iScience, Volume *23*

Supplemental Information

Perception of Our Own Body Influences Self-Concept and Self-Incoherence Impairs Episodic Memory

Pawel Tacikowski, Marieke L. Weijs, and H. Henrik Ehrsson

This document includes the following:

Transparent Methods

Supplementary References

Supplementary Figures S1-S7

Supplementary Tables S1-S5

Captions for Data S1, Data S2, and Video S1

Transparent Methods

Participants: Sixty-six healthy volunteers participated in the study (42 females; 64 right-handed; mean age: 26±5 years). The sample size was chosen based on similar previous studies (Banakou et al., 2013; Peck et al., 2018). We recruited pairs of friends (same-sex) who knew each other for at least 6 months (mean: 3.5 years). All participants had normal or corrected-to-normal vision, normal hearing, and no history of neurological or psychiatric illness. One participant was excluded from the analyses of similarity and memory data because he did not follow the instructions. All participants provided written informed consent, and the study was approved by the Regional Ethical Review Board of Stockholm (since 2019, the Swedish Ethical Review Authority).

Procedure: Pairs of friends participated in the experiment simultaneously. First, we administered the Inclusion of Other in the Self (IOS) scale (Aron et al., 1992) and the short version of the Network of Relationships Inventory (NRI): Behavioral Systems Version (Furman and Buhrmester, 2009); these questionnaires assessed different aspects of friendship and were used in the control analyses (Fig. S5 and Table S5). Next, the participants practiced how to use a numeric keypad without looking at it; this skill was required in the following self-rating task (see further). During this practice, a cardboard box covered the participant's hand and the keypad, and the task was to press a key that corresponded to a number presented on the screen (20 trials). All participants completed this practice without problems. Then, the friend-rating part was conducted during which the participants sat in front of separate computers (no body perception manipulation applied) and rated 120 trait characteristics in relation to the friend (Fig. 1D; for details, see the "Friend- and self-reference tasks" section). In the following self-rating part, the same 120 trait characteristics were randomly assigned to the syncF, asyncF, syncS, and asyncS conditions (30 traits in each) and rated by the participants in relation to the self (Fig. 1D). The two friends did not see each other's responses in either the friend- or the self-rating task. The full-body illusion conditions started with an induction phase (45 s), followed by spoken instructions and the self-rating task. During each condition, when both participants provided 10, 20, and 30 self-ratings, respectively, we simultaneously "threatened" both participants with mock knives and measured skin conductance responses during these events (Fig. 1C and 1E). Each condition lasted ~ 9 min, and the order of conditions was randomized. During breaks between conditions, the participants took off the Head Mounted Displays (HMDs) and filled out the illusion questionnaire (Fig. 1B and 1E). After the four conditions, the participants completed the last memory task while sitting in front of computers again without any body perception manipulation (Fig. 5A). Finally, a short debriefing was performed with each participant separately in which we asked for feedback and assessed naivety ("What result do you think we expect in this study?"; "Have you used any special strategy in any of the tasks, and if so, what was it?"; "Do you have any other comments or feedback?"). No participant guessed the purpose of the study or reported the hypothesized pattern of results.

Full-body illusion paradigm and visuotactile stimulation: The participants laid down on two beds and wore HMDs (Oculus Rift, Melo Park, CA, USA). The participants' necks were supported with pillows, and their heads were tilted forward (~45°), as if the participants were looking directly at their feet. Each set of HMDs was connected to two digital cameras (Grasshopper3, FLIR, Ludwigsburg, Germany) placed parallel to each other (~7 cm apart), directly behind, and above the participants' heads (Fig. 1A). This setup allowed us to present true stereoscopic, high-quality videos of the participant's own body (syncS, asyncS) or the friend's body (syncF, asyncF) recorded from a first-person perspective. During the synchronous conditions, recordings were displayed with a negligible delay (setup's intrinsic delay: <100 ms). In contrast, a 3 s delay was introduced in the asynchronous conditions. In each condition, the participants received the same number of touches on three body locations (upper legs and lower abdomen; ~13 touches per minute). Strokes were applied with white Styrofoam balls (10 cm diameter) attached to the end of thin rods. The order of touches was pseudorandom (not more than 2 consecutive touches to the same body part). The duration of each stroke was 1 s, and the interval between subsequent strokes was 2, 3, or 4 s. Each touch covered ~25 cm of the participant's body. To ensure synchrony between the two experimenters, they practiced the procedure beforehand, and during the study, they both listened to the same audio cues indicating the onset, duration, and location of each stroke. The participants did not hear these cues because they were played through the experimenters' headphones. The participants performed the self-rating task while receiving visuotactile stimulation. The stroking sequence was paused after 10, 20, and 30 completed ratings; during these pauses, the experimenters simultaneously "attacked" both participants with mock knives (see "Skin conductance responses" below). We instructed the participants to relax and move as little as possible during each condition. The participant's right hand was covered with a cardboard box to eliminate visual feedback from finger movements during the self-rating task.

Illusion questionnaire: After each condition, a questionnaire was administered to quantify the strength of the full-body illusion. Illusion and control statements were adapted from (Petkova and Ehrsson, 2008), and the participants indicated how much they agreed or disagreed with each statement (Fig. 1B; -3 "strongly disagree", +3 "strongly agree"). The illusion items concerned body ownership (I1) and referral of touch (I2, I3), which are considered to be the two core elements of the multisensory full-body illusion (Ehrsson, 2012; Kilteni et al., 2015), whereas control items (C1:C4) assessed any potential effects of suggestibility or task compliance. The L1 statement ("I felt that I was located on the other bed") was added for exploratory purposes to probe possible changes in self-location (Guterstam et al., 2015) during the friend-body swap condition. The order of statements was pseudorandom: C1, I1, C2, I2, C3, C4, I3, I4, L1. Ratings of individual statements were analyzed with pairwise Wilcoxon signed-rank tests (two-sided), and *P*-values were corrected for multiple comparisons (Benjamini-Hochberg method; FDR).

