CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/158476/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Roth, Manuel J., Lindner, Axel, Hesse, Klaus, Wildgruber, Dirk, Wong, Hong-Yu and Buehner, Marc J. 2023. Impaired perception of temporal contiguity between action and effect is associated with disorders of agency in Schizophrenia. Proceedings of the National Academy of Sciences 120 (21) , e2214327120. 10.1073/pnas.2214327120

Publishers page: https://doi.org/10.1073/pnas.2214327120

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Supplementary Figures

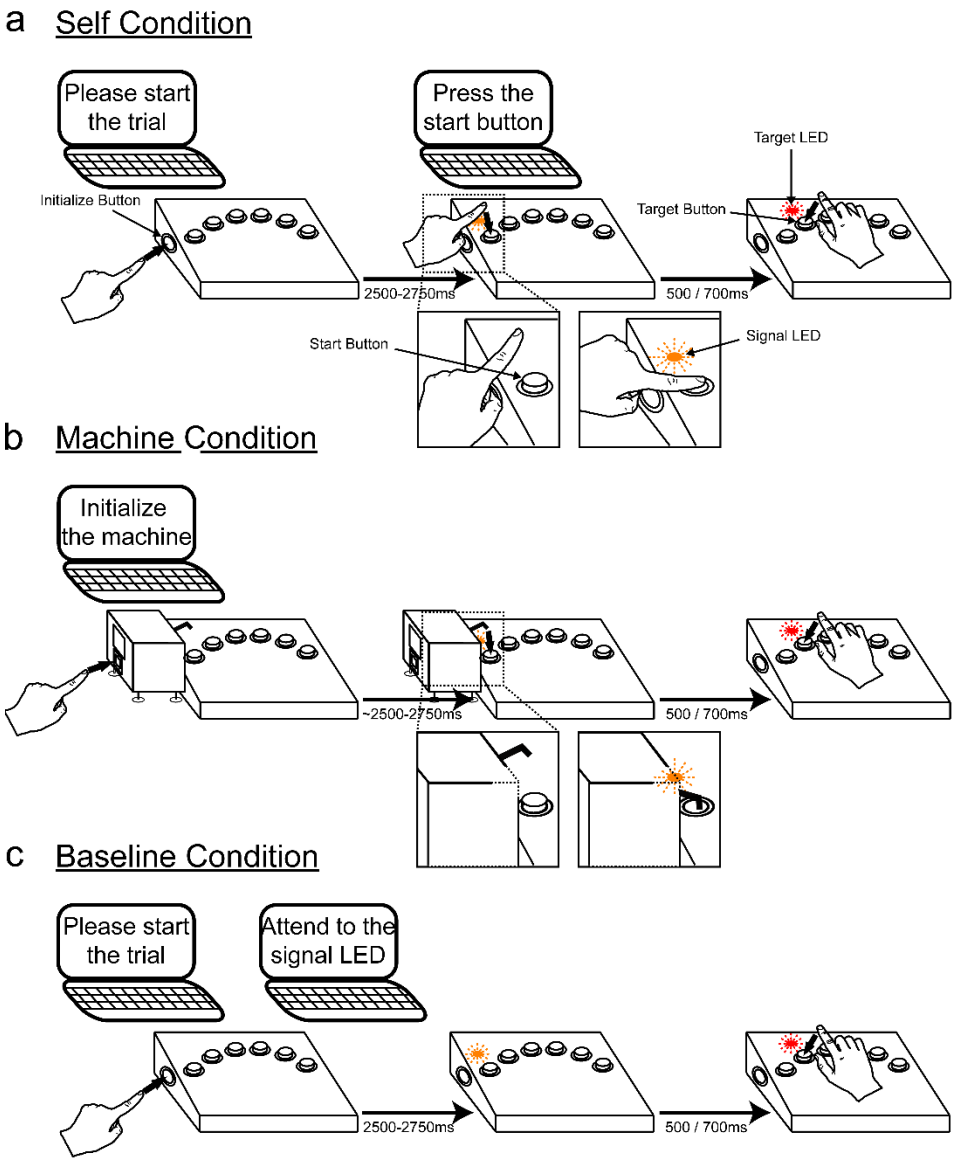

Figure S1 | Depiction of experimental setup and design. In all conditions, the subjects task was to predict the time of the target LED flash by pressing the target button at the respective time. **a**, Self condition: Participants started a trial by pressing the initialize button. After a variable delay between 2500 and 2750 ms, the computer informed them that they should press the start button located below the signal LED. Pressing the start button also produced a simultaneous flash of the signal LED (see detail magnification). The button press (along with the signal LED flash as was likewise present in all other experimental conditions) signaled the start of the target interval (500 or 700 ms, respectively, depending on the current experimental block). Next, participants were supposed to press the target button, located below the target LED, at