Cognition 206 (2021) 104489

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

To prevent means to know: Explicit but no implicit agency for prevention behavior

Roland Pfister*[,1], Solveig Tonn*[,1], Lisa Weller, Wilfried Kunde, Katharina A. Schwarz

Julius Maximilians University of Würzburg, Würzburg, Germany

ARTICLE INFO

Keywords:
Prevention
Active avoidance
Agency
Temporal binding
Action representation

ABSTRACT

Human agents draw on a variety of explicit and implicit cues to construct a sense of agency for their actions and the effects of these actions on the outside world. Associative mechanisms binding actions to their immediate effects support the evolution of agency for operant actions. However, human agents often also act to prevent a certain event from occurring. Such prevention behavior poses a critical challenge for the sense of agency, as successful prevention inherently revolves around the absence of a perceivable effect. By assessing the psychological microstructure of singular operant and prevention actions we show that this comes with profound consequences: agency for prevention actions is only evident in explicit measures but not in corresponding implicit proxies. These findings attest to an altered action representation in prevention behavior and they support recent proposals to model related processes such as avoidance learning in terms of propositional rather than associative terms.

1. Introduction

Avoidance of threatening situations is a core component of adaptive behavior. It is readily learned in a variety of species, and the study of its underlying learning mechanisms has seen a renaissance in the last decade (Krypotos, Effting, Kindt, & Beckers, 2015; LeDoux, Moscarello, Sears, & Campese, 2017). For research on humans, this renaissance was driven by novel theoretical approaches highlighting a major role of higher-order cognition for avoidance: Whereas classical accounts had explained avoidance by a sequence of classical and instrumental conditioning (Mowrer, 1960), these current theories stress the role of expectancies and mental simulations for acquiring and especially for maintaining avoidance behavior (De Houwer, Beckers, & Vandorpe, 2005; Lovibond, 2006).

Even though expectancy theories receive strong support from empirical findings (e.g., Lovibond, Mitchell, Minard, Brady, & Menzies, 2009), a pervasive issue is how avoidance is actually represented in the cognitive system. Contemporary expectancy theories typically assume propositional action representations for prevention behavior, i.e., conscious representations of how actions relate to the disappearance or prevention of threatening and harmful events (De Houwer, 2009). Evidence for this claim is currently mixed, however (Declercq & De Houwer, 2009; Sevenster, Beckers, & Kindt, 2014). In light of this database, we propose to extend the methodological repertoire of studies on avoidance: In addition to studying the acquisition of avoidance behavior through learning and test paradigms – the de facto gold standard of avoidance research (Solomon & Wynne, 1954) – we propose to adopt measures from research on perception in action to target the psychological microstructure of active avoidance, i.e., single behavioral instances that aim at preventing a stimulus from occurring (Eder & Dignath, 2014; McGraw, Larsen, Kahneman, & Schkade, 2010).

A particularly relevant type of measures relates to the agent's perception of being causally involved in a particular event; that is, the feeling of controlling upcoming events through one's actions. On an explicit, conscious level, such a sense of agency is often probed by explicit judgments of agency (Haggard & Tsakiris, 2009). Perceived causality further sparks certain perceptual illusions that have been discussed as implicit measures for the sense of agency, and this is especially the case for the phenomenon of temporal binding (Haggard, Clark, & Kalogeras, 2002; Kirsch, Kunde, & Herbort, 2019). In operant actions, i.e., actions that aim at triggering a certain sensory effect, temporal binding describes the phenomenon that the perceived points in time of action and effect are drawn together so that the action appears to unfold slightly later than it actually occurs (action binding) whereas the action-contingent effect appears to occur slightly earlier than it actually does (effect binding; Moore & Obhi, 2012). Such implicit measures carry unique information about how an action is represented as they have often been found not to correlate with explicit

* Corresponding authors at: University of Würzburg, Department of Psychology III, Röntgenring 11, 97070 Würzburg, Germany.
E-mail addresses: roland.pfister@psychologie.uni-wuerzburg.de (R. Pfister), solveig.tonn@uni-wuerzburg.de (S. Tonn).

[1] Equal author contribution.

https://doi.org/10.1016/j.cognition.2020.104489
Received 17 June 2020; Received in revised form 8 October 2020; Accepted 9 October 2020

judgments of agency, neither intra-individually across different trials nor inter-individually across different agents (e.g., Dewey & Knoblich, 2014; Saito, Takahata, Murai, & Takahashi, 2015; Schwarz, Weller, Klaffehn, & Pfister, 2019). This has led several authors to propose a distinction between explicit and implicit ("pre-reflective") components of the sense of agency though this theoretical proposal is still under discussion (David, Newen, & Vogeley, 2008; Moore & Obhi, 2012). In any case, a comprehensive account of a certain type of action thus requires both implicit and explicit measures, so that we aimed at assessing explicit judgments of agency as well as temporal binding (action binding).

