Impaired perception of temporal contiguity between action and effect is associated with disorders of agency in Schizophrenia

Manuel J. Roth *[,1,2,3], Axel Lindner *[,#,1,4], Klaus Hesse [4], Dirk Wildgruber [4], Hong-Yu Wong [5], Marc J. Buehner [#,6]

[1] Dept. Cognitive Neurology & Division of Neuropsychology, Hertie Institute for Clinical Brain Research, Tübingen, Germany
[2] International Max Planck Research School for Cognitive and Systems Neuroscience, Tübingen, Germany
[3] Dynamic Cognition Group, Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[4] Dept. Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany
[5] Philosophy of Neuroscience, Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany
[6] School of Psychology, Cardiff University, Cardiff, United Kingdom

*Correspondence should be addressed to AL (a.lindner@medizin.uni-tuebingen.de).
#Equal contributions

Abstract

Delusions of control in Schizophrenia are characterized by the striking feeling that one's actions are controlled by external forces. We here tested qualitative predictions inspired by Bayesian causal inference models, which suggest that such misattributions of agency should lead to decreased intentional binding. Intentional binding refers to the phenomenon that subjects perceive a compression of time between their intentional actions and consequent sensory events. We demonstrate that patients with delusions of control perceived less self-agency in our intentional binding task. This effect was accompanied by significant reductions of intentional binding as compared to healthy controls and patients without delusions. Furthermore, the strength of delusions of control tightly correlated with decreases in intentional binding. Our study validated a critical prediction of Bayesian accounts of intentional binding, namely that a pathological reduction of the prior likelihood of a causal relation between one's actions and consequent sensory events – here captured by delusions of control - should lead to lesser intentional binding. Moreover, our study highlights the import of an intact perception of temporal contiguity between actions and their effects for the sense of agency.

Significance statement

Delusions of control describe the experience that external forces are causing one's own thoughts and actions. Being inspired by Bayesian causal inference models, this study shows that explicit misconceptions of causal relationships in patients suffering from such delusions are paralleled by disturbed perceptions of temporal contiguity between actions and their effects. Our empirical study thus highlights the value of Bayesian accounts to explain psychosis and, more specifically, it depicts how such accounts could be used in the future to quantitatively assess and possibly predict delusions of control in Schizophrenia.

Introduction

Feeling in control over one's actions is the basis of our experience as active agents. This experience is referred to as the sense of agency (*SoA*) (1). A core dimension that allows inferring agency is *time.* For instance, any action *A* will always precede its resulting outcome *O* in time (*t*; $t_A < t_O$). If this temporal order is disturbed (such that t_O is perceived earlier than t_A), an event is no longer judged as being self-produced (compare (2)). Similarly, experimentally delaying the onset of O ($t_O >> t_A$) leads to a diminished perception of self-agency (3). Hence, establishing a SoA critically depends on the temporal order of A and O as well as on their temporal contiguity - as is true for any other causal inference (4).

Considering this import of time for establishing a SoA, it may surprise that the temporal perception of one's own actions and their consequences is heavily distorted. The perceived time (*pt*) of an experimentally delayed action-outcome O appears to be earlier than that of an identical delayed event X triggered by an external cause ($pt_O < pt_X$). At the same time, the onset of the self-action A is perceived as happening closer to its outcome O as compared to the onset of the same action in isolation ($pt_{A\_Alone} < pt_{A\_Outcome}$) (5). This relative increase in perceived temporal proximity between self-actions and their consequences is referred to as "intentional binding" (1).

Yet, rather than hindering causal agency attribution by distorting the perceived time of actions and consequent events, intentional binding may instead reflect Bayes-optimal perceptual inference. According to Hume's rule of spatiotemporal contiguity, events that are closer together in time and space are more likely perceived as causally related (4), and this rule seems to hold true for actions and their sensory consequences as well (compare (6)). Now, if a person *P* were asked to temporally estimate the onset of her action A and of a consequent outcome O, her estimates are likely to be more accurate and precise, if she combined these estimates with prior knowledge about the expected spatiotemporal contiguity of A and O. Such integration is at the heart of Bayesian perceptual inference, according to which current sensory information (i) is combined with an estimate of its prior probability (ii) to determine the most likely percept given (i) and (ii); and the result of this integration is often (close to) optimal in natural settings (for review e.g. see (7)). Thus, if O followed A with an

(artificial) delay that is longer than what she expects, integrating this prior belief will lead to intentional binding (8–10). Similarly, since A causally predicts the onset of O in time, integrating information about both events according to their relative reliability (rather than considering information about each event in isolation) would not only help reducing noise in P's temporal estimates of A *and* O. It would also attract both estimates towards one another (and thus induce intentional binding), while the estimate with the larger relative variance would get attracted more strongly (compare the framework of "optimal cue integration"; e.g., see (11)). In fact, a recent study demonstrated that the lower the signal-to-noise ratio of a sensory outcome O that followed a button press after a brief delay, the more the perceived onset of O was attracted towards A and the lesser A was attracted towards O (12). Importantly, such cue integration should happen only if the action causes the sensory event (A -> O) and action and event do not occur independently. Related Bayesian causal inference models of cue integration therefore further consider the prior probability of an action causing an event (vs. their independence) (13, 14). For example: One would not integrate one's own key press on a computer with thunder and lightning outside the window. Compatible with these latter models, Desantis and colleagues have shown that strengthening subjects' causal beliefs in self-agency does in fact increase intentional binding (15).

Given this influence of causal attribution on intentional binding, we here asked whether causally linking one's own actions to their sensory consequences is necessary for the emergence of intentional binding. Studying conditions, in which the SoA is pathologically altered, could allow critically testing predictions put forward by Bayesian causal inference. A respective condition is delusions of control in schizophrenia (DoC; sometimes also referred to as delusions of influence) (16–20). Many patients with DoC experience their actions as being controlled by someone or something else. Hence, their perceived agency is pathologically reduced and, accordingly, these patients should exhibit weaker intentional binding. In turn, there are also schizophrenia patients with grandiose delusions, i.e., delusions about having inflated power, worth, knowledge, etc. (21), that have an exaggerated SoA, which thus should lead to increased intentional binding. So far, there is some evidence for altered intentional binding in

Schizophrenia (22–25). Yet, the interpretation of these studies within the Bayesian framework is difficult because studies which acquired both explicit measures of perceived agency during the execution of the experimental tasks and individual pathology measures are missing. Thus, it is unclear whether these patients' perceived agency was diminished, exaggerated, or unaltered.

We here aimed to establish a direct link between individual psychopathology (DoC), perceived agency in our tasks, and the strength of temporal binding. To this end, we investigated schizophrenia patients with and without DoC as well as healthy controls. We hypothesized that DoC are accompanied by a reduced explicit self-attribution of agency in our task. Moreover, following the logic of Bayesian causal inference, we expected that the strength of patients' DoC should correlate with decreases in intentional binding. Accordingly, we predicted that there should be less intentional binding in patients suffering from DoC compared to healthy controls and to Schizophrenia patients that do not display this symptom.

Results

To test our predictions experimentally, we measured temporal binding in a group of 20 healthy controls and 20 patients with schizophrenia (see Methods for a power analysis and for further details about our participant groups). The experiment consisted of three different conditions, all of which required participants to predict the time of a *target-LED* flash by pressing a target button at the anticipated time of the flash (time of button press [TBP]). Depending on the experimental condition, the target-LED flash would occur with a fixed temporal interval, either (I) after a participant's active press of a start button in the *self condition*, (II) after an observed press of the start button by a machine (*machine condition*), or (III) after an observed *signal-LED* flash (*baseline condition*, Fig. 1a; also see SI Appendix, Fig. S1 for more details). Hence, in (I) the participants' intentional action caused the target-LED flash, in (II) it was caused by the machine, and in (III) there was neither an intentional link between participants' behavior and the LED flashes nor was there any other obvious causal interrelation. Please note that we minimized any other differences between our experimental conditions (compare Methods for details). Based on previous research, we