the time they expected the target LED flash to occur (remember, the target flash immediately followed after the target interval). **b**, Machine condition: Participants started a trial by pressing the initialize button on the machine. After a variable delay, which was roughly in the same range as in the self condition, the machine's lever would press the start button. Pressing the start button triggered a signal LED flash (see detail magnification) and indicated the beginning of the 500 ms/700 ms target interval. Participants next pressed the target button below the target LED when they expected the target LED flash to occur. **c**, Baseline condition: Participants started a trial by pressing the initialize button as in all other conditions. After a variable delay between 2500 and 2750 ms the signal LED would flash. This flash signaled the start of the target interval. As before, participants pressed the target button at the time they expected the target LED flash to occur. Note that we took particular care to make conditions as similar as possible.

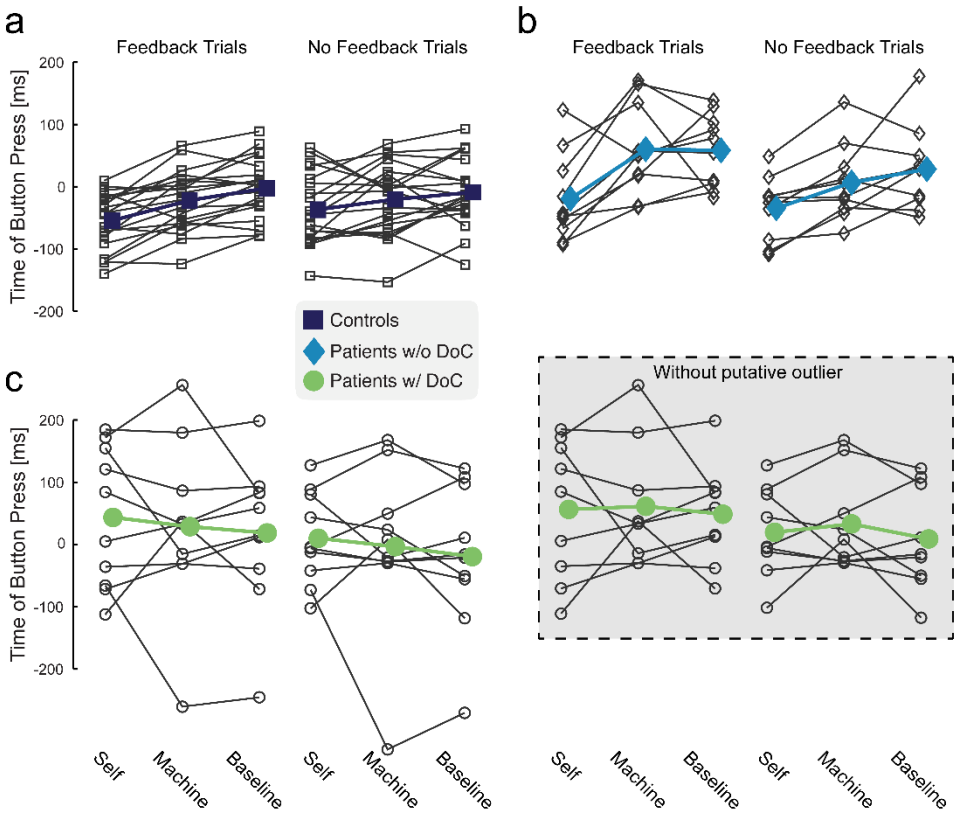

Figure S2 | Time of button press of individual participants across conditions (interval 500 ms). a, Controls (N=20). **b**, Patients without DoC (N=10). **c**, Patients with DoC (N=10). Exclusion of one putative outlier in the DoC group does not recover intentional binding (see dashed panel in **c**). Filled color symbols indicate the respective group means.