But what exactly is measured via perceptual illusions such as temporal binding? Evidence suggests that binding also occurs for causal event chains that do not involve an action of the observing agent, so that intentionality cannot be seen as a necessary precondition for temporal binding to occur (e.g., Ruess, Thomaschke, & Kiesel, 2020; Suzuki, Lush, Seth, & Roseboom, 2019). However, this does not preclude the possibility that intentionality can be sufficient to trigger binding even in the absence of directly perceivable causality. Action binding, for instance, also occurs in the absence of perceivable effects, as long as participants anticipate their action to yield a certain effect on the environment (Engbert & Wohlschläger, 2007; Moore & Haggard, 2008). The same holds true for the flipside of the coin, i.e., triggering an effect by means of deliberate omissions of an action. Such deliberate non-actions may yield temporal binding for the following effect as compared to a situation that is identical in terms of physical stimulation but differs in terms of the underlying intentions (Weller, Schwarz, Kunde, & Pfister, 2020). Temporal binding thus occurs even for single actions without immediate effects on the environment and for single events that are not preceded by an action, given that participants construe the situation as involving a causal chain of (non-)events. Due to this property, the measure of temporal binding can be applied directly to actions that aim at preventing an upcoming event as well.

In the current experiments we therefore adapted a typical setup for assessing temporal binding (Haggard et al., 2002) to a prevention context in which actions aim at preventing an event that would have occurred otherwise. Theories that assume avoidance to be based on associative learning would predict avoidance actions to yield reliable measures of agency on the explicit and the implicit level alike, just as it occurs for operant actions which aim at causing a certain event. A dissociation of explicit and implicit measures, by contrast, would support propositional rather than associative accounts, as this pattern of results would indicate a mandatory recruitment of conscious evaluations of action and prevented event without an implicit, associative component. We tested this prediction in a high-powered experiment, supported by two supplementary experiments to address potential alternative explanations of the main findings.

2. Method

2.1. Open science statement

All procedures were pre-registered prior to data collection (https://osf.io/xfs5r/). Data and analysis scripts are available at the Open Science Framework (https://osf.io/bac7s/).

2.2. Apparatus and procedure

Participants observed a 24″ monitor at a viewing distance of about 60 cm and operated a standard German QWERTZ keyboard. The monitor showed a clock face (diameter: 2.7 cm) with a rotating clock hand (frequency: one rotation per 2000 ms). Below the clock face was a loading bar (3.5 cm × 1.5 cm) that served as a visual trigger to the participants. Participants went through 3 different conditions, presented in separate blocks of a full within-subjects design. In each condition, participants could decide between pressing the response key (action) or not (omission). In the *baseline condition* neither actions nor omissions triggered any additional effects. In the *operant condition* each keypress caused a neutral tone, whereas there was no tone if the participant chose not to press the key. Crucially, in the *prevention condition* white noise was played in case of omissions and participants could prevent this stimulus by pressing upon filling of the bar. That is, not pressing the response key would lead to white noise, whereas pressing the response key would result in no tone effect. Participants were explicitly informed about the condition of the upcoming block and the corresponding consequences that their choices would entail. Each condition was presented in two blocks of 42 trials, with block order counterbalanced across participants (either O-B-P-O-B-P or P-B-O-P-B-O, with O = operant, B = baseline, P = prevention).

Each trial commenced by the loading bar filling up continuously. Across trials, the time it took for the bar was either 2000 ms or 2500 ms; we included this manipulation to ensure that participants needed to attend the bar in each trial while at the same time being able to assess the point in time in which a (non)action would be registered (see Weller et al., 2020, for a similar procedure; additional methodological considerations are discussed in the Supplementary Material). Participants could decide freely whether they would act or not upon completion of the bar. Possible sound effects were then played 300 ms later via headphones depending on the current condition and the participant's choice. This mechanic ensured that action-(non-)effect relations were constant across conditions to avoid confounds due to between-condition variance, since prior studies showed an influence of varying action-effect delays on temporal binding (e.g., Ruess, Tomaschke, & Kiesel, 2017; Wen, Yamashita, & Asama, 2015; see Supplementary Methods for a complete description). For each trial with a keypress action, participants were asked to estimate the perceived point in time of their action to assess temporal binding (action binding). They indicated their estimates by reading off the minutes on the clock face when pressing the key and entering these numbers when prompted at the end of a trial. In 1/3 of the trials we further probed for explicit agency by asking "How much did you feel that you caused the tone?" if there had been a tone in the preceding trial or "How much did you feel responsible that there was no tone?" if there was no tone. Participants gave these ratings by adjusting a slider of a visual analogue scale (VAS) with the computer mouse. To also assess this explicit agency in the baseline condition, where keypresses did not trigger any tones, we played unexpected tones at the beginning of selected trials of this condition (also in 1/3 of the trials).