153 expected that healthy participants would perceive the target-LED flash as
154 happening earlier in time in the *self condition* compared to the *baseline condition*
155 due *to temporal binding* (5). Hence, subjects' TBP should be earlier in the *self*
156 *condition* as compared to *baseline* (i.e. $TBP_{self} < TBP_{baseline}$; Fig. 1b). Following the
157 nomenclature of earlier studies, we refer to the expected temporal-difference
158 effect as *intentional binding*. Note, however, that intentional as well as mere causal
159 relationships could contribute to this effect (26–28). Because of this, we
160 introduced the *machine condition*, which – as compared to (1) – supposedly lacked
161 any intentional component. The machine condition was important, as it allowed
162 us to control whether any changes in patients' intentional binding could be
163 accounted for by changes in *causal binding* or in perceived causality. Causal
164 binding would be revealed if participants perceived the target LED flash earlier in
165 the *machine condition* than in the *baseline condition* ($TBP_{machine} < TBP_{baseline}$; Fig.
166 1b). Thus, our experimental design enabled us to independently assess the
167 amount of temporal binding as a function of the intentional and/or causal
168 relationship between cause (pressing the button) and effect (the LED flash), and
169 to investigate the predicted changes of intentional binding in schizophrenia
170 patients with DoC.
171
172 Figure 1
173
174 All experimental conditions (self, machine, baseline) were performed in separate
175 experimental blocks (60 trials each). Within each block, the flash of the target-LED
176 always occurred after a fixed target interval. In each condition we obtained one
177 block with a 500 ms interval and a block with a 700 ms interval. The 700 ms
178 intervals merely served to vary the temporal intervals across blocks and to
179 thereby prevent learning effects (note that the results for the 700 ms interval are
180 shown exclusively in our SI Appendix as robust temporal binding is only expected
181 to occur for shorter intervals (5, 27)). Sequences of blocks were pseudo-
182 randomized within each subject-group but matched between control- and patient
183 group.
184 In every single trial and for every experimental condition, subjects had to estimate
185 the temporal onset of the target-LED flash as precisely as possible by pressing a

response button at the respective time (so that, ideally, the flash and the button press would coincide). The first 15 trials of each block were considered training trials that should allow participants to adjust their responses to the current interval and we excluded these trials from later analyses. In a random subset (constituting 33%) of the remaining 45 trials, we overtly blocked participants' view of the target-LED before the beginning of a trial. These '*no feedback trials*' (NFT) served to prevent subjects from merely responding in a reactive fashion to the target-LED flash and to force them to employ a predictive response strategy. In the majority of trials (66%), however, feedback about the target-LED flash was present ('*feedback trials*'; FT; Fig. 1a). Note that comparing the TBP as a function of feedback availability further allowed us to probe for any shift from a predictive response pattern in no-feedback trials (i.e. TBP ≤ 0 ms) to a reactive response pattern whenever feedback was provided (i.e. TBP >> 0 ms). Given the evidence of impaired predictive processes (forward models) in patients with schizophrenia, we expected a more reactive response pattern in patients to compensate for this deficit if feedback is available (23, 29).

Finally, we requested explicit estimates of SoA and of the perceived causal relationships in our experiment by means of a questionnaire (also compare Methods).

Agency Questionnaire

While exhibiting a correlation between DoC and decreases of intentional binding was the main aim of our work, we first wanted to probe how healthy subjects as well as Schizophrenia patients perceived self-agency and causality in our experimental conditions. Since we expected a deficit in perceived self-agency exclusively in patients suffering from DoC, for simplicity we present the patient group as two subgroups, 10 DoC-patients and 10 NoDoC-patients, based on the presence of DoC (for further details and for the definition of DoC see Methods; note that patient-grouping was done only after collecting all data and that the experimenter was not aware of the patients' symptoms while carrying out the experiment). Apart from the group-defining positive symptomatology, there were no statistical differences between these subgroups in terms of age (independent samples t-test, two-tailed: t = -0.751, p = 0.462), education (t = 0.447, p = 0.66),

negative symptoms (30) (t = .0798, p = 0.438), general psychopathology (31) (t = 1.607, p = 0.129) or medication (t = -0.18, p = 0.859; see SI Appendix, Table S4 for information on the patient group).

To assess perceived self-agency and causality in our tasks we used a questionnaire which each participant answered separately for every condition after the experiment. To investigate the subjects' explicit assumptions of causation and agency concerning the target LED flash, we asked: "Who or what determined the time of the target LED flash?". The answers participants could choose from were 'I', 'the experimenter', 'the computer', 'the machine', 'the button box', and 'nothing/no one'. Afterwards, participants had to rate on a Likert scale how confident they felt when giving their answer, ranging from 1 (= not certain at all) to 5 (= completely certain). Since we did not reveal any group differences in our certainty measure (compare SI Appendix, Table S2; also compare (32)), we here report a combined response index for each subject group and experimental condition, i.e. self-, machine-, and baseline condition, and for each of the corresponding answers of interest, namely "I", "the machine", and "the computer", respectively. This index was calculated by multiplying the average share of the response (e.g. "I") and the average of its corresponding confidence estimates within the respective group of participants. Accordingly, this index can range from 0 (0% "I") to 5 (100% "I" times a confidence level of 5). According to our experimental setup and instructions, in healthy subjects this index should be the highest for "I" in the self condition, for "the machine" in the machine condition, and for "the computer" in the baseline condition. Visual inspection of Figure 2 and statistical analyses (also compare SI Appendix Table S1) show that this was clearly the case.

Figure 2

In line with the hypothesis of their reduced sense of agency, the pattern of results depicted in Figure 2 markedly differed between DoC-patients and healthy controls. Moreover, while the response index of NoDoC-patients seemingly resembles that of controls in both the self and the baseline condition, their pattern of results is more similar to that of DoC-patients in the machine condition. We

compared the amount of each "correct" answer in the respective condition ("I" in self condition, "machine" in machine condition, "computer" in baseline condition) between controls and the two patient subgroups using Mann-Whitney U tests (Bonferroni corrected $\alpha = 0.025$). As expected, there were no statistically significant differences in the baseline condition (response computer) between controls (3.3 ± 0.5) and either patient subgroup (DoC = 3.7 ± 0.6, $U = 86.5$, $p = 0.528$; NoDoC= 2.3 ± 0.8, $U = 82.5$, $p = 0.416$). In the self condition, however, the index of the answer "I", i.e., signifying an intentional involvement in the task, was significantly smaller (namely at zero) for DoC-patients compared to the control group (1.8 ± 0.5, $U = 60$, $p = 0.023$, $\eta^2 = 0.18$). There was no such difference between NoDoC-patients (1.0 ± 0.3) and healthy controls ($U = 83$, $p = 0.368$). Finally, the index for the (correct) answer "machine" in the machine condition was either significantly higher or with a strong trend in controls (2.6 ± 0.6) compared to both patient subgroups (DoC = 0.5 ± 0.2, $U = 56.5$, $p = 0.028$, $\eta^2 = 0.17$; NoDoC = 0.4 ± 0.1, $U = 51$, $p = 0.014$, $\eta^2 = 0.21$). Note that statistical analyses of participants' responses without multiplying them with the certainty measure led to exactly the same qualitative results. In summary, both patient subgroups did not differ from controls in their judgments concerning the baseline condition. Conversely, in the machine condition neither patient subgroup felt the machine to be responsible for the target LED flash. More importantly, compatible with the hypothesized deficit in the self-attribution of agency, patients with DoC did not feel responsible for the target LED flash in the self condition. As Figure 2 suggests, they instead assigned agency to the computer. We therefore statistically compared the index of the (wrong) answer "computer" in the self condition which was significantly higher in DoC-patients (3.1 ± 1.0) as compared to the control group (1.15 ± 0.4, $U = 50.5$, $p = 0.015$, $\eta^2 = 0.2$).

Temporal Binding Experiment

To estimate when our participants expected the target LED to flash, we calculated the median TBP relative to the actual time of the target LED flash in each individual, i.e., negative numbers indicate button presses before, positive numbers after the actual target LED flash. According to our hypothesis and in agreement with subjects' explicit agency ratings, we expected to reveal intentional (i.e.,

285 TBP_{self} < $TBP_{baseline}$) as well as causal binding (i.e. $TBP_{machine}$ < $TBP_{baseline}$) in the
286 control group. Figure 3 (see SI Appendix, Fig. S2 for performance of individual
287 participants**)** confirms our expectation: healthy participants exhibited the earliest
288 button presses in the self condition (as is indicated by more negative TBPs),
289 followed by the machine condition and the baseline condition. Performance in the
290 latter condition was highly accurate with TBPs close to zero.
291 In the healthy control group intentional binding (TBP_{self} – $TBP_{baseline}$) was
292 significant ($F_{1,19}$ = 26.6, p = <.001, η^2 = 0.59) and amounted to -52 ± 7 ms (mean ±
293 SEM.) in FT and -28 ± 11 ms in NFT across subjects. Causal binding ($TBP_{machine}$ −
294 $TBP_{baseline}$) was likewise significant ($F_{1,19}$ = 5.088, p = 0.036, η^2 = 0.21) and
295 averaged -20 ± 6 ms in FT and -11 ± 10 ms in NFT. SI Appendix Fig. S4 provides an
296 additional overview of these binding estimates (and for the 700 ms interval
297 conditions). In addition, it provides Bayes-factors (BF; also compare Methods)
298 expressing evidence in favor (BF > 1) vs. against (BF < 1) the presence of temporal
299 binding. Evidence for intentional binding in the 500 ms interval ranged from
300 strong (BF_{NFT} = 13.7) to decisive (BF_{FT} = $10^{11.3}$). Evidence was very strong for
301 causal binding in FT (BF = 67.2) but indecisive for NFT (BF = 0.50). These results
302 suggest that our task was well suited to induce intentional binding in healthy
303 controls.
304
305 Figure 3
306
307 Like for controls and in agreement with their explicit agency ratings, intentional
308 binding within the NoDoC-patient subgroup was significant and supported by
309 strong to very strong evidence ($F_{1,9}$ = 10.071, p = 0.011, η^2 = 0.53; -80 ± 28 ms in
310 FT [BF = 17.9] and -65 ± 22 ms in NFT [BF = 32]). Causal binding estimates were
311 numerically smaller than intentional binding in NoDoC and evidence in
312 favor/against causal binding was indecisive (2 ± 23 ms in FT [BF = 0.38] and -22
313 ± 20 ms in NFT [BF = 1.29]). Accordingly, causal binding also failed to reach
314 significance ($F_{1,9}$ = 0.253, p = 0.627). Importantly, intentional binding in the DoC-
315 patient subgroup showed an almost reversed pattern compared to both controls
316 and NoDoC patients: intentional binding estimates in these patients were 25 ± 27
317 ms in FT and 30 ± 26 ms in NFT and were not significant ($F_{1,9}$ = 1.239, p = 0.295).