Skin conductance responses: Data were recorded with the Biopac System (MP150, Goleta, CA, USA; sampling rate: 100 Hz). AcqKnowledge® software (Version 3.9.1.6, Biopac) was used to process the data. Two electrodes with electrode paste (Biopac, Goleta, USA) were placed on the participant's left index and left middle fingers (distal phalanges). We threatened the body by making a stabbing motion and stopping the knife just above the abdomen (Fig. 1C). Each knife threat lasted ~ 2 s. Before the study, we showed the "knives" to the participants to prevent extreme emotional stress in line with good ethical practice. Three threat events occurred in each condition when both participants rated 10, 20, and 30 items (Fig. 1E). The timings of threat events were marked in the recording file by the experimenters by pressing a key on a laptop immediately after the threat was presented.

Friend- and self-reference tasks: Trait adjectives were selected from (Anderson, 1968). We chose items that were comprehensible by nonnative English speakers and that showed the highest variability of ratings in the pilot study (N=10). Presentation® software (version 16.4, Neurobehavioral Systems, Inc., Berkeley, CA, USA) was used to present all stimuli and record responses. All items were presented through headphones worn by the participants. Responses were given by a key press on a numeric keypad ("How much does this trait refer to your friend/yourself?"; 1 "not at all"; 9 "very much"). Each item was preceded by a "fixation beep" (200 ms). Trait item duration was on average 0.8±0.1 s. After hearing each trait, the participants had a maximum of 6 s to provide a rating. The intertrial interval was 1, 1.5, or 2 s. The participants rated the same 120 trait adjectives in the friend- and self-reference tasks (30 traits in syncF, asyncF, syncS, and asyncS; see also the main text and Data S1).

Memory task: In the recognition memory task, 120 "old" items (the same as in the friend- and self-rating tasks) were randomly intermixed with 120 "new" trait adjectives (Fig. 5A; Data S1 and S2). The participants used the left and right mouse keys to indicate whether they had already heard a given word during the experiment. Key assignment was counterbalanced between the participants. Stimulus length was on average 0.7±0.1 s. After each word, the participants had a maximum of 2.5 s to give a response. After each response, the participants received feedback ("correct", "incorrect", or "too long"). The interval between trials varied (1, 1.5, or 2 s).

Analysis of illusion questionnaires: To assess the overall strength of the full-body illusion and to eliminate potential suggestibility or task-compliance effects, we calculated "illusion scores" as differences between the average illusion (I1:I3) and control (C1:C4) ratings for each participant in each condition (van der Hoort et al., 2017, 2011). These illusion scores were analyzed with the linear mixed model: score ~ 1 + synchrony + body + 1|id (Tables S1 and S2). The factors of "body" and "synchrony" had two levels each (self vs. friend and synchronous vs. asynchronous, respectively), and both of these factors were the fixed effects in the model. The "1|id" refers to the random intercept, which accounted for general variability between the participants. Follow-up tests (syncF vs. asyncF and syncS vs. asyncS) used the following linear mixed models: score ~ 1 + condition + 1|id. For the results of individual statements, see Fig. S2 and Tables S3 and S4.

Analysis of skin conductance responses: The amplitude of each response was identified as the difference between the maximum and minimum conductance values in the 0-6 s period after a knife threat. Skin conductance values were square-root-transformed, in line with common practice (Dawson et al., 2000). Data were analyzed with the following linear mixed model: response ~ 1 +

synchrony + repetition + 1|id (Table S1 and S2). The fixed effect of repetition (values from 1 to 12) indicated which knife threat a given event was during the course of the experiment. It is well established that skin conductance responses decrease with subsequent threats (Dawson et al., 2000), and we found this habituation effect as well (b=0.7; SE=0.05; t=14.8; P<0.005; Fig. S7A). Notably, a transformed repetition number (1/n) substantially improved the fit of the linear model to the data (χ^2_1=58.6; P<0.001; Fig. S7B). For the analyses presented in Fig. 3D and Fig. 4F, we extracted residuals from the following model, response ~ 1 + repetition, and calculated the difference between average responses in the syncF and asyncF conditions for each participant. In this way, we reduced the confounding habituation effect (see earlier) and measured the physiological friend-body-swap illusion more directly. For purely descriptive purposes, we further displayed the time courses of skin conductance responses (Fig. 2C). To do so, we performed the following steps: (i) we extracted data segments between -10 to 20 s around each knife threat marker; (ii) we manually selected a response onset in each segment (for "no response" trials where the difference between baseline and peak was < 0.05 μS, the "response onset" was set to the marker time); (iii) we removed a linear trend from the signal ("detrend" MATLAB function) and baseline corrected each segment (subtracted the average value from the -5 to 0 s period before the response onset); and (iv) we averaged all trials from each condition. By time-locking each response to its onset, we accounted for typical physiological variability with regard to latencies of skin conductance responses (Dawson et al., 2000).

Analysis of self- and friend-ratings: The number of personality traits that were rated both with regard to the self and the friend (i.e., traits that were used to calculate the self-to-friend similarity) was on average 29.2 per condition (min. 20 out of 30 possible traits), which shows that there were enough data points to assess multiple aspects of one's own and the friend's personalities in each of the four conditions (Fig. S7D). We also checked whether personality ratings showed desired variability (i.e., cosine similarity would not have been very meaningful if the participants used only one or two different ratings to describe their own and their friend's personalities). We found that in almost all (99.7%) condition-specific datasets, the participants used five or more different rating-values, which indicates that our choice of the similarity measure was appropriate (Fig. S7C). To account for the fact that some traits (e.g., aggressive) are generally likely to be rated low whereas other traits (e.g., nice) are generally likely to be rated high, we ran a linear mixed-model with a random intercept of trait-type (rating ~ 1|trait). This preprocessing step essentially set different "baselines" for different traits and thus made the remaining variability in ratings more relevant to our actual experimental manipulation. It is noteworthy that (i) the key findings of the present study were replicated when we used raw ratings instead and (ii) that the abovementioned preprocessing step did not bias our subsequent analyses because it was run on all friend ratings and self-ratings from all conditions combined. Residuals from the "rating ~ 1|trait" model were then used to calculate cosine similarity between friend ratings (FR) and self-ratings (SR) for each participant in each condition (*i* in the formula refers to each trait in a given dataset).

$$\text{cosine similarity} = \frac{\sum_{i=1}^{n} FR_i \; SR_i}{\sqrt{\sum_{i=1}^{n} FR_i^2} \sqrt{\sum_{i=1}^{n} SR_i^2}}$$

To account for general between-subject differences in the degree of similarity between self-ratings and friend ratings, similarity scores from each condition were corrected in the following way: similarity score from a given condition = score from this condition – average of scores from all conditions for a given participant. Structural similarity data were preprocessed in the same way as above, but the similarity between the "self" and "friend" distance matrixes in each condition was calculated with the Spearman correlation test. For pairwise comparisons at the group level (syncF vs. asyncF; syncF vs. syncS; syncF vs. asyncS), z-scored data were analyzed with the following linear mixed model: similarity ~ 1 + condition + 1|id. The analyses presented in Figs. 4E, 4F, and S4 were conducted on raw Spearman correlation coefficients.