Participants were assigned to one of two groups. One group performed the task as described above whereas the other group received an additional safety signal in the prevention condition. Here, the clock hand immediately turned blue after the program had registered a keypress, thus informing the participants that no white noise would be played for the remainder of the trial. Such safety signals have been described to have a lasting impact on avoidance behavior (Bolles & Grossen, 1969) so that we intended to explore whether they would also enhance temporal binding for successful prevention actions.

2.3. Participants

One hundred participants (mean age: 24.00 years; 83 females) were recruited for the study, 50 per group. Effect-size estimates for action binding vary in the literature so that we based power calculations on the condition that most closely resembles the current setup, i.e., the "action only" condition of Moore and Haggard (2008), in which actions aimed at triggering an effect that turned out not to occur (Cohen's $d_z = t / \sqrt{N} = 4.81 / \sqrt{10} = 1.52$). Our experiment therefore ensured a high power of $1\text{-}\beta > .99$ for detecting temporal binding in the overall sample as well as in each subgroup, and this also holds true when assuming a conservative estimate of $d_z = 0.8$. Sample sizes further ensured sufficient power for gathering evidence for absence of an effect via Bayesian analyses as well as a power of $1\text{-}\beta = .80$ for between-group

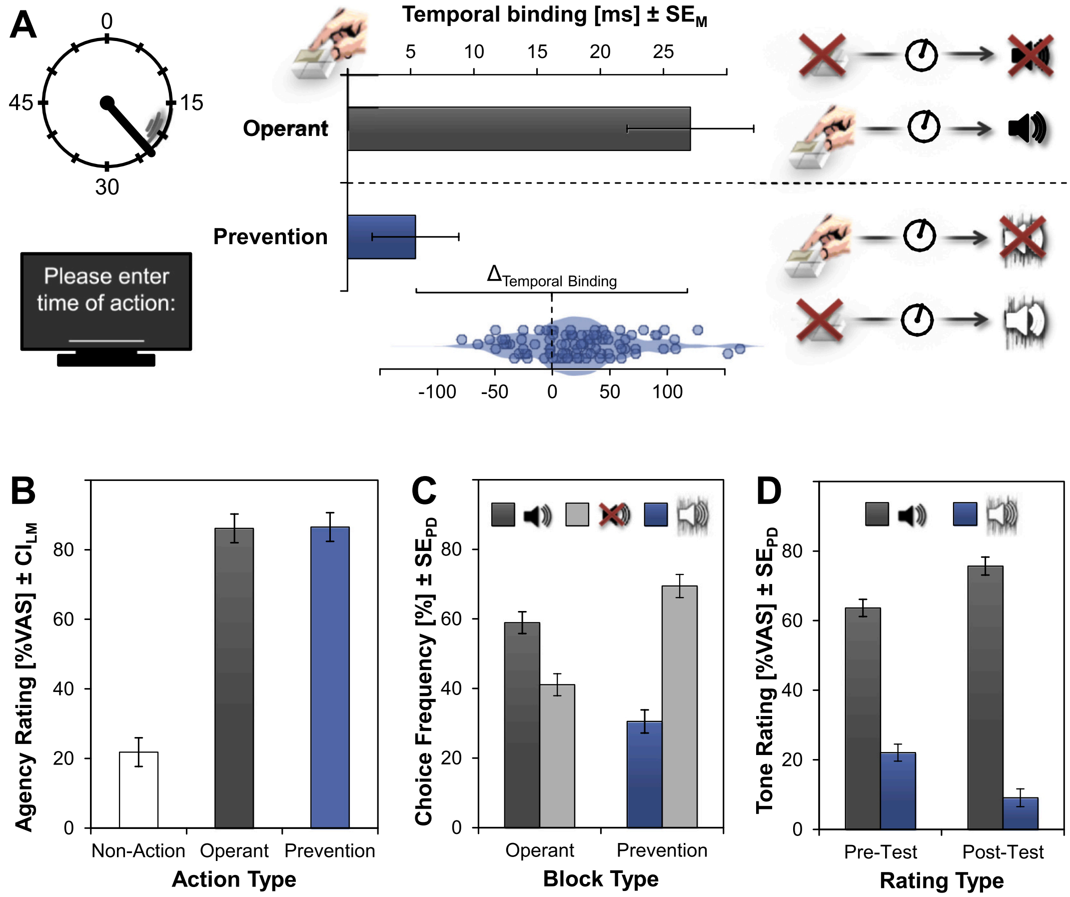

Fig. 1. Results of the main experiment. **(A)** Temporal binding for operant actions and prevention behavior. Binding scores were computed by comparing interval estimates in the prevention and the operant condition to the baseline condition (SE_M = standard error of the mean). Bayesian analyses indicated evidence for the absence of an effect in the case of prevention behavior (BF_{01} = 3.67). The scatter plot shows difference scores between both conditions for each participant and an estimate of the corresponding density function. **(B)** Mean agency ratings for baseline, operant and prevention trials (CI_{LM} = 95% confidence intervals according to the method of Loftus & Masson, 1994). VAS = visual analogue scale. **(C)** Choice frequencies (in %) for operant and prevention blocks. Operant actions triggered a neutral tone (dark grey bar) whereas action omissions in prevention blocks were followed by white noise (blue bar). No sound events occurred in the remaining conditions (light grey bars). Error bars indicate standard errors of paired differences (SE_{PD}; Pfister & Janczyk, 2013), calculated separately for each block type. **(D)** Tone ratings on a VAS (0 = negative, 100 = positive) before and after the experiment (pre-test vs. post-test). Dark grey bars show mean ratings for the neutral tone whereas blue bars show mean ratings for white noise. Error bars were calculated separately for pre- and post-test. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

effects with $d_s > 0.55$. For Bayesian *t*-tests, we chose to use a Cauchy scale parameter of 1 as a conservative prior for the expected effect size.

All participants gave written informed consent and they received payment or course credit for their participation. As per our pre-registration, seven participants had to be excluded because their data came with less than 10 usable trials for the analysis of temporal binding in at least one condition.