Note that positive numbers indicate a temporal repulsion rather than binding. This lack of significance was accompanied by substantial evidence for the absence of intentional binding in both FT (BF = 0.24) and NFT (BF = 0.20). Causal binding estimates were not significant ($F_{1,9} = 0.496$, $p = 0.499$) and likewise in the opposite direction with 10 ± 23 ms (FT) and 16 ± 19 ms (NFT). Evidence against the presence of causal binding also was substantial ($BF_{FT} = 0.28$; $BF_{NFT} = 0.16$). Note that the aforementioned results on temporal binding within each subject group are consistent with our hypotheses. Yet, they ultimately do not allow to assess whether or not the expected differences between subject groups were present (compare (33)).

Before further comparing groups for differences in temporal binding across conditions, we first wanted to make sure that across the three groups, subjects were able to provide equally accurate temporal estimates. To this end we compared the temporal estimates in the baseline condition and in the absence of feedback (NFT), as we did not expect to see any systematic group difference in this condition. A one-way ANOVA confirmed this expectation as there were no significant differences in the Baseline condition (NFT) between groups ($F_{2,37} = 1.077$, $p = 0.354$).

To statistically analyze temporal binding across groups, we conducted a mixed-design 3 x 2 x 3 ANOVA with the between-subject-factor group [controls, NoDoC-patients, DoC-patients] and the within-subject-factors feedback [FT, NFT] and condition [self, machine, baseline]. We found main effects of condition ($F_{2,74} = 5.33$, $p = 0.007$, $\eta^2 = 0.13$) and feedback ($F_{1,37} = 17.9$, $p < .001$, $\eta^2 = 0.33$), a condition x group interaction ($F_{4,37} = 4.92$, $p = 0.002$, $\eta^2 = 0.2$) and a feedback x group interaction ($F_{2,37} = 8.00$, $p = 0.001$, $\eta^2 = 0.3$). Additionally, we found a significant linear effect of condition, i.e. an increase from the *self-*, over the *machine-*, to the *baseline condition,* on the TBP ($F_{1,37} = 8.68$, $p = 0.006$, $\eta^2 = 0.19$) and an interaction between this effect and the factor group ($F_{2,37} = 7.92$, $p = 0.001$, $\eta^2 = 0.3$). To identify the specific group differences, which led to the interaction effects between condition and group as well as between feedback and group, we performed three corresponding post hoc 2 x 2 x 3 ANOVAs that allowed pairwise comparisons between subgroups (Bonferroni corrected $\alpha = 0.017$). These analyses led to the following results: Apart from a feedback x group interaction ($F_{1,28} = 11.72$, p =

0.002, $\eta^2 = 0.3$), there were no differences between NoDoC-patients and healthy controls (condition x group: $F_{2,28} = 2.73$, $p = 0.078$; condition x feedback x group: $F_{2,74} = 1.66$, $p = 0.2$). DoC-patients did, however, clearly differ from both other groups in that there were significant interactions between the factors condition and group in both respective analyses (DoC-patients vs. controls: $F_{2,28} = 5.23$, $p = 0.008$, $\eta^2 = 0.16$; DoC-patients vs. NoDoC-patients: $F_{2,18} = 5.13$, $p = 0.01$, $\eta^2 = 0.22$). The group x condition interaction was likewise present for the linear effect estimate (DoC-patients vs. controls: $F_{1,28} = 11.09$, $p = 0.002$, $\eta^2 = 0.28$; DoC-patients vs. NoDoC-patients: $F_{1,18} = 8.84$, $p = 0.008$, $\eta^2 = 0.4$). When contrasting DoC-patients and controls, there was a significant interaction between the factors group and feedback too ($F_{1,28} = 14.19$, $p = 0.001$, $\eta^2 = 0.34$). Remember, the same qualitative effect was present also when comparing NoDoC-patients and healthy controls (see above). It was absent, however, when contrasting the two patient subgroups ($F_{1,18} = 0.003$, $p = 0.957$).

These results show that intentional binding was significantly altered in Schizophrenia patients with DoC as compared to patients without DoC and as compared to healthy controls. This pattern of results well reflected the explicit agency ratings of our subject groups. The latter is true also for causal binding, which was absent for both subgroups of Schizophrenia patients. Finally, the availability of feedback had a clear influence on patients' temporal estimates across all tasks, resulting in TBPs that were significantly delayed (and thus more reactive). This effect was again present in both subgroups of patients.

SI Appendix Figure S3 provides additional information about the variability of subjects' temporal estimates across groups and conditions. Similar to the more reactive response pattern of both patient groups, as was described before, also the variability of patients' estimates was higher than that of controls. There were no further differences across groups that could have accounted for the group differences in intentional binding. Yet, there was a global effect of condition, which was explained by decreasing variabilities from the baseline- over the machine- to the self- condition.

Finally, we performed additional Bayes-factor analyses to further support the observed similarities and differences in intentional binding across groups. First, this analysis revealed substantial evidence for the absence of group

differences (BF < 1) between NoDoC-patients and healthy controls for intentional binding (BF_{FT} = 0.14; BF_{NFT} = 0.11; compare SI Appendix, Table S3). Second, and most importantly, there was substantial to very strong evidence for intentional binding of both aforementioned groups being different from that of DOC-patients (BF > 1. DoC-patients vs. controls: BF_{Ft} = 55.95; BF_{NFT} = 5.09. DoC-patients vs. NoDoC-patients: BF_{FT} = 7.75; BF_{NFT} = 8.79). As to be expected, complementary analyses of causal binding revealed that evidence in favor/against group differences were largely indecisive (compare SI Appendix, Table S3).

Correlations with Psychopathology

In a final step, we wanted to investigate the relationship between individual behavior and psychopathology through linear correlation analyses. The idea of these analyses was to exhibit specific links between (deficits in) intentional/causal binding and different classes of positive symptoms. As we already introduced in the introduction, we expected a decrease in intentional binding with an increasing strength of DoC. Yet, as at least certain hallucinations are also interpreted through an impaired SoA (e.g. see (34)), hallucinations might as well be correlated with intentional binding. Another advantage of such within-patient analyses is that they avoid systematic differences present between patients and controls (such as medication, etc.). Specifically, we performed correlation analyses between the respective temporal binding measures (i.e., intentional Binding: $TBP_{self} - TBP_{baseline}$; causal binding: $TBP_{machine} - TBP_{baseline}$) and all patients' individual measures of psychopathology, as derived from the SAPS score (35). These measures included the intensity of patients' hallucinations (SAPS item I), delusions (SAPS item II), delusions of control (a subscore of SAPS item II, here denoted as item IIa; compare Methods), residual delusions (SAPS items II – IIa), and residual positive symptoms (SAPS items III, IV, V). We thereby were particularly interested to see, whether a lesser confidence in self-agency, as quantified by the subscore assessing DoC, would correlate with intentional binding, causal binding, or both.

Figure 4

Using Kendall rank correlation, we found that DoC were indeed significantly correlated (Bonferroni corrected $\alpha = 0.01$, two-tailed) with the amount of intentional binding in both FT ($\tau_b = 0.515$, $p = 0.003$) and NFT ($\tau_b = 0.532$, $p = 0.002$). The more severe the DoC in a patient, the less intentional binding was exhibited (Fig. 4a). In addition, the overall strength of delusions (which DoC are part of) was significantly correlated with intentional binding, but in NFT only ($\tau_b = 0.495$, $p = 0.003$). There were no correlations between patients' symptomatology and causal binding (Fig. 4b). Furthermore, current medication, measured as Olanzapine equivalents, did not exhibit any correlation with intentional (FT: $\tau_b = 0.123$, $p = 0.454$; NFT: $\tau_b = 0.245$, $p = 0.135$) or causal binding (FT: $\tau_b = 0.059$, $p = 0.72$; NFT: $\tau_b = 0.171$, $p = 0.297$). In summary, our correlation analyses clearly support the hypothesized intentional binding deficit in patients suffering from DoC. There was no such correlation between the strength of DoC and causal binding.

Reaction Time Control

In order to test for general differences in reaction time between groups that could have systematically affected our TBP measure, we ran a simple manual reaction time task before the actual experiment. For every participant we computed the median reaction time. The average reaction time for controls was 238 ± 12 ms (mean ± SEM). Both patient subgroups (DoC and NoDoC) had comparable reaction times (228 ± 6 ms and 273 ± 21 ms, respectively). There were no significant differences in reaction time between patient subgroups ($U = 24$, $p = .094$) or between patient subgroups and controls (DoC: $U = 81.5$, $p = 0.856$; NoDoC: $U = 60$, $p = 0.112$). There was also no influence of medication on patients' reaction time ($\tau_b = -0.012$, $p = 0.944$). Hence, our main results cannot be explained by a difference in subjects' ability to provide timely responses.