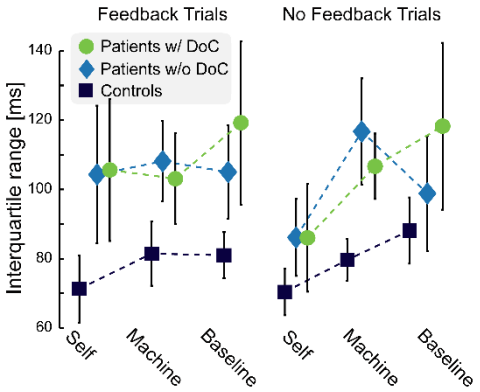

Figure S3 | Response variability (interval 500 ms; mean ± SEM). To analyze the variability in participants' responses we estimated the interquartile range (IQR) for every participant. Both patient subgroups showed much larger response variability compared to controls. For statistical analysis we used log-transformed data because of otherwise non-normally distributed residuals. A 3 x 3 x 2 ANOVA with the between subject factor group (controls, DoC, NoDoC), and within subject factors condition (self, machine, baseline), and feedback (FT, NFT) revealed a main effect of condition ($F_{2,74} = 3.4$, $p = .038$) and a significant group difference ($F_{2,37} = 3.8$, $p = .032$) between controls and patients. The larger variability in patients with schizophrenia is in accordance with previous studies on patients with schizophrenia (for review see (1)). However, our subject groups were equally precise in terms of response accuracy (see Baseline estimates in Figure 2 and compare Results), which is decisive for the reliable estimation of temporal binding effects. Controls: N = 20, patient subgroups: N = 10, each.