3. Results

3.1. Temporal binding and explicit ratings

Action binding was determined for each participant by subtracting the mean estimation error of the baseline condition from the mean estimation error of the operant and the prevention condition, respectively, so that positive values indicate that actions in the latter two conditions were perceived to have occurred later than in the baseline condition (see the Supplementary Material for details on data pre-processing). The resulting binding scores were then tested against 0 with a two-tailed, one-sample *t*-test (see Fig. 1A).

The operant condition showed a reliable action binding effect (27 ms), $t(92) = 5.42$, $p < .001$, $d_z = 0.56$, whereas the prevention condition did not, $t(92) = 1.57$, $p = .120$, $d_z = 0.16$. Follow-up analyses of Bayes Factors supported these findings by indicating strong support for the presence of an effect in the operant condition, $BF_{10} = 24596.19$, and evidence for the absence of an effect in the prevention condition, $BF_{10} = 0.27$ ($BF_{01} = 3.67$; computed as JZS Bayes Factors with a non-directional test). A direct comparison of both conditions revealed a significant difference in action binding (Δ = 22 ms), $t(92) = 4.60$, $p < .001$, $d_z = 0.48$, suggesting a marked difference in temporal binding between operant and prevention behavior, $BF_{10} = 1011.71$.

Explicit agency ratings, by contrast, were equally high in the operant condition and the prevention condition (Fig. 1B), $t(92) = 0.22$, $p = .826$, $d_z = 0.02$, $BF_{10} = 0.08$ ($BF_{01} = 11.92$). Both conditions differed markedly from the baseline, $ps < .001$, $d_z > 2.25$, $BFs_{10} > 5.90 \times 10^{34}$, as also reflected in an omnibus analysis of variance (ANOVA) across all three conditions, $F(2, 184) = 416.16$, $p < .001$, $\eta_p^2 = .82$ (corrected according to Greenhouse and Geisser's method; $\varepsilon = .772$), suggesting high subjective agency in operant and prevention behavior alike.

3.2. Manipulation check

To assess whether the white noise stimulus was sufficiently aversive to trigger prevention behavior we analyzed the distribution of action versus non-action choices in the prevention condition as compared to the operant condition (see Fig. 1C). We thus computed one-sample *t*-tests to compare the individual values to the choice frequency that would be expected by chance (50%). In the operant condition, participants chose to act in 59%, suggesting a choice bias towards acting (and/or towards producing a neutral tone effect), $t(92) = 5.68$, $p < .001$, $d_z = 0.59$. This bias towards acting was also evident in the prevention condition (69%), $t(92) = 11.74$, $p < .001$, $d_z = 1.22$, and a comparison of both conditions indicated that participants showed a stronger bias towards preventing the white noise than towards producing the neutral tone, $t(92) = 6.40$, $p < .001$, $d_z = 0.66$. This pattern was mirrored in subjective ratings of the tone valence which participants gave at the beginning (pre-test) and at the end of the experiment (post-test; Fig. 1D). A 2 (tone: neutral vs. white noise) x 2 (position: pre-test vs. post-test) repeated-measures ANOVA indicated that the white noise was rated as negative at both time points, $F(1, 92) = 841.40$, $p < .001$, $\eta_p^2 = .90$, while this difference increased over the experiment, $F(1, 92) = 64.58$, $p < .001$, $\eta_p^2 = .41$ (main effect of position: $F(1, 92) = 7.44$, $p = .008$, $\eta_p^2 = .07$).