Discussion

We showed that the unique experience of patients with schizophrenia suffering from delusions of control (DoC) – the feeling that someone or something else is controlling their very own movements – is associated with absent temporal binding between their intentional actions and these actions' effects. More

precisely, as we could show by means of our questionnaire, altered perception due to DoC was present in the respective patients while performing the self condition. Not only was these patients' perceived self-agency significantly reduced, but at the same time they were convinced that the computer caused the LED flash. In agreement with the predictions put forward on the basis of Bayesian accounts of intentional binding, this alteration of DoC patients' SoA was accompanied by an absence of temporal binding. This relationship between DoC and intentional binding was further supported by specific correlations between symptom severity and reductions in intentional binding. Thus, our results provide support for a relationship between intentional binding and the conscious experience of agency. Importantly, the performance in the reaction time task along with the equally precise response accuracies across groups in the baseline condition without feedback show that the patients were able to perform the task and that any between-group differences could not be explained by general motor or perceptual impairments, or medication. This was particularly important to demonstrate, as our study solely studied the binding of outcomes to actions (but not vice versa).

So far only a few studies applied temporal binding paradigms in Schizophrenia patients (22, 23, 25). Notably, the earliest of these studies reported a "hyperbinding" effect, namely a stronger intentional binding in patients with schizophrenia compared to healthy controls (22) (also compare (23)). While such hyperbinding might appear in stark contrast to our findings, it is on the contrary not unexpected: Hyperbinding might result from an exaggerated SoA in case of grandiose delusions, i.e. delusions about having inflated power, worth, knowledge, etc., which affect almost half of patients with schizophrenia (21). Unfortunately, the aforementioned binding study lacks a detailed characterization of patients' individual psychopathology, which could have helped clarify this point. Moreover, an explicit rating on how subjects perceived agency in the respective experimental tasks was missing in all earlier studies that reported hyperbinding (22, 23). We now could show that patients with DoC did lack a conscious feeling of "being in control" in our intentional binding task. At the same time, these DoC patients did lack intentional binding. Whether grandiose delusions would lead to a respective hyperbinding between actions and effects was not at the focus of our study and is still open. However, others have at least shown that the same subjects with

putative psychotic prodrome who showed hyperbinding (24), also exhibited an explicit over-attribution of external events to the self (though in an independent task) (36). Hence, consistent with predictions of Bayesian frameworks of intentional binding (10, 13), there is a direct link between increases/decreases of intentional binding in psychosis and an exaggerated/attenuated likelihood of perceived self-agency, respectively.

Importantly, we do not want to suggest the conscious judgement of agency as being the decisive factor for the occurrence of intentional binding. Instead, we suggest that both intentional binding and subjects' perceived agency largely (but not only) depend on the outcome of an operation in the central nervous system that allows it to differentiate between self- and externally- caused sensory inputs. This basic operation provides the basis of perceptual stability (37) and precise motor control (38), among many other things. Specifically, self- and externally produced sensory inputs can be distinguished by the comparison of an internal forward model, predicting the sensory consequences of one's actions, with the actual sensory information available. In case of a match, self-agency would be assumed; in case of a mismatch, the residual sensory information would be attributed to external causes (6, 38–40). Through such an automatic comparison it is possible to infer one's own influence on the environment and this comparator is therefore considered a crucial mechanism informing the SoA (6, 18, 41). Please note that on an abstract level of description, forward models solve a subset of the same problems addressed by Bayesian causal inference models, namely, to infer the most likely causal structure that explains sensory information (14). Moreover, forward models themselves can be thought of "complex" priors in Bayesian terms (42, 43).

Accordingly, disorders of agency, as observed in Schizophrenia patients with delusions of control could be a result of an impaired forward model mechanism or of an altered prior. In fact, previous studies showed that forward model predictions about the consequences of self-action are unreliable in patients with schizophrenia suffering from DoC (18, 19, 29, 44). Moreover, the stronger patients suffered from DoC, the less reliable their forward models (29, 44). Consequently, imprecise forward models (or priors) could reflect a common cause that led to reduced explicit agency ratings in our patients with DoC as well as to

their lack of intentional binding. Unfortunately, our study does not allow verifying this assumption experimentally, as we did not monitor the precision of subjects' internal models in our experiment. Yet, this will be an interesting endeavor for future research. Such research could also address whether intentional temporal binding is merely a reflection of a SoA, whether it is a consequence of the SoA, or, lastly, whether it could even further perceived self-agency by increasing temporal contiguity between actions and their effects. If intentional binding is at least partially dependent on forward models (and not only on the perceived SoA), a furthering of the SoA through temporal binding seems conceivable.

While perceived agency (and, supposedly, precise internal models) do obviously further intentional binding, temporal binding can also be present while observing externally caused events (26–28). We were able to capture these specific contributions of perceived causality during intentional actions compared to perceived causality in the absence of self-action through our measures for intentional and causal binding, respectively. In particular, our experiment consisted of two 'causal conditions', namely the self condition and the machine condition, but only the self condition was also an 'intentional condition'. In contrast, the baseline condition mimicked a case in which there was a correlation (in time) between two events (as in all other conditions), but where there was no obvious causal (and/or intentional) interrelation between these events. In the control group we observed significant temporal binding between causes (button presses) and events (LED flashes) in the self- and in the machine condition. Hence, temporal binding *per se* cannot be considered a proxy of the SoA – as the effect might solely be based on perceived causality in the absence of any intentional action. However, this does not preclude the additional presence of an intentional binding component in the self condition. In fact, our measures of causal and intentional binding differed significantly, with intentional binding being the larger one. Such an 'intentional boost' has also been observed in previous research (27, 28). The presence of an "intentional boost" could be easily explained by the availability of additional (causal) cues in the self vs. the machine condition such as proprioception and forward models. The significant reduction of variability in subjects' time estimates from the baseline over the machine to the self-condition is compatible with this interpretation (compare SI Appendix, Fig. S3 and SI

549 Appendix, supplementary Discussion). It is an interesting question, though,
550 whether it would make a difference in temporal binding whether an external
551 agent or the machine elicited a sensory event via a button press. So far, we only
552 know that temporal binding is also present when observing actions (and their
553 consequences) of external agents (45). Yet, whether such binding could solely be
554 accounted for by "external causal cues" or, in addition, could be explained by
555 attributing intentionality to the external agent is still open. In any case, it is the
556 intentional boost over and above mere causal "machine binding" which should be
557 considered as an implicit marker of the SoA (or of the attribution of a SoA to
558 another agent).
559 Interestingly, both groups of Schizophrenia patients had a significantly
560 lowered response index for the machine in our causal control condition and,
561 accordingly, did not exhibit causal binding. Furthermore, there was no correlation
562 between causal binding and DoC. This does not rule out the possibility that
563 Schizophrenia patients also exhibit deficits in other causal inference mechanism,
564 as were suggested by earlier research (46). Ultimately, our study was not designed
565 to this possibility. However, it at least suggests that the observed deficit in
566 intentional binding in patients with DoC is not secondary to a deficit in causal
567 binding but does occur independently.
568 Apart from the deficit in intentional binding, which was specific to
569 Schizophrenia patients with DoC, we additionally exhibited a strong effect that
570 generalized across tasks and patient subgroups. In all patients, we found that the
571 time of the target button press was significantly delayed in FT (i.e., when the target
572 LED was visible) and, therefore, patients exhibited a more reactive response
573 pattern as compared to controls. This result was not unexpected – at least for
574 patients with DoC: Since internal forward models about self-actions are unreliable
575 in patients with DoC (19, 29, 41, 44, 47), these "priors" should weigh less than
576 external sensory cues (29, 48, 49). Accordingly, in our task patients might delay
577 their responses until more sensory evidence is available. In fact, during feedback
578 trials (FT) patients would delay their button presses up to a time at which reactive
579 responses to the LED target flash were more likely to occur. During NFT, on the
580 other hand, patients had to base their temporal estimate solely on internal cues
581 (as there was no visual feedback). The absence of feedback led to mean response

582 times more similar to those of controls, but with higher variability (compare Fig.
583 3 and SI Appendix, Fig. S3). This also well resembles the fact that forward models
584 in Schizophrenia are imprecise but (on average) are as accurate as those of
585 controls (29) and that respective patients cannot benefit from intact Bayesian
586 causal inference, accordingly. Importantly, there is converging evidence for such
587 pathologically increased integration of retrospective sensory information from a
588 study investigating the influence of predictive and retrospective mechanisms on
589 intentional binding in Schizophrenia: while temporal binding in healthy subjects
590 depended on action-effect predictability, binding in Schizophrenia patients solely
591 built on the presence (vs. the absence) of a sensory action effect (23). The authors
592 of this study likewise concluded that patients' "experience of agency appeared to
593 be driven by immediate sensory evidence (…), without any reference to an
594 internal model" ((23), p. 3110). One remaining question is why the feedback effect
595 in our study was visible in all conditions and in all patients (with and without
596 DoC). One explanation could be that there is a more global deficit in forward
597 models in Schizophrenia patients with positive symptoms. In fact, internal
598 forward models not only inform about the consequences of self-action. The
599 nervous system also represents and updates forward models about external
600 events (e.g., compare (50)). It might well be that forward models (or priors) are
601 more generally impaired in Schizophrenia patients with positive symptoms (16,
602 51), an idea that is also reverberated by recent Bayesian accounts of psychosis in
603 Schizophrenia ((52–54); compare below). The general increase in response
604 variability across all patient groups, irrespective of the experimental task (SI
605 Appendix, Fig. S3), could certainly be explained through the additional presence
606 of disease/medication in patients. However, this finding is also compatible with
607 the aforementioned idea of a more general deficit in forward models or priors.
608 Though being an empirical study, our work was critically inspired by
609 qualitative predictions of Bayesian perceptual inference, and it allowed for
610 verifying some of these predictions experimentally. Throughout our work we also
611 highlight how these Bayesian accounts could, in turn, help to explain delusions of
612 control in Schizophrenia. Yet, rather than being a narrow framework that can only
613 explain a rather specific class of symptoms (such as "classical" forward model
614 approaches did only explain delusions of control and certain hallucinations (16,

34)), Bayesian perceptual inference allows quantitatively capturing and predicting a larger range of psychotic symptoms within the same general framework (for reviews compare (52–54)).