Analysis of memory data: Only "old" traits that were rated in the self- and friend-reference tasks (i.e., traits followed by a button press) were included in the analysis of memory data (n=7593 out of 7800). In this way, we ensured (i) that similarity and memory datasets were fully compatible (Fig. 5C) and (ii) that all traits had been heard and noticed during stimulus encoding. Behavioral performance during the self-reference task did not differ significantly between conditions, which further indicates that all conditions were associated with similar attentional engagement (number of "misses": syncF vs. asyncF; t_{64}=-0.74; P=0.46; BF_{01}=5.67; syncF vs. syncS; t_{64}=-1.37; P=0.18; BF_{01}=3.03; syncF vs. asyncS; t_{64}=0.78; P=0.44; BF_{01}=5.49; reaction times: syncF vs. asyncF; t_{64}=-1.2; P=0.24; BF_{01}=3.72; syncF vs. syncS; t_{64}=0.76; P=0.45; BF_{01}=5.57; syncF vs. asyncS; t_{64}=1.55; P=0.13; BF_{01}=2.37; paired t-tests; two-sided; N=65). All "new" traits were included in the analysis of memory data (n=7800). For the main analysis, we calculated "d-primes" for each participant in each condition separately. These indexes assessed how well the participants were able to discriminate between the "new" and "old" items in an unbiased way (Wickens, 2002). The average d-prime from all participants and all conditions combined was 2.51±0.07, which is well above the chance level (t_{64}=35.03; P<0.005; one-sided). This shows that, in general, the participants performed very well in discriminating between the old and new words. D-primes from each condition were corrected in the following way: d-prime from a given condition = d-prime from this condition – average of d-primes from all conditions for a given participant. This correction accounted for the between-subject variability in the overall memory capacity. For planned comparisons (syncF vs. asyncF, syncF vs. syncS, and syncF vs. asyncS), we used the following linear mixed model: d-prime ~ 1 + condition + 1|id. Furthermore, the “criterion” values did not differ significantly between the four conditions (syncF: -0.623 ± 0.025; syncS: -0.633 ± 0.025; asyncF: -0.634 ± 0.026; asyncS: -0.631 ± 0.026; $F_{3,256}$=0.038; P=0.99).

General statistical information: All analyses were performed in RStudio and R software (Version 3.3.3, The R Foundation for Statistical Computing, https://www.r-project.org). Linear mixed models were estimated using the “lme4” package. For analyses that focused specifically on the effect of illusory ownership of the friend’s body, which can only vary between the syncF and asyncF conditions, we used similarity indexes from the same two conditions (Fig. 3C, 3D, 4E, 4F). In turn, for the analysis that tested how the updating of self-concept during syncF affects memory performance in this condition (Fig. 5C), we used the difference between syncF versus all control conditions because this index captures what is unique to syncF also compared to the conditions with one’s own body. Model selection was performed with the “lmerTest” package (the “step” function; Table S1). *P*-values for the F-tests were based on Satterthwaite’s approximation to degrees of freedom as implemented by the “lmerTest” package (Table S2). *P*-values for all correlations and planned pairwise tests were obtained with the bootstrapping technique (“boot” package; 10000 simulations). D-primes were calculated with the “psycho” package and Bayes factors with the “BayesFactor” package (r=0.707). For hierarchical clustering (Fig. 4C), we used the “hclust” R function.

Supplementary references:

Anderson, N.H., (1968). Likableness ratings of 555 personality-trait words. J. Pers. Soc. Psychol. 9, 272–279.

Aron, A., Aron, E.N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. J. Pers. Soc. Psychol. 63, 596–612.

Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. Proc. Natl. Acad. Sci. 110, 12846–12851.

Dawson, M.E., Schell, A.M., and Filion, D.L., (2000). The electrodermal system. In Handbook of Psychophysiology, J.T. Cacioppo, L. Tassinary, G. Berntson, ed. (Cambridge University Press), pp. 200–223.

Ehrsson, H.H. (2012). The concept of body ownership and its relation to multisensory integration. In The New Handbook of Multisensory Processes, B.E., Stein, ed. (MIT Press), pp. 775–792.

Furman, W., and Buhrmester, D. (2009). The network of relationships inventory: behavioral systems version. Int. J. Behav. Dev. 33, 470–478.

Guterstam, A., Björnsdotter, M., Gentile, G., and Ehrsson, H. (2015). Posterior cingulate cortex integrates the senses of self-location and body ownership. Curr. Biol. 25, 1416–1425.

Kilteni, K., Maselli, A., Kording, K.P., and Slater, M. (2015). Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. Front. Hum. Neurosci. 9, 141.

Peck, T.C., Doan, M., Bourne, K.A., and Good, J.J. (2018). The effect of gender body-swap illusions on working memory and stereotype threat. IEEE Trans. Vis. Comput. Graph. 24, 1604–1612.

Petkova, V.I., and Ehrsson, H.H. (2008). If I were you: perceptual illusion of body swapping. PLoS One 3, e3832.

van der Hoort, B., Guterstam, A., and Ehrsson, H.H. (2011). Being Barbie: the size of one's own body determines the perceived size of the world. PLoS One 6, e20195.

van der Hoort, B., Reingardt, M., and Ehrsson, H.H. (2017). Body ownership promotes visual awareness. Elife 6:e26022, 1–22.

Wickens, T.D. (2002). Elementary signal detection theory (Oxford University Press).