3.3. Safety signals

In a follow-up analysis we compared the group without safety signals and the group with safety signals to determine whether the presence of safety signals would foster implicit measures of agency (Fig. 2). Whereas action binding was present in the operant condition of both groups, $ps < .001$, $d_z > 0.48$, $BFs_{10} > 10.52$, action binding was consistently absent in the prevention condition of both groups, $ps > .255$, $d_z < 0.18$, $BFs_{10} < 0.23$ ($BFs_{01} > 4.52$). Accordingly, a 2 (group: safety signal vs. no safety signal) x 2 (condition: operant vs. prevention) ANOVA did not show an interaction, $F(1, 91) = 0.17$, $p = .682$, $\eta_p^2 = .00$, and a Bayesian ANOVA indicated evidence in favor of the null hypothesis of no between-group difference, $BF_{01} = 4.13$, and decisive evidence for an overall reduction of temporal binding for prevention behavior as compared to operant actions, $BF_{10} = 1389.09$.

3.4. Supplementary experiments

One potential limitation of the main experiment is that the use of strict timing might have drawn the participants' attention away from the fact that their actions prevented a certain stimulus from occurring. Such attentional limitations would especially affect the prevention condition as it takes additional cognitive effort to register the absence of an event as compared to the occurrence of an effect (Horváth, Müller, Weise, & Schröger, 2010). To replicate the results in the absence of timing constraints, Exp. S1 employed the same design as the main experiment but participants did not have to time their action using a loading bar ($n = 48–2$; see the Supplementary Material for details). The results replicated the findings of the main experiment with robust action binding in the operant condition (12 ms), $t(45) = 3.85$, $p < .001$, $d_z = 0.57$, $BF_{10} = 62.09$, but no action binding in the prevention condition (4 ms), $t(45) = 1.33$, $p = .192$, $d_z = 0.20$, $BF_{10} = 0.27$ ($BF_{01} = 3.72$), and a significant difference between both conditions, $t(45) = 2.45$, $p = .018$, $d_z = 0.36$, $BF_{10} = 1.82$ (see Fig. S1 in the Supplementary Material).

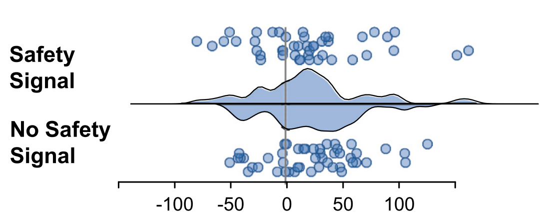

Fig. 2. Individual temporal binding scores for participants in the group with safety signal (upper half) and the group without safety signal (lower half). Positive values indicate action binding and the shaded area indicates the estimated density function for the binding scores of each group.

A second potential objection holds that the negative valence of to-be-prevented events is responsible for the absence of action binding in the prevention condition rather than the specifics of prevention behavior itself. Due to the mixed evidence on the role of effect valence for temporal binding (Moreton, Callan, & Hughes, 2017; Takahata et al., 2012; Yoshie & Haggard, 2017), we therefore conducted Exp. S2 in which neutral tones of different pitch could be produced in the operant condition and prevented in the prevention condition (300 vs. 600 Hz; duration: 100 ms; no white noise stimuli were used, see the Supplementary Material for details). This study ($n = 32–4$) again yielded robust action binding in the operant condition (15 ms), $t(27) = 2.94$, $p = .007$, $d_z = 0.56$, $BF_{10} = 5.62$, but not in the prevention condition (1 ms), $t(27) = 0.26$, $p = .801$, $d_z = 0.05$, $BF_{10} = 0.14$ ($BF_{01} = 6.98$), and a significant difference between both conditions, $t(27) = 2.74$, $p = .011$, $d_z = 0.52$, $BF_{10} = 3.62$.

Finally, a pooled analysis of all three experiments ($N = 167$) for a power of $1\text{-}\beta = 0.97$ even for small effects of $d_z = 0.3$ corroborated this assessment by again yielding evidence for the absence of temporal binding for prevention actions, $BF_{10} = 0.21$ ($BF_{01} = 4.77$).