In conclusion, patients with DoC did misperceive agency in our intentional binding task, namely that they did not feel responsible for causing the LED flash through their button press. At the same time, they also did not exhibit intentional binding in the respective condition. These results agree with predictions put forward by Bayesian models of intentional binding that suggest that the likelihood of self-agency has a direct impact on the amount of intentional binding. Importantly, this relationship was further supported by the correlation between DoC symptom strength and reductions in intentional binding. Building on previous research on Schizophrenia patients, we propose that imprecise forward models (or priors) could explain both, the pathological alteration of perceived agency in DoC as well as the deficit in intentional binding. We further suggest that by enhancing the perceived spatiotemporal contiguity between one's actions and their sensory consequences, temporal binding could strengthen the subjective experience of self-agency. Absent temporal binding could, in turn, be another cause for the loss of an *agentive self,* as is experienced by numerous patients with Schizophrenia.

Methods and Materials

Participants

Sample size was guided by a power analysis that built on our previous research in schizophrenia patients, exhibiting a tight correlation between the strength of delusions of control (DoC) and the (un)reliability of forward models (29, 44). Note, that we assumed that such deficits in forward models also reflect the basis for altered intentional binding (compare Discussion). Based on the average effect size in these previous studies ($r = 0.64$), the estimated sample size amounted to 20 Schizophrenia patients, given an alpha-level of 0.01 (one-tailed; due to our prior hypothesis that intentional binding decreases with the strength of DoC) and a power of 0.8. Initially, we therefore recruited a total of 21 patients with schizophrenia and 22 matched (see below for details) controls. One patient was excluded because in the debriefing it became apparent that he did not understand

the task correctly. One control had to be excluded because he only reacted to the target LED flash as opposed to predicting it, which he described himself in the debriefing after the experiment. Another control was excluded because of a diagnosed psychiatric disorder other than schizophrenia. This led to a group of 20 patients with schizophrenia spectrum disorders (6 females, 14 males; age 33.8 ± 2 years (± s.e.m.); 12 ± 0.3 years of education (primary + secondary school); DoC subgroup: 3 females, 7 males; age 32.3 ± 1.4 years; 12.1 ± 0.3 years of education; NoDoC subgroup: 3 females, 7 males; age 35.3 ± 3.5; 11.8 ± 0.5 years of education) and 20 age-matched healthy controls with equal levels of education (6 females, 14 males; age 33.3 ± 2.4 years; 11.9 ± 0.3 years of education; see SI Appendix, Table S4 and S5) that were used for analysis. All subjects had normal or corrected-to-normal visual acuity and gave their written informed consent. Patients with schizophrenia from in- and outpatient treatment at the Department of Psychiatry and Psychotherapy of the University of Tübingen were recruited from the Psychiatric University Hospital Tübingen, Germany. The local ethics committee approved the study. Patients were eligible for participation when they fulfilled the diagnostic criteria of schizophrenia or schizoaffective disorder (patients # 1, 3, 4, 6, 11, 14; compare SI Appendix, Table S4) according to DSM-IV, diagnosis was confirmed by a structured clinical interview (SCID-I). All patients had stable medication for at least one week. Further inclusion criteria were no mental retardation or current substance use disorder. All of the patients were treated with second-generation antipsychotics.

Hallucinations and delusions were quantified by the Scale for the Assessment of Positive Symptoms (SAPS) (35). The mean SAPS rating amounted to 19.7 ± 5.2. We additionally assessed the following subscores: score I hallucinations (SAPS questions 1 to 7), 3.3 ± 1.6; score II delusions (questions 8 to 20), 11.4 ± 2.8; score IIa delusions of control (questions 15 to 19; as defined previously [6]), 3.8 ± 1.4; score IIb residual delusions (score II – score IIa, i.e., questions 8 to 14 and 20), 7.6 ± 1.6; score III to V residual positive symptoms (questions 21 to 35), 5 ± 2.1. The SAPS was acquired by an independent clinician in close temporal proximity to the experiment (either on the same day or +/- one day). Importantly, the experimenter was blind to this score when performing the experiment.

Experimental Setup

All experiments were conducted on a MacBook Pro (mid 2010 model) using PsyScope X (55) and an ioLab Systems USB (www.iolab.co.uk) response box. The machine that we used in the machine condition was custom built and has been described in detail elsewhere (27). In order to diminish auditory cues, we had our participants wear noise-isolating earmuffs during the whole experiment.

Experimental Design

We told participants that their main task throughout the whole experiment was to predict the time of the target LED flash with a press on the target button so that, in an ideal case, both of these events would coincide. The predictive cues that informed subjects' temporal judgments varied across our three experimental conditions, as will be detailed below. Accordingly, we provided our subjects with detailed instructions for each of the three different conditions (Fig. 1 and SI Appendix, Fig. S1). Each condition was performed twice in dedicated experimental blocks, once with a 500 ms interval between the predictive cues and the target LED flash and once with a 700 ms interval. We included the 700 ms intervals to diminish learning across blocks. Given that previous research suggested robust temporal binding to only occur for shorter intervals (5, 27), the results from the 700 ms blocks are shown exclusively in our SI Appendix. Each block consisted of 60 trials. The first 15 of these trials were considered training trials and provided visible feedback about the target LED flash (feedback trials FT). The remaining 45 trials were comprised of 30 FT and 15 no feedback trials (NFT). We included NFT trials to ensure that (at least in this subset of trials) subjects' temporal estimates of target LED onset would not be reactive (i.e., be triggered by the actual LED flash in a given trial). FT and NFT trials were presented in randomized order. The sequence of six experimental blocks (3 conditions x 2 temporal intervals) was pseudorandomized within each subject-group but matched between control- and patient group. Before the experiment, each participant completed a training session with four trials (three FT and one NFT) of each condition.

Self Condition

In order to start the trial participants had to press the initialize button with their

left hand, which was the leftmost button (on the left side) of the button box (SI Appendix, Fig. S1a). After a variable delay between 2500 to 2750 ms, participants could press the start button with their left hand. This button press signaled the start of the interval until the flash of the target LED (500 or 700 ms, depending on the current block) that subjects had to predict. Note that whenever the start button was pressed the *signal LED* was flashed to make conditions more similar (see baseline condition). After participants pressed the target button with their right hand in order to report the estimated onset of the target LED flash, a new trial could be initiated by pressing the initialize button on the left.

In case the current trial was a NFT, the experimenter blocked the subjects' view of the target LED with a piece of black cardboard *before* the beginning of the trial. Hence, in NFT participants would always know in advance that during the following trial no visual feedback from the target LED would be visible.

Machine Condition

In the machine condition, in order to start the trial, participants had to press a button on the machine with their left hand that 'initialized' the machine (the button was located on the left side, mimicking the spatial layout of the other two conditions; SI Appendix, Fig. S1b). Participants were told that by pressing the initialize button, a random number generator inside the machine was activated which would lead to a press of the machine lever on the start button after a randomly selected interval. To visually support this mechanism, the press on the initialize button started a stopwatch on a display on the machine, which stopped at the time that the lever moved. However, in reality the experimenter triggered the lever movement with a small remote control and in a way that could not be noticed by the subject. The working mode of the machine was shown and explained in detail to the participants before the experiment and handling of the machine (i.e., pressing the initialize button on the machine) was practiced during training. Debriefing of participants showed that none of them had any suspicion about how the machine worked. The machine has been described in detail before (27).

After the machine pressed the start button, which as before also made the signal LED flash, the target LED would flash after an interval of either 500 or 700 ms

(depending on the current experimental block). As before, participants tried to match their press on the target button to the target LED flash. The procedure in NFT was the same as in the self condition.