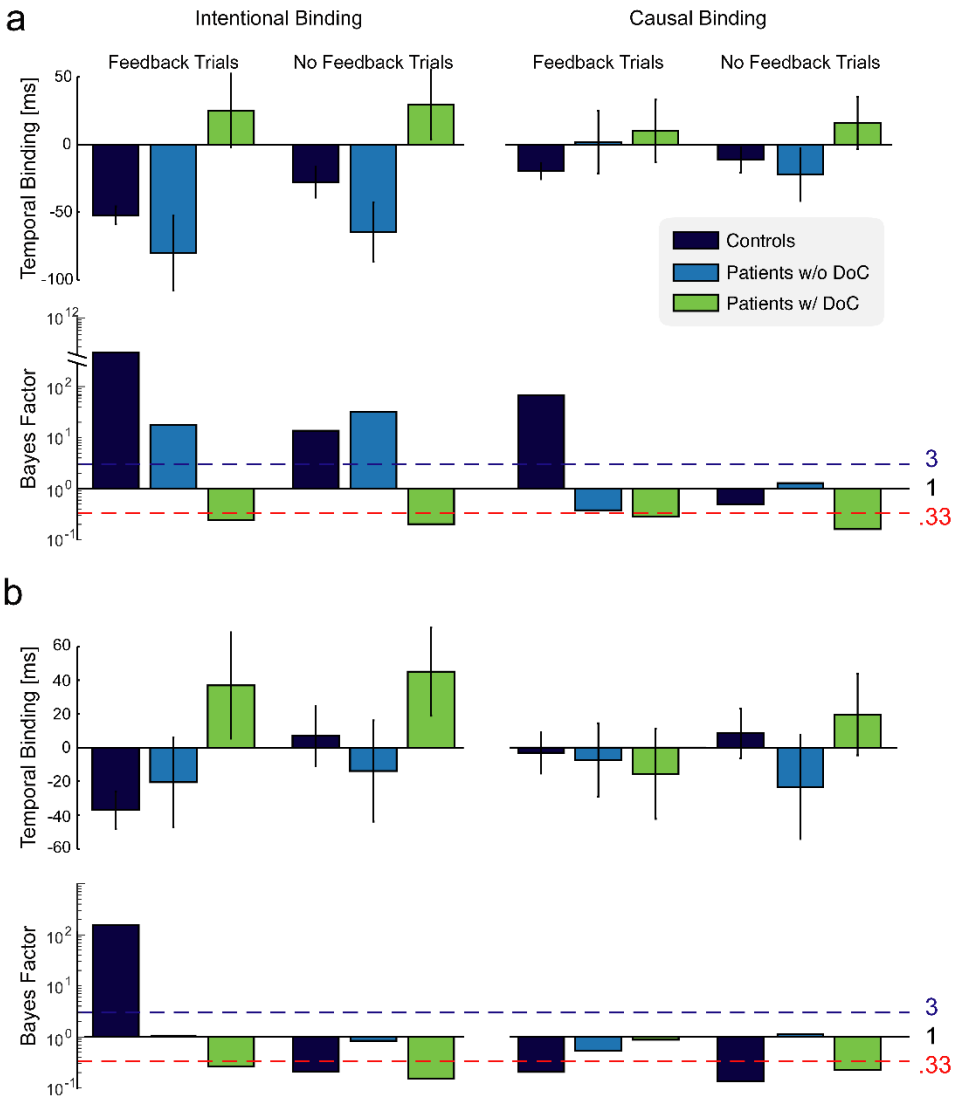

Figure S4 | Amount of intentional- and causal binding (means ± SEM) and their respective Bayes Factors. a, Binding in the 500 ms interval. DoC-patients showed a strong numerical difference in intentional binding compared to both controls and NoDoC-patients, namely a temporal repulsion instead of a temporal binding effect. Visible inspection suggests, however, that there were no comparable between-group differences in causal binding. The respective Bayes Factors support this conclusion in that there is evidence for intentional binding in both controls and patients without DoC while there is evidence for no intentional binding in patients with DoC. There is evidence for causal binding only in the controls in FT and evidence for no causal binding in patients with DoC. Note that Bayes factors above 3 (blue dashed line) and below 0.33 (red dashed line) provide substantial evidence in favor and against temporal binding, respectively, while Bayes factors in between are considered undecisive. **b**, Intentional and causal binding in the 700 ms interval and the corresponding Bayes Factors. Controls: N = 20, patient subgroups: N = 10 each.

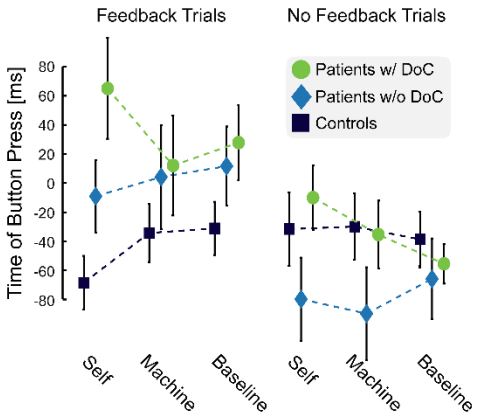

Figure S5 | Time of button press across conditions and subject groups (interval 700ms; mean ± SEM). **a**, Time of button press in controls (dark blue) and the overall group of schizophrenia patients (red). The overall pattern of results mimics the one for the 500 ms interval apart from the fact that intentional binding in healthy controls was less pronounced in FT and even absent in NFT (compare Fig. 2a). TBPs were on average too early relative to the actual time of the target LED flash. Overall, we did not exhibit strong temporal binding for the longer target interval (as was to be expected (2, 3)). **b**, Time of button press in controls (dark blue) and the subgroups of Schizophrenia patients with DoC (green) and without DoC (light blue). Both patient subgroups showed a clear difference in performance in relation to the availability of feedback (FT, NFT), namely later button presses in FT and therefore a more reactive response pattern. Visual inspection suggests that the results are quite similar to those for the 500ms interval (compare Figure 2b) and indicates intentional binding in patients without DoC but no intentional binding (or rather a repulsion) in patients with DoC. Controls: N = 20, all patients: N = 20, patient subgroups: N = 10 each.

Supplementary Tables

Table S1 | Comparison of judgments of agency in healthy subjects between conditions. To check whether our control participants showed the expected bias in explicit agency judgments in the various experimental conditions (i.e., highest response index for "I" in the self-condition, highest response index for "the machine" in the machine-condition, and highest response index for "the computer" in the baseline condition), we first compared the respective indices for each answer in the three conditions using Friedman tests (one per response). All three tests revealed statistically significant differences in the response index between conditions (response "I": $\chi^2(2) = 16.00$, $p < .000$; response "machine": $\chi^2(2) = 22.00$, $p < .000$; response "computer": $\chi^2(2) = 19.47$, $p < .000$). Three post hoc analyses with Wilcoxon signed-rank tests per response were conducted with a Bonferroni correction applied ($\alpha = 0.0167$). All tests revealed significant results supporting our hypothesis about control participants' assumptions about intentionality and causality in our experimental conditions.

Response:	**I**			**Machine**			**Computer**		
Comparison	Self vs. Machine	Self vs. Baseline	Machine vs. Baseline	Machine vs. Self	Machine vs. Baseline	Self vs. Baseline	Baseline vs. Self	Baseline vs. Machine	Self vs. Machine
Z	-2.588	-2.588	0	-3.071	-3.071	0	-3.022	-2.913	-0.447
p	0.01	0.01	1	0.002	0.002	1	0.003	0.004	0.655

Table S2 | Comparison of the confidence ratings in the agency questionnaire between groups. To check whether differences in confidence between participant groups exist we compared the confidence values using Kruskal-Wallis Test. The results are shown in the table below. There were no significant differences across the control group and both patient subgroups in any condition.

