads participants to underestimate the temporal interval between personally controlled actions and their outcomes, while durations of action-outcome intervals between outcomes following externally triggered actions are in turn over-estimated (see Moore & Obhi, 2012 for review). Both explicit and implicit measures of agency have been widely employed in research on volition and agency, but have not previously been used to study counterfactual comparisons in decision-making tasks. While explicit measures reflect sense of agency judgments informed by contextual cues and propositional beliefs, intentional binding is assumed to arise from an integration of lower-level perceptual and motor cues which are less susceptible to cognitive bias (Synofzik, Vosgerau, & Newen, 2008).

We were particularly interested in the interplay between counterfactual action and outcome alternatives and the influence these manipulations would have on our measures of

sense of agency, as this relation would clarify how these two components of the “*could have done otherwise*” requirement affect the sense of personal control.

2. Experiment 1

2.1. Material and Methods

2.1.1. Participants

19 participants (12 females, average age 23 years) took part in Experiment 1. The sample size was determined by an a priori power analysis (G*Power; Faul, Erdfelder, Lang, & Buchner, 2007) in order to detect a strong ($f = .40$) significant main effect or interaction with the settings of power = .90 and alpha level = .05. Participants provided informed consent and received £10.50 for participation. The study was approved by the local ethics review board.

2.1.2. Procedure

Participants were playing a risky decision-making game on a computer (see **Figure 1**). Each trial presented a cartoon mouse facing two mouse holes and participants selected to which mouse hole the mouse would go. If both mouse holes had identical shades of grey, this indicated that they would have identical outcomes, whereas different shades of grey indicated opposite outcomes. In half the trials participants chose freely between the two mouse holes. In the remaining trials participants were instructed by a black circle that only one mouse hole could be selected. Participants pressed the left or right arrow on a keyboard with their right hand, corresponding to the left and right mouse holes, respectively. After a variable time interval (100/400/700 ms) the outcomes of both chosen and un-chosen actions were presented visually by a change in colour of the mouse holes. Two outcomes were possible: holes that turned green indicated a positive outcome, signifying that the mouse survived and participants were rewarded 20p. Red holes represented negative outcomes, indicating that the mouse died and participants lost 18p. A tone (500 Hz, 50 ms) accompanied the colour change. After outcome presentation, participants estimated the time interval that passed between their key-press and the tone, which constituted the interval estimation measure of intentional binding (for detailed methods see Caspar, Cleeremans, & Haggard, 2015). Finally, participants gave an explicit rating over the amount of personal control they perceived over the obtained outcome on a 9-point scale.

The task design consisted of three factors: action selection (free, instructed) × outcome valence (positive, negative) × counterfactual outcome (identical, opposite), giving a total of eight experimental conditions. Each condition was presented 36 times per participant, resulting in 288 trials in total. Trials were presented randomised in three blocks of 96 trials each. The probability of obtaining a positive or negative outcome was 50% throughout the whole experiment and did not differ between conditions. Restricting outcome probability to chance prevented participants from learning to predict outcomes and adjusting action selection accordingly. Counterfactual information thus could not assist task performance. All combinations of experimental condition, shadings of grey and time intervals were counterbalanced and appeared with equal frequency.

2.2. Results

2.2.1. *Explicit control ratings*

Explicit control ratings were entered in a $2 \times 2 \times 2$ repeated measures ANOVA (see **Figure 2**). Participants reported higher agency in free choice compared to instructed trials ($F(1,18) = 33.09, p < .001, \eta_p^2 = .65$). Furthermore, agency ratings were higher in the presence of opposite compared to identical counterfactual outcomes ($F(1,18) = 20.53, p < .001, \eta_p^2 = .53$). However, opposite counterfactual outcomes only increased agency in free ($t(18) = 4.68, p < .001$) but not instructed trials ($t(18) = 0$), as indicated by an action selection \times counterfactual outcome interaction ($F(1,18) = 22.52, p < .001, \eta_p^2 = .56$). Control ratings were generally higher following positive compared to negative outcomes ($F(1,18) = 21.05, p < .001, \eta_p^2 = .54$). This effect was stronger in free ($t(18) = 4.70, p < .001$) compared to instructed trials ($t(18) = 3.63, p = .002$), as indicated by an action selection \times outcome valence interaction ($F(1,18) = 4.80, p = .04, \eta_p^2 = .21$).