Baseline Condition

In order to start the trial participants had to press the initialize button with their left hand (SI Appendix, Fig. S1c). After a variable delay between 2500 to 2750 ms the *signal LED* flashed, which signaled the start of the temporal interval until the flash of the target LED (500 or 700 ms, depending on the current block) that subjects had to predict. At the same time when the signal LED flashed a 'click sound' was played through the computer speakers. This sound was recorded from the machine's action in the machine condition, and we used this auditory cue in order to make the baseline condition more similar to the machine and the self condition because in both of the latter conditions a button was pressed, producing a gentle click sound. In addition, we tried to minimize any auditory influence using noise-isolating earmuffs, as was noted above. After participants pressed the target button with their right hand in order to report the estimated onset of the target LED flash, a new trial could be initiated by pressing the initialize button on the left. The procedure in NFT was the same as in the self condition.

Reaction Time Task

In order to control for differences in manual reaction time between our subject groups, we included a reaction time experiment before the main experiment. This additional experimental block consisted of 30 trials in which participants had to start a trial by pressing the leftmost button, just as in the main experiment, and after a random interval between 1500 to 3530 ms, the target LED would flash. The task for participants was to press the target button as fast as possible in response to the target LED flash. Due to technical problems, one patient (patient #5) and one control (control #20) could not complete the reaction time experiment.

Questionnaire

After the experiment, we asked participants the same four questions for all the three experimental conditions. The scope of the questionnaire was to test for

differences in subjects' explicit interpretation of the causal and intentional relationships inherent to our experiment (question 1). Furthermore, we wanted to control for possible confounds, namely that patients with Schizophrenia might feel differently about who "wants" the target LED to flash (e.g. patients might feel control over the LED; question 2) or about the machine and its working mechanism (e.g. patients might feel controlled by or in control of the machine; questions 3 and 4). The questionnaire consisted of the following four questions.

1. Who or what determined in your opinion the time of the target LED flash?
2. Who or what wanted in your opinion the target LED to flash? Whose intention was it?
3. Who or what in your opinion controlled the machine?
4. Who or what was in your opinion controlled by the machine?

Answers to choose from were:

1. I
2. The Experimenter
3. The Computer
4. The Machine
5. The Signal LED/the Button Box
6. Nothing/No One

Additionally, participants had to rate how confident they felt with their answer on a Likert scale from 1 (= not sure at all) to 5 (= completely sure). Importantly, we stressed that the questions only apply to the time after (!) the trial was initialized by pressing the respective buttons (on the response box or the machine) so that the button press to initialize a trial would not factor into subjects' answers. Debriefing after the experiment revealed that the vast majority of participants had difficulties with answering question 2. For that reason, we discarded question 2 and did not analyze it further. As described before, questions 3 and 4 solely controlled for delusions of patients concerning the machine. We did not find any indication that such delusions were present (no single patient reported being controlled by the machine or feeling in control over the machine).

Analyses

All analyses were performed using MATLAB R2014a (The MathWorks, Inc.). To analyze participants' responses in the temporal binding paradigm we discarded the first 15 trials of every block (i.e., the initial 15 feedback trials [FT] that served to familiarize subjects with a given temporal interval) and used the remaining 45 trials, comprised of 30 FT and 15 NFT, for all analyses. We expressed the time of button press (TBP) relative to the actual target LED flash and took the median of every participant's time of button press (TBP). For group results we always depicted the mean of the individual subjects' medians and the standard error of the mean (SEM) as a measure for variability. To estimate participants' reaction times, we took the median of all 30 trials for every participant and averaged across these individual estimates for depicting our group results (i.e., mean+/-SEM).

In order to analyze the responses of the questionnaire we calculated a combined response index reflecting the confidence with which participants rated their answers as well as the share of the specific answer itself. More precisely, the index was calculated by multiplying the share of the response (e.g. 70% answer "I") with the average corresponding confidence (e.g. 3). Accordingly, this index can range from 0 (0% "I") to 5 (100% 'I' times a confidence level of 5).

Statistics

Statistical analyses were performed using Matlab R2014a (The MathWorks, Inc.) and SPSS 24 (IBM). Participants' performance in the temporal binding paradigm was analyzed by means of ANOVAs. We tested for sphericity (Mauchly's test) and adjusted the F statistic using Huynh-Feldt-correction when the assumption of sphericity was not met (mentioned in main text). We furthermore confirmed the assumption of normality by Shapiro-Wilk tests ($p>0.01$; no correction for multiple comparisons). Effect sizes are expressed as (partial) η^2. To control for multiple comparisons, we applied Bonferroni correction when necessary (see main text).

To assess the evidence in our data in favor of the hypothesis that temporal binding was present vs. the evidence in favor of the alternative hypothesis that there was no temporal binding, we calculated Bayes factors within each subject group. Bayesian statistics can confirm whether (or not) in those instances where we reported a non-significant orthodox statistical test on temporal binding, there was substantial evidence in favor of the null-hypothesis. This was important, as it

documents that we did not miss any relevant differences simply due to a lack of statistical sensitivity. Based on previous reports using comparable (450 ms) delays and the Libet clock procedure (5) as well as procedures directly comparable to ours (27), we expected temporal binding of about 33 ms on average. Following the recommendations by Dienes (56), we modelled the prediction for the hypothesis that there is temporal binding as a two-tailed normal distribution with a mean of 33 ms and a standard deviation of half the mean. The Bayes factor BF was then calculated using the routines provided by the same author. Note that a Bayes factors of more than 1 provides evidence for the hypothesis over the alternative hypothesis whereas factors of less than 1 favor the alternative hypothesis. Following the nomenclature proposed by Jeffreys (57), we considered Bayes factors above 3 and below 1/3 as substantial evidence. Finally, we also quantified evidence in favor of the hypothesis for group differences in temporal binding vs. the evidence in favor of the absence of such differences (alternative hypothesis). Respective Bayes factors were calculated using the bayesFactor toolbox for Matlab by Krekelberg (Bart Krekelberg (2023). BayesFactor (https://github.com/klabhub/bayesFactor), GitHub. Retrieved February 2, 2023).

Correlations between temporal binding and Psychopathology

To correlate the amount of patients' temporal binding with individual psychopathology we used Kendal rank correlation coefficients because of non-normally distributed residuals in linear correlation analyses. Statistics on the Pearson correlation coefficients of these prior regression analyses led, however, to the same qualitative results. To control for multiple comparisons, we applied Bonferroni correction when necessary (see main text).

Agency Questionnaire

We checked for the expected response pattern in controls by first using Friedman tests to establish statistically significant differences in each of the three responses of interest (I, computer, machine) between conditions, as our response indices were not normally distributed. Wilcoxon signed-rank tests were applied for post hoc analyses (see SI Appendix, Table S1). For comparison between groups, we

used Mann-Whitney U tests. As an effect size measure, we calculated η^2 using the following formula: $\eta^2 = \frac{z^2}{N-1}$

To control for multiple comparisons, we applied Bonferroni corrections (see main text).

Reaction Time

The distribution of reaction times was non-normally distributed. Therefore, we used Mann-Whitney U tests to compare reaction times between groups.

Code availability

In this study we utilized standard software and published analytical routines, as are specified in detail in our methods section. Related Matlab and PsyScope codes are available from the corresponding author upon reasonable request.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgment

This work was supported by a grant from the German Research Council (DFG CIN to HYW and AL).

References

1. P. Haggard, Sense of agency in the human brain. *Nat. Rev. Neurosci.* **18**, 197–208 (2017).
2. C. Stetson, X. Cui, P. R. Montague, D. M. Eagleman, Motor-Sensory Recalibration Leads to an Illusory Reversal of Action and Sensation. *Neuron* **51**, 651–659 (2006).
3. D. T. Leube, *et al.*, The neural correlates of perceiving one's own movements. *Neuroimage* (2003) https:/doi.org/10.1016/j.neuroimage.2003.07.033.
4. D. Hume, *A treatise of human nature*, L. A. Selby-Bigge, Ed. (Clarendon Press, 1896).

5. P. Haggard, S. Clark, J. Kalogeras, Voluntary action and conscious awareness. *Nat. Neurosci.* **5**, 382–385 (2002).
6. S. J. Blakemore, C. D. Frith, D. M. Wolpert, Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cogn. Neurosci.* **11**, 551–9 (1999).
7. R. L. Seilheimer, A. Rosenberg, D. E. Angelaki, Models and processes of multisensory cue combination. *Curr. Opin. Neurobiol.* (2014) https:/doi.org/10.1016/j.conb.2013.11.008.
8. M. J. Buehner, G. R. Humphreys, Causal contraction: Spatial binding in the perception of collision events. *Psychol. Sci.* **21**, 44–48 (2010).
9. D. M. Eagleman, A. O. Holcombe, Causation and the perception of time. *TRENDS Cogn. Sci.* **6**, 323–325 (2002).
10. C. Hoerl, *et al.*, Temporal Binding, Causation, and Agency: Developing a New Theoretical Framework. *Cogn. Sci.* (2020) https:/doi.org/10.1111/cogs.12843.
11. M. O. Ernst, M. S. Banks, Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
12. N. Wolpe, P. Haggard, H. R. Siebner, J. B. Rowe, Cue integration and the perception of action in intentional binding. *Exp. Brain Res.* (2013) https:/doi.org/10.1007/s00221-013-3419-2.
13. R. Legaspi, T. Toyoizumi, A Bayesian psychophysics model of sense of agency. *Nat. Commun.* (2019) https:/doi.org/10.1038/s41467-019-12170-0.
14. L. Shams, U. Beierholm, Bayesian causal inference: A unifying neuroscience theory. *Neurosci. Biobehav. Rev.* **137**, 104619 (2022).
15. A. Desantis, C. Roussel, F. Waszak, On the influence of causal beliefs on the feeling of agency. *Conscious. Cogn.* (2011) https:/doi.org/10.1016/j.concog.2011.02.012.
16. D. M. Frith, Chris; Blakemore, Sarah-Jayne; Wolpert, Explaining the symptoms of schizophrenia: Abnormal awareness of action. *Brain Res. Reveiws*, 357–363 (2000).
17. C. D. Frith, S.-J. Blakemore, D. M. Wolpert, Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* **355**, 1771

LP – 1788 (2000).