4. Discussion

Our participants reported high explicit ratings of agency for prevention actions, whereas effects on the implicit measure of temporal binding were consistently absent for this type of behavior. These results attest to an altered action representation for prevention as compared to operant behavior, supporting views that suggest a fully propositional representation in this case (Lovibond et al., 2009; Mitchell, De Houwer, & Lovibond, 2009). This sophisticated representation likely involves situation models, which detail expectations about causal (non)action-(non)event sequences and require explicit, conscious recollection (Ranganath & Ritchey, 2012; Schneider, Albert, & Ritter, 2020).

From a methodological point of view, the consistent absence of temporal binding for prevention behavior is especially notable because our experimental design ensured optimal conditions for this effect to occur. Each operant action triggered a perfectly predictable effect with 100% contingency and, likewise, each prevention action yielded the anticipated omission of the event with 100% contingency. Furthermore, the potential onset of the to-be-prevented effect was constant and easily accessible. In more externally valid scenarios, such perfect action-effect contingencies and clear temporal structures will not be present and may even change with the agent's environment (Behrens, Woolrich, Walton, & Rushworth, 2007). These considerations suggest that the picture obtained in the present study likely generalizes to most settings that involve prevention behavior, and a lack of agency on an implicit level will thus also accompany more extensive processes such as avoidance learning.

Observing evidence for the absence of temporal binding for prevention behavior can also be reconciled with recent cue integration models of the binding phenomenon (Kawabe, Roseboom, & Nishida, 2013; Kirsch et al., 2019; Legaspi & Toyoizumi, 2019; Wolpe, Haggard, Siebner, & Rowe, 2013). These models suggest that temporal binding reflects the weighted integration of different multisensory events. Each of these events comes with a specific sensory or representational precision so that, e.g., a precisely accessible representation of an action exerts a strong pull on a following, less reliable effect and vice versa (Lush et al., 2019). Arguably, the events involved in this process of multisensory fusion need not reflect physical stimulation but any internal event such as the intentional decision not to act may instigate binding just as well (Weller et al., 2020). Cue integration models would thus also predict temporal binding to occur if participants had formed a sufficiently precise representation of the to-be-prevented event during prevention behavior. Absent binding thus indicates a lack of multimodal fusion when viewed from the perspective of cue integration models, suggesting that upcoming effects are only represented during operant actions but not during prevention behavior.

The results further support larger-scale dissociations on a motivational level such as the distinction between prevention and promotion foci in human motivation, i.e., motivation that either aims at producing certain outcomes (promotion) or at preventing undesired events (prevention; Higgins, 1997). Even though a motivational focus on preventing certain outcomes will not necessarily involve active avoidance behavior (Scholer & Higgins, 2008), it seems tempting to ground such motivational processes in systematic differences in action representations as suggested by the present findings. This view also suggests novel hypotheses for common motivational effects such as stronger persistence under a promotion rather than under a prevention focus (Roney, Higgins, & Shah, 1995). These findings are typically explained via differences in goal setting with a promotion focus biasing towards maximal outcomes and a prevention focus biasing towards minimal outcomes (Higgins, 1998). Alternatively, decreased persistence under a prevention focus as compared to a promotion focus might also derive from a consistent lack of implicit cues to agency, as agency has been directly related to the effectiveness of goal pursuit (Higgins, 2015).

The proposed connection of agency and goal pursuit also highlights that the sense of agency for an action should not be seen only as a retrospective assessment. Rather, agency influences goal-setting and action control for upcoming situations as well (Gozli, 2019; Gozli & Dolcini, 2018). If operant actions predictably come with a richer experience of agency than prevention actions, then the human action control system might support the former mode of operation. Alternatively, as temporal binding has been identified to strongly rely on causal inference rather than agency per se (Buehner, 2012; Kirsch et al., 2019; Schwarz, Weller, Pfister, & Kunde, 2019), the present data might indicate that perceived causality is notably stronger when an action leads to a sensory effect compared to when the action prevents a sensory event from occurring (i.e., actually no sensory changes in the environment occur). Even a prevented effect as perfectly predictable and openly advertised as in the present experiments could be seen as more vague in its temporal occurrence and harder to distinguish from the lack of sensory events before and after, unrelated to the participants' actions. The difference in action binding between operant and prevention condition might therefore reflect the participants' difficulty to incorporate prevented effects on an implicit level into a causal chain despite conscious knowledge of the causality involved.

Author contributions

Conceived study and methodology: S.T. with R.P. and L.W.; programmed experiments and collected data: S.T., L.W.; analyzed data: R.P., S.T.; curated data: R.P., S.T.; visualized results: R.P., S.T.; wrote the paper: R.P., S.T., L.W., W. K., K.S.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2020.104489.