2.2.2. *Interval estimation*

Mean normalised interval duration estimations between action and outcome were entered into a $2 \times 2 \times 2$ repeated measures ANOVA (**Figure 2**). Time intervals were perceived as longer in free choice trials ($F(1,18) = 5.98, p = .025, \eta_p^2 = .25$). Opposite counterfactual outcomes also increased interval estimations ($F(1,18) = 15.24, p = .001, \eta_p^2 = .46$). Intentional binding effects thus did not correspond with explicit control ratings. In particular, factors that increased explicit agency ratings did not produce shorter perceived intervals, as might be predicted, but showed the reverse pattern. To identify potential confounding factors, we analysed action selection decision-times.

2.2.3. *Decision-times*

Decision-times were defined as the interval between trial onset and button press. After excluding responses faster than 100 ms and longer than 2 *SD* above mean, decision-times were analysed with a 2 (action selection) \times 2 (counterfactual outcome) repeated measures ANOVA (**Figure 2**). Both main effects of action selection ($F(1,18) = 17.44, p = .001, \eta_p^2 = .49$) and counterfactual outcome ($F(1,18) = 15.73, p = .001, \eta_p^2 = .47$) were qualified by a significant interaction ($F(1,18) = 24.30, p < .001, \eta_p^2 = .58$), suggesting longer decision-times in trials with opposite compared to identical counterfactual outcomes, but only if actions were selected freely ($t(18) = 4.46, p < .001$) rather than being instructed ($t(18) < 1$).

2.3. Discussion Experiment 1

Experiment 1 showed that available between-action and between-outcome comparisons increased sense of agency. Explicit sense of agency was higher in the presence of available action alternatives. In contrast, the availability of outcome alternatives only increased SoA when the action that would have led to the counterfactual outcome had previously been available (i.e., in free choice but not instructed choice conditions). These findings provide first evidence that the availability of counterfactual comparisons can increase the sense of personal control and agency. However, Experiment 1 does not distinguish between two possible mechanisms of agency, a prospective or a retrospective one (Moore & Haggard, 2008). As the availability of action and outcome alternatives was evident at the start of the trial and prior to action selection, participants could have evaluated the different actions and outcomes in advance. This possibly allowed prospective estimates of control based on the extent to which action-selection could influence events. Decision-time data suggests that the combination of free choice and a differential outcome indeed engaged a specific process of

outcome-based action selection. Alternatively, agency judgments could also have been evaluated after action-selection, based on comparing the actual outcome with the outcome that could have been obtained. On a strict interpretation, only the latter, retrospective, evaluation would constitute a truly counterfactual effect (Byrne & Egan, 2004). We set out to distinguish between these two components in Experiment 2.

3. Experiment 2

Experiment 2 involved a modified version of the task, in which presentation of outcome alternatives was delayed until after action-selection. If the increase in SoA observed in Experiment 1 reflected a true counterfactual, as opposed to prospective evaluation of alternative outcomes known in advance, then Experiment 2 should reproduce the agency pattern observed in Experiment 1.

3.1. Material and Methods

3.1.1. Participants

20 participants (10 females, average age 23 years) participated in Experiment 2. Just as in Experiment 1, participants provided informed consent and received £10.50 for participation.

3.1.2. Procedure

The procedure of Experiment 2 was identical to Experiment 1, except that counterfactual outcome information was only available at the point of outcome presentation (see **Figure 1**). All trials started with a selection between two differently shaded mouse holes. After action selection, identical or opposite outcome colours revealed whether chosen and un-chosen actions had identical or opposite outcomes. This manipulation ensured that counterfactual outcome information informed sense of agency in purely retrospective way.

3.2. Results

3.2.1. Explicit control ratings

Similar to Experiment 1, in Experiment 2 agency ratings were higher in free compared to instructed trials ($F(1,19) = 32.57, p < .001, \eta_p^2 = .63$) (see **Figure 2**). Agency was higher in trials with opposite compared to identical counterfactual outcomes ($F(1,19) = 12.46, p = .002, \eta_p^2 = .40$). Again, opposite counterfactual outcomes only increased control ratings in free ($t(19) = 3.93, p = .001$) but not instructed trials ($t(19) < 1$), as indicated by an action selection \times outcome alternative interaction ($F(1,19) = 13.44, p = .002, \eta_p^2 = .41$). Agency ratings were higher following positive compared to negative outcomes ($F(1,19) = 24.55, p < .001, \eta_p^2 = .56$).