18. S. J. Blakemore, D. Wolpert, C. Frith, Why can't you tickle yourself? *Neuroreport* **11**, R11-6 (2000).

19. S. S. Shergill, G. Samson, P. M. Bays, C. D. Frith, D. M. Wolpert, Evidence for sensory prediction deficits in schizophrenia. *Am. J. Psychiatry* **162**, 2384–2386 (2005).

20. N. Franck, *et al.*, Defective recognition of one's own actions in patients with schizophrenia. *Am. J. Psychiatry* **158**, 454–459 (2001).

21. R. Knowles, S. McCarthy-Jones, G. Rowse, Grandiose delusions: A review and theoretical integration of cognitive and affective perspectives. *Clin. Psychol. Rev.* **31**, 684–696 (2011).

22. P. Haggard, F. Martin, M. Taylor-Clarke, M. Jeannerod, N. Franck, Awareness of action in schizophrenia. *Neuroreport* **14**, 1081–1085 (2003).

23. M. Voss, *et al.*, Altered awareness of action in schizophrenia: A specific deficit in predicting action consequences. *Brain* **133**, 3104–3112 (2010).

24. M. Hauser, *et al.*, Altered sense of agency in schizophrenia and the putative psychotic prodrome. *Psychiatry Res.* (2011) https:/doi.org/10.1016/j.psychres.2010.08.003.

25. K. T. Graham-Schmidt, M. T. Martin-Iverson, N. P. Holmes, F. A. V. Waters, When one's sense of agency goes wrong: Absent modulation of time perception by voluntary actions and reduction of perceived length of intervals in passivity symptoms in schizophrenia. *Conscious. Cogn.* **45**, 9–23 (2016).

26. M. J. Buehner, G. R. Humphreys, Causal binding of actions to their effects: Research article. *Psychol. Sci.* **20**, 1221–1228 (2009).

27. M. J. Buehner, Understanding the Past, Predicting the Future: Causation, Not Intentional Action, Is the Root of Temporal Binding. *Psychol. Sci.* **23**, 1490–1497 (2012).

28. M. J. Buehner, Awareness of voluntary and involuntary causal actions and their outcomes. *Psychol. Conscious. Theory, Res. Pract.* **2**, 237–252 (2015).

29. M. Synofzik, P. Thier, D. T. Leube, P. Schlotterbeck, A. Lindner, Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain* **133**,

977 262–271 (2010).
978 30. N. C. Andreasen, The Scale for the Assessment of Negative Symptoms
979 (SANS): Conceptual and Theoretical Foundations. *Br. J. Psychiatry* **155**, 49–
980 52 (1989).
981 31. S. R. Kay, A. Fiszbein, L. A. Opler, The positive and negative syndrome scale
982 (PANSS) for schizophrenia. *Schizophr. Bull.* **13** (1987).
983 32. A. R. Krugwasser, Y. Stern, N. Faivre, E. V. Harel, R. Salomon, Impaired sense
984 of agency and associated confidence in psychosis. *Schizophrenia* **8** (2022).
985 33. S. Nieuwenhuis, B. U. Forstmann, E. J. Wagenmakers, Erroneous analyses of
986 interactions in neuroscience: A problem of significance. *Nat. Neurosci.*
987 (2011) https:/doi.org/10.1038/nn.2886.
988 34. J. M. Ford, D. H. Mathalon, Electrophysiological evidence of corollary
989 discharge dysfunction in schizophrenia during talking and thinking in
990 *Journal of Psychiatric Research*, (2004) https:/doi.org/10.1016/S0022-
991 3956(03)00095-5.
992 35. N. Andreasen, The Scale for the Assessment of Positive Symptoms (SAPS).
993 *Univ. Iowa*, 1–2 (1984).
994 36. M. Hauser, *et al.*, Sense of agency is altered in patients with a putative
995 psychotic prodrome. *Schizophr. Res.* (2011)
996 https:/doi.org/10.1016/j.schres.2010.10.031.
997 37. T. Haarmeier, F. Bunjes, A. Lindner, E. Berret, P. Thier, Optimizing visual
998 motion perception during eye movements. *Neuron* **32**, 527–535 (2001).
999 38. D. M. Wolpert, Z. Ghahramani, Computational principles of movement
1000 neuroscience. *Nat. Neurosci.* **3**, 1212–1217 (2000).
1001 39. R. C. Miall, W. D. M., Forward Moldels for Physiological Motor Control.
1002 *Neural Networks* **9**, 1265–1279 (1996).
1003 40. D. M. M. Wolpert, Z. Ghahramani, M. I. I. Jordan, An internal model for
1004 sensorimotor integration. *Science (80-.).* **269**, 1880–1882 (1995).
1005 41. S. J. Blakemore, D. M. Wolpert, C. D. Frith, Central cancellation of self-
1006 produced tickle sensation. *Nat. Neurosci.* **1**, 635–640 (1998).
1007 42. L. Shams, U. R. Beierholm, Causal inference in perception. *Trends Cogn. Sci.*
1008 **14**, 425–432 (2010).
1009 43. K. Friston, What is optimal about motor control? *Neuron* (2011)

https:/doi.org/10.1016/j.neuron.2011.10.018.

44. A. Lindner, P. Thier, T. T. J. Kircher, T. Haarmeier, D. T. Leube, Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. *Curr. Biol.* **15**, 1119–1124 (2005).

45. S. Lorimer, *et al.*, The developmental profile of temporal binding: From childhood to adulthood. *Q. J. Exp. Psychol.* (2020) https:/doi.org/10.1177/1747021820925075.

46. S. Moritz, T. S. Woodward, Jumping to conclusions in delusional and non-delusional schizophrenic patients. *Br. J. Clin. Psychol.* **44**, 193–207 (2005).

47. S. J. Blakemore, J. Smith, R. Steel, E. C. Johnstone, C. D. Frith, The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: Evidence for a breakdown in self-monitoring. *Psychol. Med.* **30**, 1131–1139 (2000).

48. M. Synofzik, G. Vosgerau, A. Lindner, Me or not me - An optimal integration of agency cues? *Conscious. Cogn.* **18**, 1065–1068 (2009).

49. J. Metcalfe, *et al.*, Judgements of agency in schizophrenia : an impairment in autonoetic metacognition Judgements of agency in schizophrenia : an impairment in autonoetic metacognition. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* **367**, 1391–1400 (2012).

50. M. J. Roth, M. Synofzik, A. Lindner, The cerebellum optimizes perceptual predictions about external sensory events. *Curr. Biol.* **23**, 930–5 (2013).

51. P. C. Fletcher, C. D. Frith, Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* **10**, 48–58 (2009).

52. P. Sterzer, *et al.*, The Predictive Coding Account of Psychosis. *Biol. Psychiatry* **84**, 634–643 (2018).

53. R. A. Adams, K. E. Stephan, H. R. Brown, C. D. Frith, K. J. Friston, The computational anatomy of psychosis. *Front. Psychiatry* (2013) https:/doi.org/10.3389/fpsyt.2013.00047.

54. C. E. Notredame, D. Pins, S. Deneve, R. Jardri, What visual illusions teach us about schizophrenia. *Front. Integr. Neurosci.* (2014) https:/doi.org/10.3389/fnint.2014.00063.

55. J. Cohen, B. MacWhinney, M. Flatt, J. Provost, PsyScope: An interactive

graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behav. Res. Methods, Instruments, Comput.* **25**, 257–271 (1993).

56. Z. Dienes, Bayesian versus orthodox statistics: Which side are you on? *Perspect. Psychol. Sci.* (2011) https:/doi.org/10.1177/1745691611406920.