References

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*, 1214–1221. https://doi.org/10.1038/nn1954.

Bolles, R. C., & Grossen, N. E. (1969). Effects of an informational stimulus on the acquisition of avoidance behavior in rats. *Journal of Comparative and Physiological Psychology, 68*(1), 90–99. https://doi.org/10.1037/h0027677.

Buehner, M. J. (2012). Understanding the past, predicting the future: Causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*(12), 1490–1497. https://doi.org/10.1177/0956797612444612.

David, N., Newen, A., & Vogeley, K. (2008). The "sense of agency" and its underlying cognitive and neural mechanisms. *Consciousness and Cognition, 17*(2), 523–534. https://doi.org/10.1016/j.concog.2008.03.004.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37*(1), 1–20. https://doi.org/10.3758/LB.37.1.1.

De Houwer, J., Beckers, T., & Vandorpe, S. (2005). Evidence for the role of higher order reasoning processes in cue competition and other learning phenomena. *Learning & Behavior, 33*(2), 239–249. https://doi.org/10.3758/BF03196066.

Declercq, M., & De Houwer, J. (2009). Evidence for a hierarchical structure underlying avoidance behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 35*(1), 123–128. https://doi.org/10.1037/a0012927.

Dewey, J. A., & Knoblich, G. (2014). Do implicit and explicit measures of the sense of agency measure the same thing? *PLoS One, 9*(10), Article e110118. https://doi.org/10.1371/journal.pone.0110118.

Eder, A. B., & Dignath, D. (2014). I like to get nothing: Implicit and explicit evaluation of avoided negative outcomes. *Journal of Experimental Psychology: Animal Behavior Processes, 40*, 55–62. https://doi.org/10.1037/xan0000005.

Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and Cognition, 16*(2), 255–264. https://doi.org/10.1016/j.concog.2006.09.010.

Gozli, D. (2019). *Experimental psychology and human agency.* Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-20422-8.

Gozli, D., & Dolcini, N. (2018). Reaching into the unknown: Actions, goal, hierarchies, and explorative agency. *Frontiers in Psychology, 9*, 266. https://doi.org/10.3389/fpsyg.2018.00266.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382–385. https://doi.org/10.1038/nn827.

Haggard, P., & Tsakiris, M. (2009). The experience of agency: Feelings, judgements, and responsibility. *Current Directions in Psychological Science, 18*(4), 242–246. https://doi.org/10.1111/j.1467-8721.2009.01644.x.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280–1300. https://doi.org/10.1037/0003-066X.52.12.1280.

Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. *Advances in Experimental Social Psychology, 30*, 1–46. https://doi.org/10.1016/S0065-2601(08)60381-0.

Higgins, E. T. (2015). Control and truth working together: The agentic experience of "going in the right direction". In P. Haggard, & B. Eitam (Eds.). *The sense of agency* (pp. 327–344). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190267278.003.0015.

Horváth, J., Müller, D., Weise, A., & Schröger, E. (2010). Omission mismatch negativity builds up late. *NeuroReport, 21*(7), 537–541. https://doi.org/10.1097/WNR.0b013e3283398094.

Kawabe, T., Roseboom, W., & Nishida, S. Y. (2013). The sense of agency is action-effect causality perception based on cross-modal grouping. *Proceedings of the Royal Society B: Biological Sciences, 280*(1763), 20130991. https://doi.org/10.1098/rspb.2013.0991.

Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional binding is unrelated to action intention. *Journal of Experimental Psychology: Human Perception and Performance, 45*(3), 378–385. https://doi.org/10.1037/xhp0000612.

Krypotos, A. M., Effting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience, 9*, 189. https://doi.org/10.3389/fnbeh.2015.00189.

LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: A reconceptualization of a troubled paradigm. *Molecular Psychiatry, 22*, 24–36. https://doi.org/10.1038/mp.2016.166.

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications, 10*, 4250. https://doi.org/10.1038/s41467-019-12170-0.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*(4), 476–490. https://doi.org/10.3758/BF03210951.

Lovibond, P. (2006). Fear and avoidance: An integrated expectancy model. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.). *Fear and learning: From basic processes to clinical implications* (pp. 117–132). American Psychological Association. https://doi.org/10.1037/11474-006.

Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy, 47*(8), 716–720. https://doi.org/10.1016/j.brat.2009.04.013.

Lush, P., Roseboom, W., Cleeremans, A., Scott, R. B., Seth, A. K., & Dienes, Z. (2019). Intentional binding as Bayesian cue combination: Testing predictions with trait individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 45*(9), 1206–1217. https://doi.org/10.1037/xhp0000661.