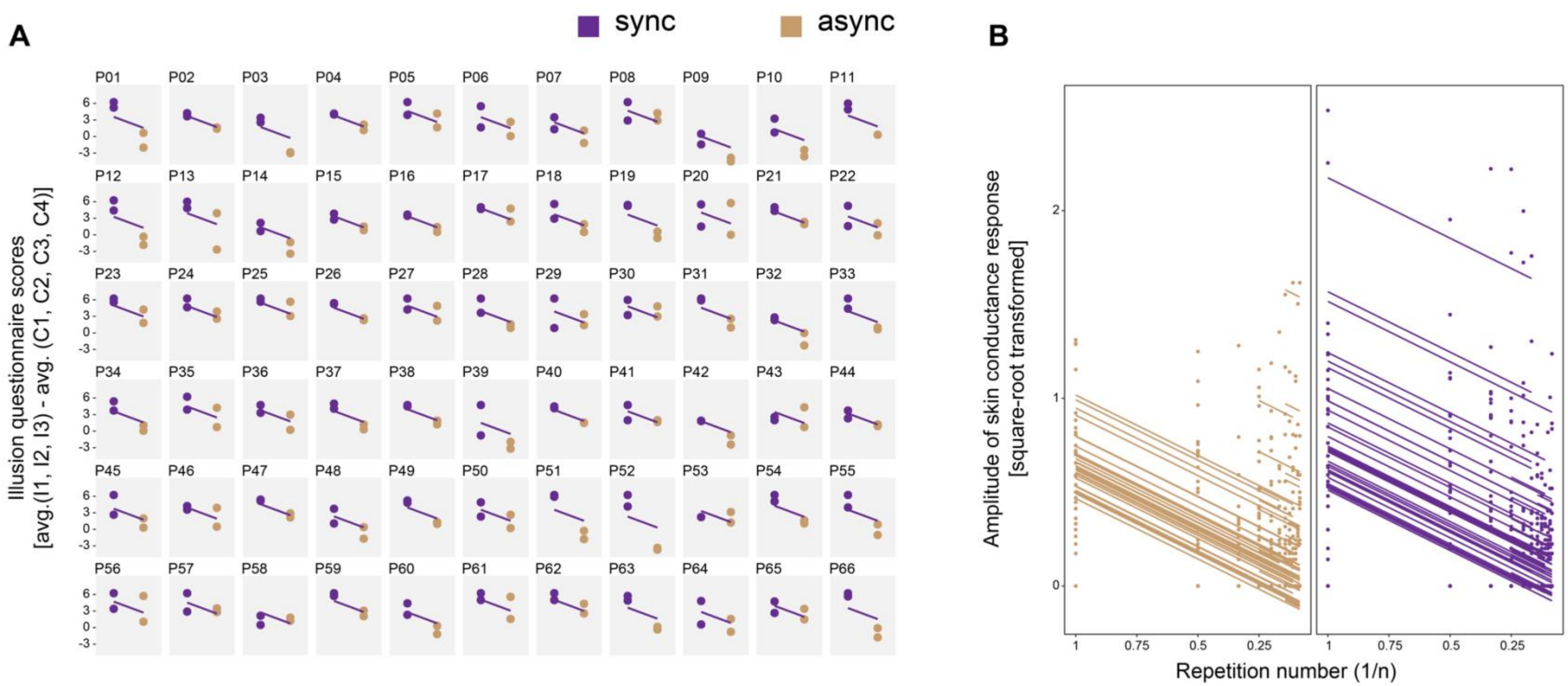

Fig. S1. (A) Illusion questionnaire scores. Related to Fig. 2. Each panel shows data from one participant, where each point corresponds to an illusion score from one condition. The lines represent the model's predictions of the main effect of synchrony. For most participants, the illusion scores were higher in the synchronous than the asynchronous conditions. **(B) Skin conductance responses.** Each line corresponds to the model's prediction of the main effect of repetition for one participant. Points correspond to skin conductance responses in individual knife threat trials. Skin conductance responses during the synchronous conditions (right panel) were generally higher than during the asynchronous conditions (left panel).

Fig. S2. Illusion questionnaire results for individual items. Related to Fig. 2A. Plots show means±*SE*. For medians and ranges, see Tables S3 and S4. Data for each statement were analyzed by pairwise comparisons (Wilcoxon signed-rank tests; *P*-values were FDR-corrected; N=66). Notably, the ratings of individual illusion statements (I1-I3) were significantly higher in the synchronous condition than in the corresponding asynchronous condition (syncF vs. asyncF and syncS vs. asyncS). Some control statements showed significant differences between the synchronous and asynchronous conditions as well, but in those cases, the ratings from synchronous conditions indicated stronger disagreement (i.e., ratings below zero) than the already low ratings from asynchronous conditions.

Fig. S3. (A) Item-by-item similarity data. Related to Fig. 3B. Each panel shows data from one participant, where each point represents the degree of similarity between self-ratings and friend ratings from one condition. The green, blue, and yellow lines represent the models' predictions of the following differences: syncF vs. asyncF, syncF vs. syncS, syncF vs. asyncS, respectively. Self-ratings and friend ratings were usually more similar in the syncF than in the other conditions. **(B) D-prime indexes of episodic recognition memory. Related to Fig. 5B.** The display convention is analogous to panel A, but each point represents a d-prime value from a given condition (see the legend). Trait adjectives encountered during the syncF condition were generally remembered worse than traits encountered in the other conditions.