To compare the two experiments directly, a Bayesian repeated measures ANOVA (JSASP, JASP Team, 2016) was performed. It showed moderate to strong support for the null hypothesis regarding all interactions including the between-subject factor experiment (all $BF_{01} > 6$), suggesting that control rating patterns did not differ between Experiments 1 and 2.

3.2.2. Interval estimation

Mean normalised interval duration estimations between action and outcome were entered into a $2 \times 2 \times 2$ repeated measures ANOVA (**Figure 2**). An outcome valence \times counterfactual

outcome interaction ($F(1,19) = 5.30, p = .033, \eta_p^2 = .22$) indicated different directions of the outcome valence effect as a function of counterfactual outcome. However, pairwise comparison between positive and negative outcomes did neither reach significance in identical ($t(19) = 1.10, p = .28$) nor in opposite counterfactual outcomes ($t(19) = -1.20, p = .24$).

3.2.3. Decision-times

A main effect of action selection ($F(1,19) = 24.21, p < .001, \eta_p^2 = .56$) indicated longer decision-times in free choice trials (see **Figure 2**). Similar to Experiment 1, in Experiment 2 decision-times thus showed a pattern that was similar to the pattern observed for interval estimations: choices which required longer decisions were followed by longer interval estimations. Heightened attention and arousal in trials with high anticipated agency might explain this relationship, as these factors can increase perceived temporal durations (Eagleman, 2008). Intentional binding, the implicit measure of SoA, can thus be assumed to be confounded with decision-times, which possibly occludes further effects. We therefore refrain from interpreting the intentional binding results of both experiments and to focus the discussion on explicit agency rating findings.

3.3. Discussion Experiment 2

Experiment 2 replicated the pattern of agency ratings that were observed in Experiment 1. This suggests that a purely retrospective, genuinely counterfactual, evaluation of between-outcome alternatives is sufficient to increase subjective sense of agency ratings.

4. General Discussion

In the present paper we investigated how the sense of agency is influenced by the presence or absence of counterfactual comparisons between alternative actions, and between alternative outcomes that these actions produce. These comparisons constitute two essential yet distinct aspects of the *could have done otherwise* requirement of personal control and responsibility. Our results reveal two novel findings.

First, we found that explicit sense of control over the outcome of one's action was highest following trials which allowed both between-action and between-outcome comparisons. This is consistent with the role of explicit counterfactual thinking in the evaluation of outcome preventability (Mandel & Lehman, 1996). Preventable or inhibitory causes are the personally controllable subset of all causes that contribute to an event, and are therefore informative cues to personal causation (Lagnado, Gerstenberg & Zultan, 2013). Accordingly, sense of agency was highest when outcomes were perceived as preventable, that is when the agent could have performed another action which in turn would have resulted in a different outcome. This finding allows to reformulate the CDO requirement to the perception of personal causation and agency in a more precise way: "*I could have performed another action, thereby obtaining another outcome*". Interestingly, counterfactual outcome alternatives only informed agency if they were consequences of actually viable (i.e., potentially selectable) action alternatives, whereas counterfactual outcomes of unavailable actions did not affect SoA. Counterfactual comparisons thus focused on realistic alternatives constrained by task structure, following the nearest possible world constraint (Lewis, 1973; Woodward, 2011). In this sense, explicit judgements of agency are well-aligned with the concept of causal control.

The second main finding clarifies the time-course of counterfactual influences on sense of agency. We show that counterfactual outcome information influences sense of agency whether it is presented early in the action-outcome sequence, at the point of action selection (Experiment 1), or only later when action is complete and the outcome is presented (Experiment 2). This suggests that counterfactual alternatives need not necessarily be considered at the point of action selection to influence sense of agency – consideration at the point of outcome presentation is sufficient. Counterfactual outcome alternatives thus affect judgements of agency over actions already made in a purely retrospective way, and without reference to action-selection processes.

At first sight it might seem surprising that retrospectively available outcome information increases sense of agency. This is at odds with findings that thinking counterfactually about situations in which one “could not have known better” can reduce perceived control. Such excuse-making effects have been shown to emphasise that an outcome was inevitable because of lacking relevant information (Markman & Tetlock, 2000). Similar reasoning patterns may underlie the tendency to preserve a positive self-evaluation in the face of failure (Festinger, 1962). While some studies show that such short-term affective benefits can be detrimental to future performance by decreasing the motivation to prevent future negative outcomes (McCrea, 2008), other findings suggest that counterfactuals that increase the feeling of control can in fact increase subsequent motivation and performance (Nasco & Marsh, 1999).