57. H. Jeffreys, *The Theory of Probability*, 3rd Ed. (Oxford University Press, 1940).

Figures

Figure 1

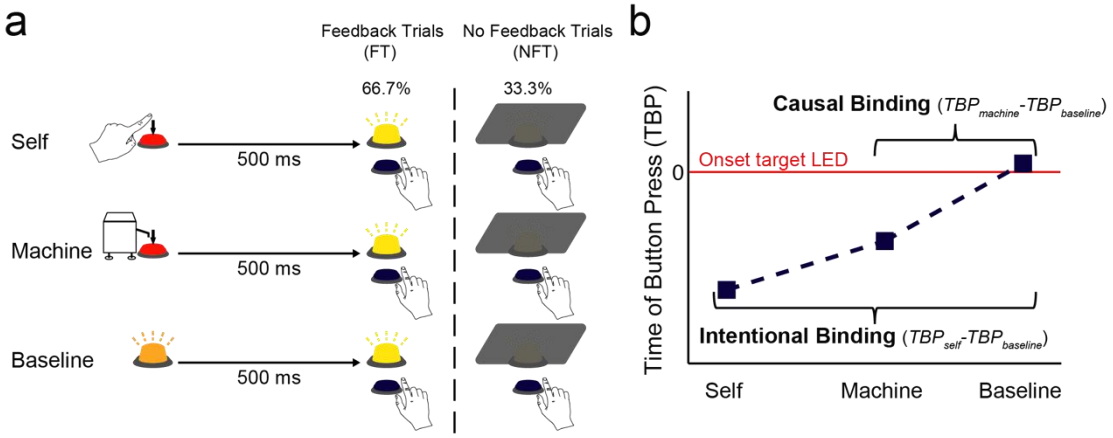

Fig 1 | Experimental paradigm and predicted results in healthy individuals. a, Participants had to predict the time of a target-LED flash (depicted in yellow) and were instructed to press a target button (blue) simultaneously with the flash. Depending on the experimental condition, the target-LED flash was caused by the participant's press on the start button (self condition), caused by a machine's press on the start button (machine condition), or it was only associated with a signal-LED flash (baseline condition). In all cases, the initial event preceded the target LED flash by a fixed temporal interval of 500ms. In two thirds of trials participants could see the target LED flash (feedback trials, FT) while in the remaining trials view of the target-LED was explicitly blocked by the experimenter before a trial started (no feedback trials, NFT). These NFTs prevented subjects from responding merely in a reactive fashion to the target-LED flash and forced them to engage in a predictive response strategy. **b**, Based on previous research, we expected temporal binding and, therefore, earlier times of button press (TBP) in the self condition compared to baseline. We refer to this difference as intentional binding. TBPs in the machine condition should also occur earlier compared to baseline, an effect we here denote as causal binding. Given the lack of an active intentional component in the machine condition, we predict TBPs in the machine condition to be closer to baseline than those in the self condition. For further details please refer to the Discussion.

Figure 2

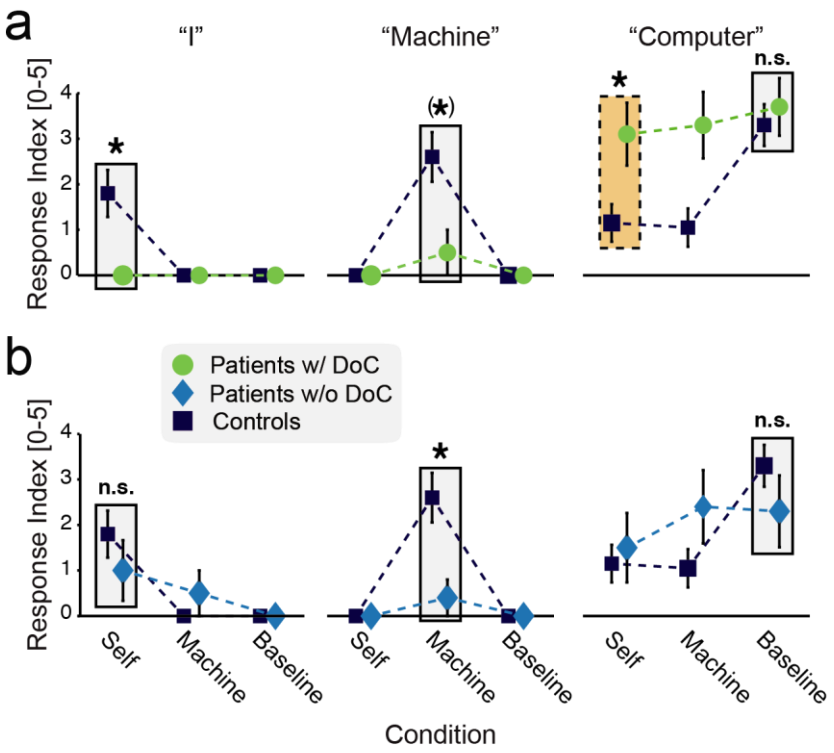

Fig. 2 | Explicit judgments of agency/causality in response to the question "Who or what determined the time of the target LED flash?" (mean ± SEM). a, Controls showed the expected pattern of results with the response index for "I" being highest in the self-condition, the index for "the machine" being highest in the machine-condition, and the index for "the computer" being highest in the baseline condition. Patients with schizophrenia suffering from DoC did not feel responsible for causing the target-LED to flash. As compared to controls this led to a significantly smaller response index for the answer "I" in the self condition (which was actually zero). Instead, they attributed significantly stronger levels of agency to the computer in this condition. In the machine condition, patients with DoC attributed significantly less agency to the machine than controls. **b**, Patients without DoC had a very similar response pattern as controls, the only significant difference being a lower response index for the machine in the machine condition. Significant group differences are indicated (Mann-Whitney U tests; n.s. = not significant, * $p < 0.05$ corrected, ** $p < 0.01$). Control group: N = 20, patient subgroups: N = 10, each.

Figure 3

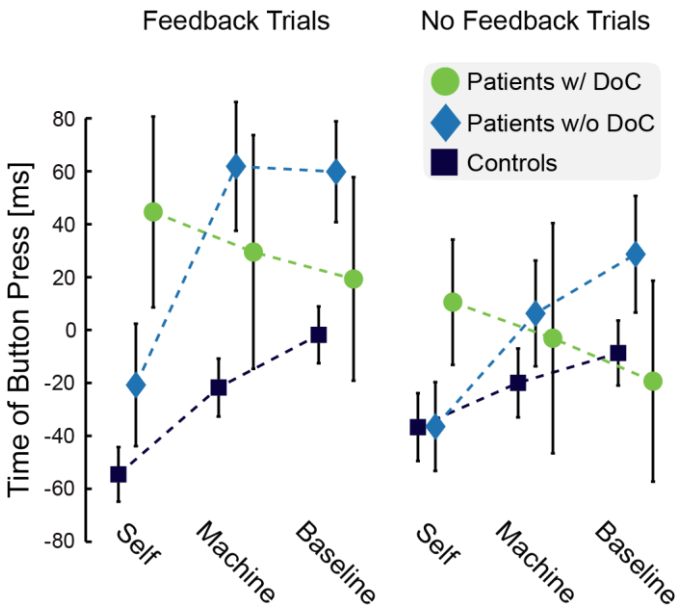

Fig. 3 | Time of button press across conditions and subject groups (interval 500 ms; mean ± SEM). Time of button press in controls (dark blue) and the subgroups of Schizophrenia patients with DoC (green) and without DoC (light blue). Note that patients without DoC showed strong intentional binding. This effect was absent and numerically even reversed (repulsion) in patients with DoC. Both patient subgroups exhibited later TBPs in feedback trials (FT). For details refer to Results. Controls: N = 20, all patients: N = 20, patient subgroups: N = 10, each.

Figure 4

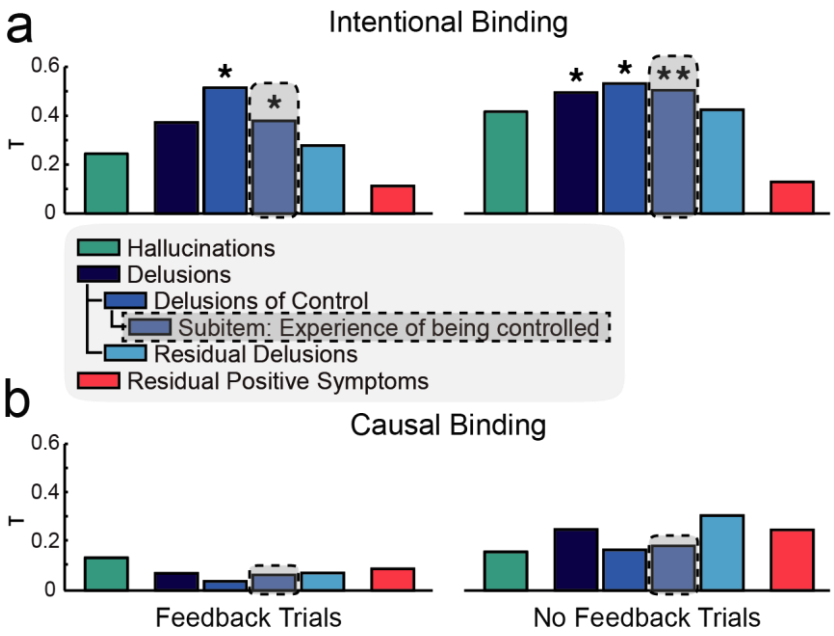

Fig. 4 | Correlation between intentional or causal binding and individual psychopathology. a, Reductions in intentional binding were positively correlated with the strength of patients' DoC (and in one case also with the strength of overall delusions which DoC are a part of). **b**, There was no relationship between the amount of causal binding and DoC or other psychopathologies investigated here. Significant correlations are indicated (Kendall rank correlation, two-tailed; * $p < 0.05$ corrected). N = 20, each. **a and b**, shaded bars represent an exploratory correlation with only question 15 of the SAPS which directly asks for the experience of being controlled (Kendall rank correlation, two-tailed; * $p < 0.05$, ** $p < 0.01$).