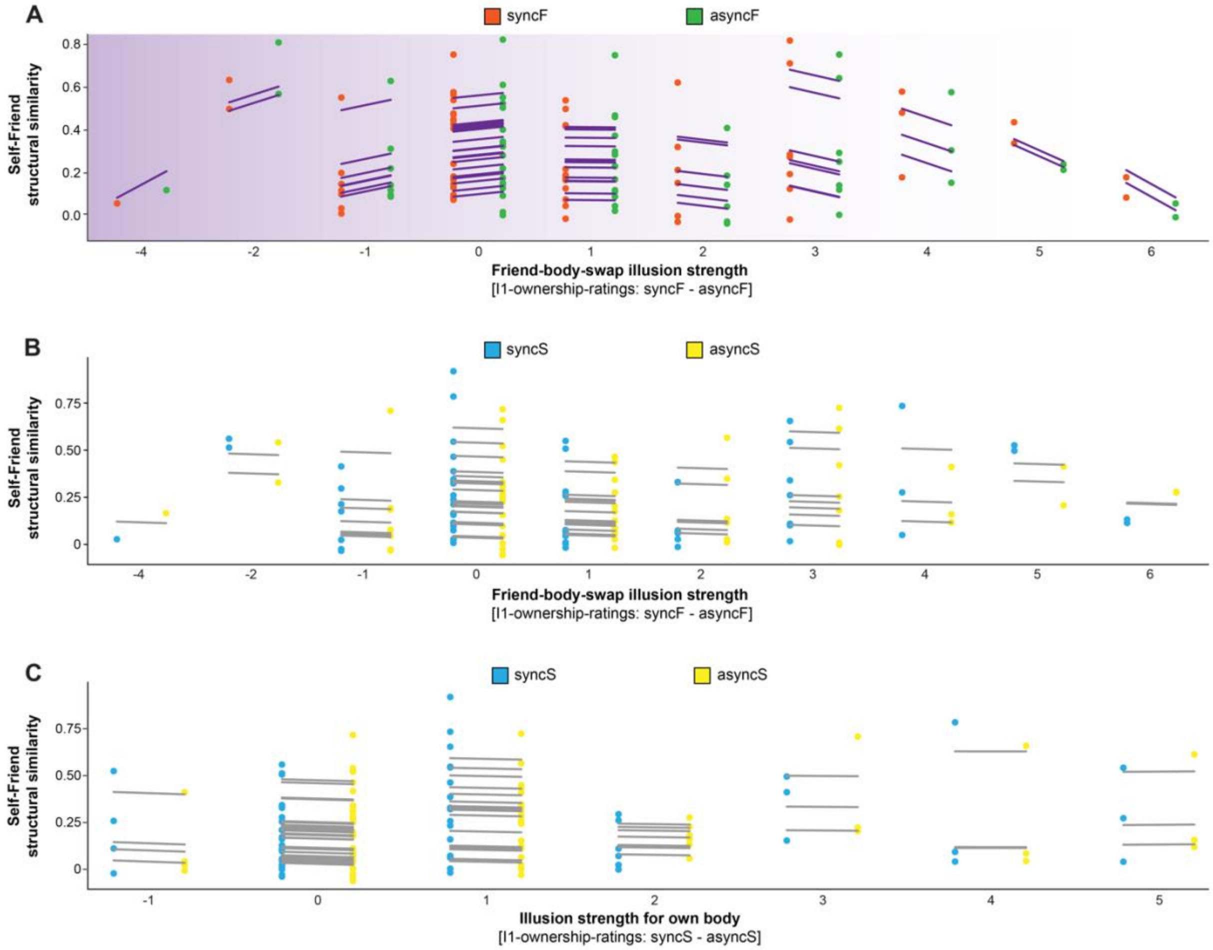

Fig. S4. Additional analyses of structural similarity data. Related to Fig. 4. (A) Strong illusory ownership of the friend's body during syncF was related to increased structural similarity in the syncF condition compared to the asyncF condition, suggesting that the "new" bodily self updated beliefs about the participant's own personality so that they became more similar to beliefs about the friend's personality [condition × ownership: $F_{1,65}$=4.01; P=0.047; LMM: similarity ~ condition × ownership + (1|id); two-sided; N=65]. **(B)** A control analysis showed that there was no significant relationship between the friend-body-swap illusion strength and the degree of structural similarity in the syncS and asyncS conditions, which indicates that the effect shown on panel A was specific to syncF [condition × ownership: $F_{1,65}$=0.001; P=0.99; LMM: similarity ~ condition × ownership + (1|id); two-sided; N=65]. **(C)** Another control analysis demonstrated that there was no significant relationship between ownership of one's own actual body in syncS and the degree of self-friend similarity in syncS versus asyncS, which suggests that our main finding (panel A) was related to illusory ownership of the friend's body specifically and not to body ownership more generally [condition × ownership: $F_{1,65}$=0.03; P=0.87; LMM: similarity ~ condition × ownership + (1|id); two-sided; N=65]. Individual lines in each plot represent the models' predictions of the main effect of condition at different levels of body ownership. Each dot indicates structural similarity for one participant in one condition.