McGraw, A. P., Larsen, J. T., Kahneman, D., & Schkade, D. (2010). Comparing gains and losses. *Psychological Science, 21*, 1438–1445. https://doi.org/10.1177/0956797610381504.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences, 32*, 183–246. https://doi.org/10.1017/S0140525X09000855.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition, 17*(1), 136–144. https://doi.org/10.1016/j.concog.2006.12.004.

Moore, J. W., & Obhi, S. (2012). Intentional binding and the sense of agency: A review.

Consciousness and Cognition, 21(1), 546–561. https://doi.org/10.1016/j.concog.2011.12.002.

Moreton, J., Callan, M. J., & Hughes, G. (2017). How much does emotional valence of action outcomes affect temporal binding? *Consciousness and Cognition, 49*, 25–34. https://doi.org/10.1016/j.concog.2016.12.008.

Mowrer, O. H. (1960). *Learning theory and behavior.* New York: John Wiley & Sons. https://doi.org/10.1037/10802-000.

Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation and a few simple rules. *Advances in Cognitive Psychology, 9*(2), 74–80. https://doi.org/10.5709/acp-0133-x.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience, 13*, 713–726. https://doi.org/10.1038/nrn3338.

Roney, C. J. R., Higgins, E. T., & Shah, J. (1995). Goals and framing: How outcome focus influences motivation and emotion. *Personality and Social Psychology Bulletin, 21*(11), 1151–1160. https://doi.org/10.1177/01461672952111003.

Ruess, M., Thomaschke, R., & Kiesel, A. (2020). Intentional binding for unintended effects. *Timing and Time Perception*. https://doi.org/10.1163/22134468-bja10005.

Ruess, M., Tomaschke, R., & Kiesel, A. (2017). The time course of intentional binding. *Attention, Perception, & Psychophysics, 79*(4), 1123–1131. https://doi.org/10.3758/s13414-017-1292-y.

Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. *Consciousness and Cognition, 37*, 1–7. https://doi.org/10.1016/j.concog.2015.07.011.

Schneider, W. X., Albert, J., & Ritter, H. (2020). Enabling cognitive behavior of humans, animals, and machines: A situation model framework. *ZiF-Mitteilungen, 2020*(1), 21–34. https://www.uni-bielefeld.de/(de)/ZIF/Publikationen/Mitteilungen/Aufsaetze/2020-1-Schneider_Albert_Ritter.pdf.

Scholer, A. A., & Higgins, E. T. (2008). Distinguishing levels of approach and avoidance: An analysis using regulatory focus theory. In A. J. Elliot (Ed.). *Handbook of approach and avoidance motivation* (pp. 489–503). New York: Psychology Press. https://doi.org/10.4324/9780203888148.

Schwarz, K. A., Weller, L., Klaffehn, A. L., & Pfister, R. (2019). The effects of action choice on temporal binding, agency ratings, and their correlation. *Consciousness and Cognition, 75*, 102807. https://doi.org/10.1016/j.concog.2019.102807.

Schwarz, K. A., Weller, L., Pfister, R., & Kunde, W. (2019). Connecting action control and agency: Does action-effect binding affect temporal binding? *Consciousness and Cognition, 76*, 102833. https://doi.org/10.1016/j.concog.2019.102833.

Sevenster, D., Beckers, T., & Kindt, M. (2014). Fear conditioning of SCR but not the startle reflex requires conscious discrimination of threat and safety. *Frontiers in Behavioral Neuroscience, 8*(32), https://doi.org/10.3389/fnbeh.2014.00032.

Solomon, R. L., & Wynne, L. C. (1954). Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychological Review, 61*(6), 353–385. https://doi.org/10.1037/h0054540.

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional binding without intentional action. *Psychological Science, 30*(6), 842–853. https://doi.org/10.1177/0956797619842191.

Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's not my fault: Postdictive modulation of intentional binding by monetary gains and losses. *PLoS One, 7*(12), Article e53421. https://doi.org/10.1371/journal.pone.0053421.

Weller, L., Schwarz, K., Kunde, W., & Pfister, R. (2020). Something from nothing: Agency for deliberate nonactions. *Cognition, 196*, 104136. https://doi.org/10.1016/j.cognition.2019.104136.

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of action-outcome delay and arousal on sense of agency and the intentional binding effect. *Consciousness and Cognition, 36*, 87–95. https://doi.org/10.1016/j.concog.2015.06.004.

Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research, 229*(3), 467–474. https://doi.org/10.1007/s00221-013-3419-2.

Yoshie, M., & Haggard, P. (2017). Effects of emotional valence on sense of agency require a predictive model. *Scientific Reports, 7*(1), 8733. https://doi.org/10.1038/s41598-017-08803-3.