On standard views, explicit counterfactual thoughts are primarily activated by negative outcomes such as failed goals and serve to evaluate how these outcomes could have turned out better (Epstude & Roese, 2008). These counterfactuals are termed upward counterfactuals and are assumed to regulate behaviour via two mechanisms: by transforming previous experience into behavioural plans for the future (i.e., learning) and by increasing “affect-fuelled motivation” (Epstude & Roese, 2011, p. 21) as outlined above. However, neither mechanism seems to fully explain the present findings. Firstly, our task precluded the possibility of using counterfactual information as a learning signal because of random action-outcome mapping. Available counterfactual comparisons thus could not have activated behavioural intentions that in turn increased objective performance and control in subsequent trials (Smallman, 2013, Smallman & Roese, 2009). In this sense, our design broke the conventional linkage between counterfactuals and outcome-based learning. Second, while motivational or affective regulation may be adaptive after an error, its effects should reverse following positive outcomes, leading to downward counterfactuals which evaluate less desirable potential courses of events. As such, positive outcomes should have reduced perceived control by making potential failures more salient. However, this pattern was not observed in our studies. In particular, control ratings were not affected by actual outcome valence and the resulting direction of counterfactual comparisons.

This suggests that the implicit CDO type counterfactual comparisons encouraged by our task differed from previously studied instances of motivational-affective regulation by explicit counterfactual thinking. In particular, the active nature of the employed decision-making task and the personally relevant and tangible outcomes might have led to a differential type of processing compared to hypothetical vignettes previously used in the investigation of explicit counterfactual thinking and control. We hypothesise that the present implicit counterfactual comparisons were related to the online tracking of agency and causal control during active decision-making. Nevertheless, the retrospective nature of these

evaluations shown in Experiment 2 is compatible with a view of explicit counterfactual evaluations as a part of the post-actional phase (Epstude & Roese 2008).

Taken together, our findings demonstrate that the availability of counterfactual comparisons in an active choice task directly influences SoA, even if this counterfactual information promotes neither learning nor affective-motivational regulation. We found that CDO type evaluations increase the perceived control over action outcomes when two independent conditions are met. First, the agent could have chosen an alternative action. Second, this action would have changed the outcome. This finding suggests a strong, possibly automatic mind-set for agency and personal causation, even in the absence of objective control over action outcomes.

In addition to these key findings, explicit agency ratings in both experiments confirmed two well-documented evaluative biases, namely the self-serving bias and the illusion of control. Regarding self-serving bias, we found higher agency ratings for positive compared to negative outcomes (Bandura, 1984). This effect persisted even in instructed trials, suggesting that it was only influenced by actually obtained outcomes, while neither the actions leading to these outcomes nor counterfactual outcomes were considered. Self-serving bias was thus decoupled from actual control and CDO evaluations. The second bias – illusion of control – refers to stronger perceived control over the outcome in the presence of free choices, even when outcomes are solely determined by chance (Langer, 1975). In the present study it manifested in higher agency ratings in free compared to instructed choices between identical alternatives, as here choice could not affect outcome. The illusion of control has previously been attributed to a general confusion of skill and chance (Stefan & David, 2013). However, it also involves a counterfactual between-action comparison (“*I could have chosen another action*”). Free choice situations therefore partly satisfy the CDO requirement, which possibly increases perceived agency. The illusion of control might thus be a by-product of CDO-type evaluation of action alternatives which helps to monitor agents’ SoA during active decision-making. As such, available between-action CDO evaluations seem to influence sense of agency, even when there is clear evidence against between-outcome CDO.

5. Conclusion

The present study tested the direct influence of available counterfactual comparisons on the sense of agency. By independently manipulating information about available counterfactual actions and counterfactual outcomes, we identified the specific contribution of each. We thereby clarified the *could have done otherwise* requirement for judgements of agency. Our key result is that SoA was strongest when participants knew both that they could choose between alternative actions, and that this choice would produce alternative outcomes. Importantly, this combined effect of both alternative actions and alternative outcomes was present even when outcome alternatives were presented after actions were already made. This finding strongly suggests that CDO effects on SoA are retrospective reconstructions. In this sense, the CDO aspect of SoA seems to relate to general beliefs about one’s own causal capacity rather than to the online control of the current action. These new findings allow a more precise phrasing of the CDO requirement, which we express by the counterfactual statement that the agent *could have performed another action, thereby obtaining another outcome*.