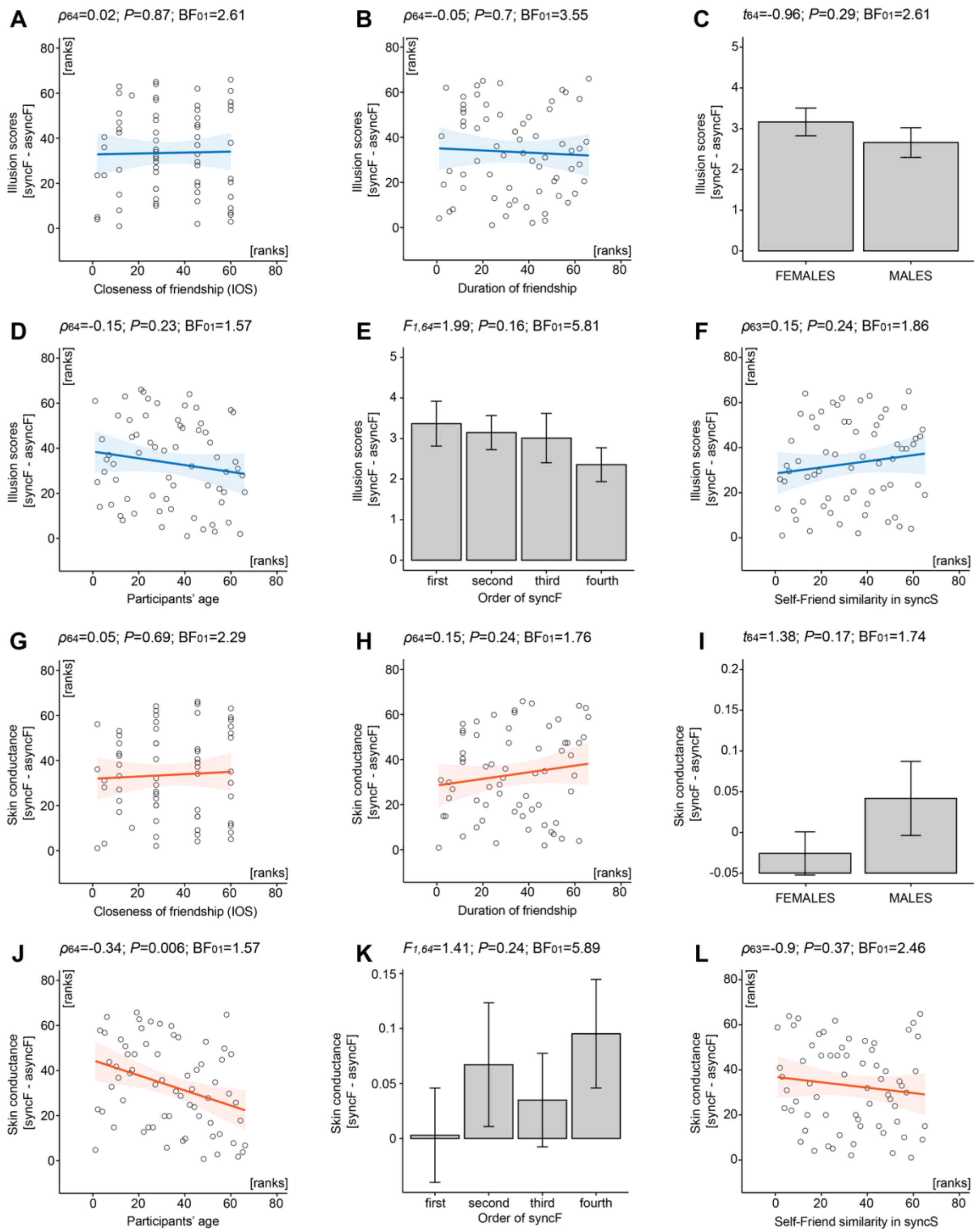

Fig. S5. Control analyses of potential confounding factors that could affect the strength of the friend-body-swap illusion. Related to Fig. 2. There was no significant relationship between illusion scores in the syncF condition and closeness of friendship (IOS; Inclusion of Other in the Self scale) **(A)**, duration of friendship **(B)**, participants' sex **(C)**, participants' age **(D)**, condition order **(E)**, or similarity between ratings of one's own and the friend's personalities in the syncS baseline condition **(F)**. A similar pattern of results was present for the skin conductance measure of the friend-body-swap illusion **(G-L)**. Please note that the participant's age correlated significantly with skin conductance responses **(J)** but not with illusion scores **(D)**; thus, future studies are needed to determine whether age consistently modulates the strength of full-body illusions. To analyze continuous variables, we used Spearman's correlation tests. The effect of participants' sex was assessed with an independent-samples t-test. Condition order was analyzed with a one-way between-subjects ANOVA. Bayes factors (BF_{01}) indicate support for the null hypothesis. All *P*-values are two-sided. Bar plots correspond to means±*SE*.

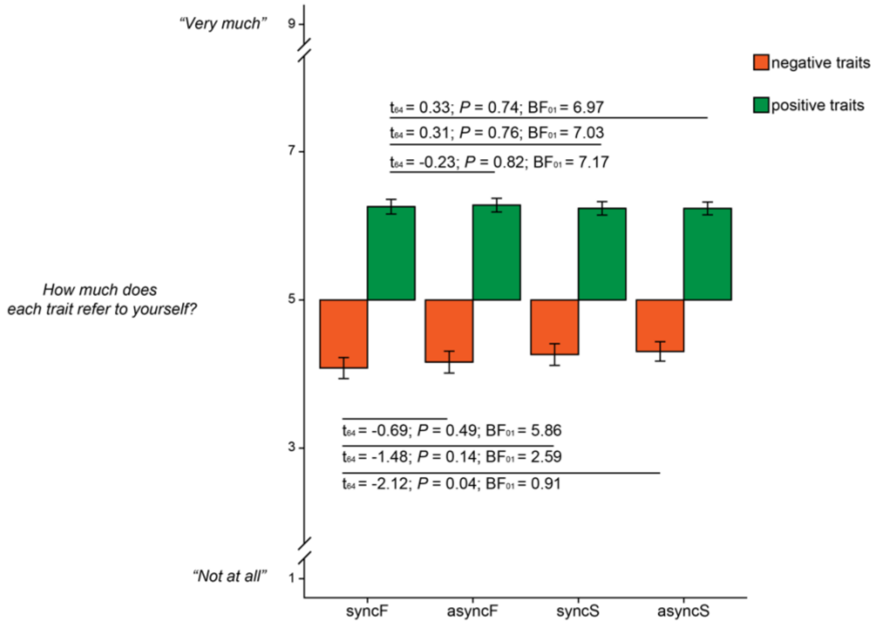

Fig. S6. Control analyses showed that increased similarity between ratings of one's own and friend's personalities in syncF was not associated with generally more negative ratings of one's own personality in this condition. Related to Fig. 3. One could argue that uncertainty about one's own body, presumably induced by the friend-body-swap illusion, could reduce the general tendency to evaluate oneself more positively than others ("self-enhancement bias"). Such a potential reduction of the self-enhancement bias could by itself increase the similarity between ratings of one's own and the friend's personalities. To test this possibility, we asked five independent raters to indicate whether each trait from the experiment was positive, negative, or neutral in their opinion. If the majority of raters indicated the same category, a given trait was assigned to this category ("ties" were assigned to the neutral category). This procedure resulted in 61 traits classified as positive, 38 traits classified as negative, and 21 traits classified as neutral. We found that self-ratings of negative traits did not increase significantly in syncF as compared to other conditions (i.e., self-views did not become more negative) and self-ratings of positive traits did not significantly decrease in syncF as compared to other conditions (self-views did not become less positive). These results speak against the possibility that illusory ownership of the friend's body reduced the self-enhancement bias and instead support our main interpretation that the illusion dynamically updated the multidimensional content of self-concept. Pairwise comparisons used paired t-tests (two-sided). Bayes factors (BF_{01}) indicate support for the null hypothesis. Bar plots correspond to means±*SE*.