Condition	**Self**	**Machine**	**Baseline**
H(2)	3.610	1.705	1.434
p	0.164	0.426	0.488

Table S3 | Bayes-Factors expressing evidence for vs. against differences in temporal binding between groups. In order to support the conclusions of our main group analyses, we determined Bayes-factors capturing evidence for (BF > 1) vs. against (BF < 1; compare Methods for the interpretation of Bayes-factor values) group differences in intentional and causal binding, respectively. These results corroborate our other group analyses in that there is substantial to very strong evidence for group differences in intentional binding between controls and patients with DoC and for differences between patients with and without DoC, while there is substantial evidence for no difference between controls and patients without DoC. Results concerning causal binding are often undecisive and only indicate substantial evidence for a difference between controls and patients with DoC in NFT.

		Controls vs. Patients w/ DoC	**Controls vs. Patients w/o DoC**	**Patients w/ DoC vs. Patients w/o DoC**
Intentional Binding	FT	55.95	0.138	7.75
	NFT	5.088	0.106	8.787
Causal Binding	FT	1.813	1.071	0.484
	NFT	1.371	0.233	1.399

Table S4 | Characteristics of patients. Subgroup: 1 = DoC, 2 = NoDoC. m = male; f = female. Years of education = primary + secondary school. SAPS/SANS score = Scale for the Assessment of Positive/Negative Symptoms (4, 5). PANSS = Positive and Negative Syndrome Scale (6). Medication: Calculation of Olanzapine equivalents using the defined daily doses (DDDs) of the World Health Organization (https://www.whocc.no/).

Patient #	Subgroup	Sex	Age (years)	Education (years)	SAPS: Hallucinations	Delusions	Bizarre Behavior	Positive Formal Thought Disorder	Composite Score	SANS: Composite Score	PANSS: Positive Scale	Negative Scale	General Psychopathology	Medication: Olanzapine Equivalents
3	1	m	43	13	16	26	0	7	49	54	37	36	67	5
5	1	f	34	13	1	29	0	0	30	41	16	32	40	50
6	1	m	36	10	0	7	0	0	10	6	15	10	36	5.3
7	1	m	35	10	0	13	0	3	18	46	17	27	25	2.5
10	1	f	27	13	0	11	0	0	11	22	21	7	40	1.6
12	1	f	32	13	0	14	1	2	17	3	12	7	24	0.3
13	1	m	29	10	0	5	0	0	5	19	27	20	41	2
14	1	m	20	13	0	22	0	0	22	31	22	22	45	1.5
16	1	m	37	13	19	25	0	16	60	26	15	10	44	0.3
19	1	m	30	13	24	47	3	7	84	13	?	?	?	1.7
Mean			**32**	**12**	**6**	**20**	**0**	**4**	**31**	**26**	**20**	**19**	**40**	**7**
1	2	m	59	13	0	2	2	0	4	35	10	20	35	1
2	2	f	38	10	0	0	0	0	0	20	14	12	25	2.3
4	2	m	45	10	0	3	0	1	4	36	19	26	50	0.8
8	2	m	29	13	0	0	0	0	0	14	14	13	32	1.5
9	2	m	35	10	5	2	0	0	7	5	13	7	24	0.4
11	2	m	34	10	0	12	14	22	51	23	21	13	42	51.5
15	2	m	25	13	0	0	0	0	0	5	7	10	20	0
17	2	f	38	13	0	2	0	2	4	32	?	?	?	22.0
18	2	f	20	13	0	5	0	0	5	23	9	10	20	1.5
20	2	m	30	13	0	3	6	3	12	17	?	?	?	2
Mean			**35**	**12**	**1**	**3**	**2**	**3**	**9**	**21**	**13**	**14**	**31**	**8**
Grand Mean			**34**	**12**	**3**	**11**	**1**	**3**	**20**	**24**	**17**	**17**	**36**	**8**

Table S5 | Characteristics of controls. m = male; f = female. Years of education = primary + secondary school.