Acknowledgements

This work was supported by a European Research Council Advanced Grant (HUMVOL 323943) and by an Arts and Humanities Research Council grant (AH/L015145/1) to PH.

References

- Bandura, A. (1984). Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research*, 8(3), 231-255.
- Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge: MIT Press.
- Byrne, R. M. (2016). Counterfactual Thought: From Conditional Reasoning to Moral Judgment. *Annual Review of Psychology*, 67(1).
- Byrne, R. M., & Egan, S. M. (2004). Counterfactual and prefactual conditionals. *Canadian Journal of Experimental Psychology*, 58(2), 113.
- Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674), 1167-1170.
- Caspar, E. A., Cleeremans, A., & Haggard, P. (2015). The relationship between human agency and embodiment. *Consciousness and Cognition*, 33, 226-236.
- Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2013). An online neural substrate for sense of agency. *Cerebral Cortex*, 23(5), 1031-1037.
- Dannenberg, L., Förster, J., & Jostmann, N. B. (2012). "If only...": When counterfactual thoughts can reduce illusions of personal authorship. *Consciousness and Cognition*, 21(1), 456-463.
- Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition*, 20(4), 1211-1220.
- de Biran, M. (1841). *Oeuvres philosophiques*. Paris: Ladrangé.
- Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, 18(2), 131-136.
- Engbert, K., Wohlschläger, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition*, 107(2), 693-704.
- Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, 12(2), 168-192.
- Epstude, K., & Roese, N. J. (2011). When goal pursuit fails. *Social Psychology*, 42(1), 19-27.
- Festinger, L. (1962). *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.
- Frankfurt, H. G. (1988). *The importance of what we care about: Philosophical essays*. Cambridge: Cambridge University Press.
- Frith, C. D. (2014). Action, agency and responsibility. *Neuropsychologia*, 55, 137-142.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5(4), 382-385.
- Haggard, P., & Tsakiris, M. (2009). The experience of agency feelings, judgments, and responsibility. *Current Directions in Psychological Science*, 18(4), 242-246.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93(2), 136.
- Lagnado, D. A., Gerstenberg, T., & Zultan, R. I. (2013). Causal Responsibility and Counterfactuals. *Cognitive Science*, 37(6), 1036-1073.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.

- Lewis, D. (1986). Causal explanation. *Philosophical Papers*, 2, 214-240.
- Mackie, J. (1974). *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.
- Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, 71(3), 450.
- Markman, K. D., & Tetlock, P. E. (2000). 'I couldn't have known': accountability, foreseeability and counterfactual denials of responsibility. *British Journal of Social Psychology*, 39(3), 313-325.
- McCrea, S. M. (2008). Self-handicapping, excuse making, and counterfactual thinking: consequences for self-esteem and future motivation. *Journal of Personality and Social Psychology*, 95(2), 274.
- Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition*, 17(1), 136-144.
- Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Consciousness and Cognition*, 21(1), 546-561.
- Nasco, S. A., & Marsh, K. L. (1999). Gaining control through counterfactual thinking. *Personality and Social Psychology Bulletin*, 25(5), 557-569.
- Pereboom, D. (2011). Free-will skepticism and meaning in life. In R. Kane (Ed.), *The Oxford Handbook of Free Will*. Oxford, UK: Oxford University Press.
- Smallman, R. (2013). It's what's inside that counts: The role of counterfactual content in intention formation. *Journal of Experimental Social Psychology*, 49(5), 842-851.
- Smallman, R., & Roese, N. J. (2009). Counterfactual thinking facilitates behavioral intentions. *Journal of Experimental Social Psychology*, 45(4), 845-852.
- Stefan, S., & David, D. (2013). Recent developments in the experimental investigation of the illusion of control. A meta-analytic review. *Journal of Applied Social Psychology*, 43(2), 377-386.
- Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1), 219-239.
- Woodward, J. (2011). Psychological studies of causal and counterfactual reasoning. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation. Issues in philosophy and psychology* (pp. 16-53). Oxford, UK: Oxford University Press.
- Zeelenberg, M., van Dijk, W. W., Van der Pligt, J., Manstead, A. S., Van Empelen, P., & Reinderman, D. (1998). Emotional reactions to the outcomes of decisions: The role of counterfactual thought in the experience of regret and disappointment. *Organizational Behavior and Human Decision Processes*, 75(2), 117-141.