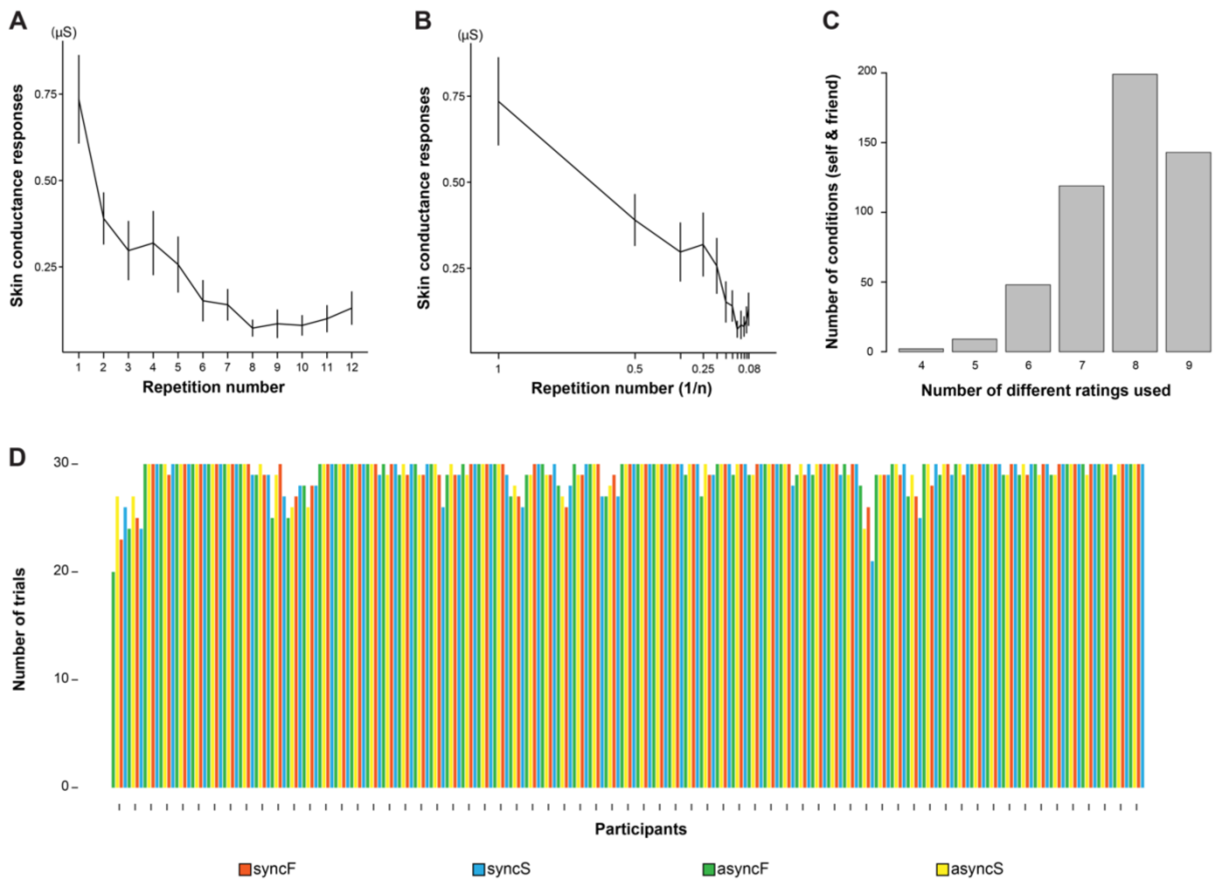

Fig. S7. Data quality checks. Related to Fig. 2-6. (A) Skin conductance responses decreased exponentially with subsequent knife threats (means±*SE*; data combined from all conditions and all participants). **(B)** The transformed repetition number (1/n) "linearized" this decrease and provided a substantially better fit of the linear mixed model to the data (χ^2_1=58.6; *P*<0.001). **(C)** In almost all single-condition datasets (99.7%), the participants used 5 or more different values to rate their own or their friend's personality, which validates our choice of similarity measures. **(D)** The number of traits rated with regard to one's own and the friend's personalities (i.e., only these traits were used to calculate the self-to-friend similarity) was sufficiently high to assess multiple aspects of one's own and the friend's personalities (i.e., min. 20 out of 30 possible traits per condition; mean = 29.2).

Table S1. Model selection^. Related to Fig. 2.

	full model	df	AIC	selected model	df	AIC
IQS	score ~ sync × body + (1|id)	6	1049	score ~ sync + body + (1|id)	5	1047
SCR	scr ~ sync × body + rep + (1|id)	7	20	scr ~ sync + rep + (1|id)	5	16

^ – For model selection, we used the "lmerTest" package ("step" function). All models included fixed and random intercepts. Models including interactions also included main effects; for example, "sync×body" is equivalent to "1 + sync + body + sync×body".
Abbreviations: AIC – Akaike information criterion; **body** – factor with two levels: own body vs. friend's body; **df** – degrees of freedom; **id** – participants; **IQS** – illusion questionnaire scores (avg. (I1+I2+I3) – avg. (C1+C2+C3+C4); **rep** – SCR repetition number; **SCR** – skin-conductance responses; **sync** – factor with two levels: synchronous vs. asynchronous.

Table S2. Statistical analysis of illusion questionnaire scores and skin conductance responses. Related to Fig. 2.

	Model	Effect	dfN	dfD	*F*	*P*
IQS	score ~ sync + body + (1|id)	sync	1	198	296.43	<0.005
		body	1	198	37.91	<0.005
	syncF vs. asyncF: score ~ sync + (1|ID)	sync	1	66	140.61	<0.005
	syncS vs. asyncS: score ~ sync + (1|ID)	sync	1	66	138.52	<0.005
SCR	scr ~ sync + rep + (1|ID)	sync	1	726	9.00	<0.005
		rep	1	726	459.48	<0.005
	syncF vs. asyncF: scr ~ sync + rep + (1|ID)	sync	1	330	10.41	<0.005
		rep	1	344	134.66	<0.005
	syncS vs. asyncS: scr ~ sync + rep + (1|ID)	sync	1	330	4.49	0.035
		rep	1	346	54.60	<0.005

Abbreviations: asyncF – synchronous-Friend condition; **asyncS** – synchronous-Self condition; **body** – factor with two levels: own body vs. friend's body; **dfN** – degrees of freedom in the numerator; **dfD** – degrees of freedom in the denominator; ***F*** – F-ratio; **id** – participants; **IQS** – illusion questionnaire scores (avg. (I1+I2+I3) – avg. (C1+C2+C3+C4); ***P*** – *P*-value; **rep** – SCR repetition number; **SCR** – skin-conductance responses; **sync** – factor with two levels: synchronous vs. asynchronous; **syncF** – synchronous-Friend condition; **syncS** – synchronous-Self condition.