Control #	Sex	Age (years)	Education (years)
1	f	21	13
2	f	24	12
3	m	21	11
4	m	24	12
5	m	31	12
6	m	22	9
7	m	25	13
8	f	32	13
9	m	26	13
10	m	63	13
11	f	33	13
12	f	35	13
13	m	35	13
14	m	49	12
15	m	37	10
16	f	43	10
17	m	44	13
18	m	41	11
19	m	30	10
20	m	29	12
mean		**33**	**12**

Supplementary Results

Temporal Binding and Feedback Availability

The 2 x 2 x 3 ANOVA (group [SZ, CTR], feedback [FT, NFT], condition [Self, Machine, Baseline]) described in the main results section yielded a significant feedback x group interaction. To illuminate this finding in more detail we ran two 2 x 3 post-hoc ANOVAs (feedback, condition; Bonferroni corrected $\alpha = 0.025$), one per group, to investigate the effect of interest, namely feedback, in the two groups separately. This revealed a significant main effect of feedback ($F_{1,19} = 17.47$, $p = 0.001$) in the patient, but not in the control group ($F_{1,19} = 0.77$, $p = 0.389$) proving the influence of feedback availability on patients' performance.

Supplementary Discussion

In the discussion section of our main paper we state that increased temporal binding in the self- vs. the machine condition, as was observed in healthy control subjects, could be easily explained by the availability of additional (causal) cues in the self vs. the machine condition such as proprioception and forward models. We further state that a significant reduction of variability in subjects' time estimates from the baseline over the machine to the self-condition would be compatible with this interpretation. We here expand these considerations in greater detail. While the final response is identical across all experimental tasks, the available cues around the beginning of the 500/700 ms delay vary greatly. While in the baseline condition the start of this delay is only signaled by the start-signal LED going off, additional cues are available in the machine and in the self-condition. Specifically, in the machine condition additional (visual) sensory cues related to the machine pressing the button come into play. In the self-condition even further efferent and afferent information sources related to the subjects' finger movements are in principle available (proprioception, efference copies, somatosensory input, etc.). Following the logic of optimal cue integration, the perception of the time-point of the start of the delay should be the more reliable, the more informative cues can be integrated. One indirect consequence of this is that subjects' responses should be most accurate in the self-condition, followed by the machine condition, and the baseline condition. This is because across these latter conditions the number of informative cues about the start signal and thus its reliability increases from the latter to the earlier. This is exactly what we found in all subject groups (compare Fig. S3). A second consequence would be that if temporal information about the signal LED were combined with temporal information about the target LED flash (to assess a more robust estimate of the latter due to cue combination), then the target LED flash should be perceived the earlier, the more reliable the start signal (also compare (7) and our introduction). In fact, in healthy subjects this is exactly what we found: the target LED flash was perceived earliest in the self-condition, followed by the machine condition and then the baseline condition. But despite the fact that the reliability of subjects' responses across conditions was indistinguishable between our subject groups, the patterns of temporal binding in patients clearly differed from those of controls. Hence, optimal cue integration perhaps contributes to temporal binding but cannot explain the full picture of our results. The crucial addition that allows

qualitatively explaining our results is captured by the so-called Bayesian causal inference models of cue integration ((8, 9); also compare our introduction in the main manuscript). The additional step that is introduced by these models is (in the context of our study) the estimation of the likelihood that the start signal and the target LED are causally related (since only then the optimal integration of the two cues is meaningful; also remember the counter-example that integrating thunder and lightning with one's key press typically makes no sense). In fact, this is exactly the processing step that seems distinct between patients and controls and which could account for the presence vs. the absence of intentional binding in controls vs. DoC patients, as the deficit in inferring causal self-agency in DoC patients in the self-condition is associated with a lack/diminished intentional binding (aka optimal cue integration). Crucially, we assume that this deficit in agency attribution is based on imprecise forward models (or unreliable "complex" priors), as we explain in further detail in the discussions section of our paper.

Supplementary References

1. S. Thoenes, D. Oberfeld, Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy. *Clin. Psychol. Rev.* **54**, 44–64 (2017).
2. P. Haggard, S. Clark, J. Kalogeras, Voluntary action and conscious awareness. *Nat. Neurosci.* **5**, 382–385 (2002).
3. 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).
4. N. Andreasen, The Scale for the Assessment of Positive Symptoms (SAPS). *Univ. Iowa*, 1–2 (1984).
5. N. C. Andreasen, The Scale for the Assessment of Negative Symptoms (SANS): Conceptual and Theoretical Foundations. *Br. J. Psychiatry* **155**, 49–52 (1989).
6. S. R. Kay, A. Fiszbein, L. A. Opler, The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13** (1987).
7. 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.
8. L. Shams, U. Beierholm, Bayesian causal inference: A unifying neuroscience theory. *Neurosci. Biobehav. Rev.* **137**, 104619 (2022).
9. L. Shams, U. R. Beierholm, Causal inference in perception. *Trends Cogn. Sci.* **14**, 425–432 (2010).