Table S3. Questionnaire results for individual items in the syncF and asyncF conditions. Related to Fig. 2A and S2.

	syncF						asyncF							
Items:	**min**	**Q1**	**Q2**	**M**	**Q3**	**max**	**min**	**Q1**	**Q2**	**M**	**Q3**	**max**	**Z^**	***P*^^**
I1: It felt as if the body I saw was my own body.	-3.0	1.0	2.0	1.4	2.0	3.0	-3.0	-1.0	1.0	0.4	2.0	3.0	3.99	<0.005
I2: It felt as if the stick I saw caused the touch I experienced.	-3.0	2.0	3.0	2.2	3.0	3.0	-3.0	-3.0	-2.0	-0.9	1.0	3.0	6.49	<0.005
I3: It seemed that the touch I felt was applied to the body I saw.	-3.0	2.0	3.0	2.1	3.0	3.0	-3.0	-2.8	-1.0	-0.8	1.0	3.0	6.49	<0.005
C1: It felt as if I had two bodies at the same time.	-3.0	-3.0	-2.0	-1.0	1.0	3.0	-3.0	-1.0	1.0	0.0	1.8	3.0	-3.84	<0.005
C2: It felt like I had no body.	-3.0	-3.0	-3.0	-2.1	-1.3	1.0	-3.0	-3.0	-2.0	-1.4	0.0	2.0	-3.22	<0.005
C3: It felt as if my body was turning artificial.	-3.0	-3.0	-1.0	-0.9	1.0	3.0	-3.0	-2.0	0.0	-0.5	1.0	3.0	-2.29	0.029
C4: It felt as if my body was empty inside.	-3.0	-3.0	-2.0	-1.7	-1.0	2.0	-3.0	-3.0	-2.0	-1.3	0.8	3.0	-2.21	0.031
L1: It felt as if I was located on the other bed.	-3.0	-2.0	1.0	0.3	2.0	3.0	-3.0	-1.0	1.0	0.6	2.0	3.0	-0.99	0.322

Please note that the illusion statements (I1-I3) in syncF were affirmed by most participants (Q2/median ≥ +2), whereas the control statements (C1-C4) were typically rejected with negative median rating scores. (^) Wilcoxon signed-rank tests (N=66). (^^) FDR-corrected *P*-values (two-sided). **Abbreviations:** M – mean; Q1-Q3 – quartiles.

Table S4. Questionnaire results for individual items in the syncS and asyncS conditions. Related to Fig. 2A and S2.

	syncS						asyncS							
Items:	**min**	**Q1**	**Q2**	**M**	**Q3**	**max**	**min**	**Q1**	**Q2**	**M**	**Q3**	**max**	**Z^**	***P*^^**
I1: It felt as if the body I saw was my own body.	-2.0	2.3	3.0	2.6	3.0	3.0	-2.0	1.0	2.0	1.6	3.0	3.0	4.83	<0.005
I1: It felt as if the stick I saw caused the touch I experienced.	-3.0	3.0	3.0	2.6	3.0	3.0	-3.0	-3.0	-1.0	-0.5	1.0	3.0	6.48	<0.005
I1: It seemed that the touch I felt was applied to the body I saw.	-3.0	2.0	3.0	2.4	3.0	3.0	-3.0	-2.0	0.0	0.0	2.0	3.0	5.87	<0.005
C1: It felt as if I had two bodies at the same time.	-3.0	-3.0	-3.0	-2.5	-2.3	2.0	-3.0	-2.0	0.0	-0.3	1.0	3.0	-6.17	<0.005
C2: It felt like I had no body.	-3.0	-3.0	-3.0	-2.4	-3.0	3.0	-3.0	-3.0	-3.0	-1.8	0.0	2.0	-2.83	0.006
C3: It felt as if my body was turning artificial.	-3.0	-3.0	-2.0	-1.3	0.8	3.0	-3.0	-2.0	0.0	-0.5	1.0	3.0	-3.28	<0.005
C4: It felt as if my body was empty inside.	-3.0	-3.0	-3.0	-1.8	-1.0	3.0	-3.0	-3.0	-2.0	-1.3	0.0	3.0	-1.94	0.053
L1: It felt as if I was located on the other bed.	-3.0	-3.0	-3.0	-2.8	-3.0	0.0	-3.0	-3.0	-3.0	-2.5	-3.0	1.0	-2.19	0.033

Please note that the illusion statements (I1-I3) in syncS were affirmed by most participants (Q2/median ≥ +2), whereas the control statements (C1-C4) were typically rejected with negative median rating scores. (^) Wilcoxon signed-rank tests (N=66). (^^) FDR-corrected *P*-values (two-sided). **Abbreviations:** M – mean value; Q1-Q3 – quartiles.

Table S5. Control analyses for the participants who showed strong vs. weak updating of self-concept in syncF. Related to Fig. 5D.

	Strong updating	**Weak updating**	**Chi / t**	***P***	**BF$_{01}$**
N	33	32			
female	20	22	0.18	0.67	2.77
age (years)	25 ± 1	27 ± 1	-1.35	0.18	1.82
friendship (months)	45 ± 8	42 ± 6	0.31	0.76	3.79
IOS	5.3 ± 0.2	5.3 ± 0.3	0.02	0.98	3.94
NRI (support)	3.4 ± 0.1	3.3 ± 0.1	0.81	0.42	2.99
NRI (negative)	1.5 ± 0.1	1.4 ± 0.1	0.9	0.37	2.79
control scores (avg. C1:C4 in syncF)	-1.2 ± 0.2	-1.6 ± 0.2	1.22	0.23	2.09

Values are counts or means±*SE*. The proportion of females was tested with the equality of proportions chi-square test. The remaining variables were tested with two-sample t-tests (two-sided). Bayes factors report evidence for the null hypothesis (BF$_{01}$). **Abbreviations:** IOS – Inclusion of Other in the Self Scale; NRI – the Network of Relationships Inventory.