Supplementary Material

Enhanced self-reported affect and prosocial behaviour without differential physiological responses in mirror-sensory synesthesia

Kalliopi Ioumpa[a,b,*], Sarah A. Graham[c], Tommy Clausner[a], Simon E. Fisher[a,c], Rob van Lier[a], Tessa M. van Leeuwen[a,*]

a Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands

b Netherlands Institute for Neuroscience, Royal Netherlands Academy for Arts and Sciences, Amsterdam, the Netherlands

c Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

* Corresponding authors:
Kalliopi Ioumpa (ioumpa.k@gmail.com)
Tessa M. van Leeuwen (tesvlee@gmail.com)

Supplementary Methods

Supplementary Results

Supplementary Figures S1-S6

Supplementary Table S1-S2

Supplementary Methods

Mirror-sensory synesthesia verification experiment

In this vision-touch interference paradigm [1], participants were asked to report the location of actual touch applied to their cheeks by an electrical tactile device while observing videos of another person or object being touched. For mirror-sensory synesthetes, the additional experience of synesthetic touch leads to more errors in situations where the actual touch and touch as observed in the video do not match.

Set up. Each trial started with a video showing either a boy, a girl, or an object (apple) being touched either on the right, left, both sides, or no touch (Fig. S1A). Simultaneously to the moment of touch in the video, a tactile device (Fig. S1B) applied real touch (left, right, both sides, or no touch) to the participant's face in a way that was perceived as spatially congruent (observed touch same as felt) or incongruent (observed touch different from what was felt); see Figure S1 for a schematic. Synesthetes, apart from the real touch, also felt the synesthetic touch on their face, which corresponded to the observed touch in the video. For some synesthetes an observed touch on e.g. the left cheek triggers a synesthetic sensation on their left cheek (anatomical correspondence), but for others the synesthetic sensation is then felt on the right cheek (i.e. as if they were looking in a mirror, a specular correspondence) [1]. In our study, congruency of the touch location (specular or anatomical) was determined according to self-report. Participants reported the location of the real touch by pressing one of the keyboard arrow keys (left - left touch, right - right touch, up - touch felt at both sides and down - no feeling of touch) as fast as possible. Their reaction times (RT) and error rates were recorded.

For synesthetes, on congruent trials the real touch was on same side of the face as the synesthetic experience. On incongruent trials, actual touch would be e.g. on the right cheek but synesthetic touch, induced by the video, could be e.g. on the left cheek. In this case the correct answer would be 'right' and a 'mirror touch error' would be answering 'both'. An example of a *no touch* trial would be not receiving any actual touch from the device but observing touch in the video, which would elicit a synesthetic touch for the synesthetes. Previous research has shown that synesthetes tend to get confused during the incongruent and *no touch* trials in this type of set-up: they make more mistakes and have longer reaction times than controls [1, 2].

Stimuli. In total, 252 trials were presented: 78 congruent, 112 incongruent and 60 no touch trials were used for each participant. Four synesthetes experienced specular mirror-touch synesthesia and 14 synesthetes experienced synesthetic touch in anatomical correspondence. For each one of the conditions, 40% of the trials involved videos applying touch to a female actor, 40% of the trials touch to a male actor, and 20% of the trials involving observed touch to an apple. The order of trials was pseudo-randomised so that the same videos (max 2 in a row), actors (max 5 repetitions), touch sides (max 3 repetitions) and congruency (max 6 repetitions) did not repeat too often. Stimuli were presented in three blocks of 84 trials each. The videos were presented centrally on a Benq XL2420Z 24" monitor with a refresh rate of 120 Hz, had an approximate duration of 4 seconds, and had a height of 10.8 cm and a width of 19.20 cm. Participants could give their response while the video was still playing. The next video started only after a response was recorded. The inter-trial interval was 1500 ms (fixation cross). In our initial videos the touch was applied by a fingertip. After one synesthete reported feeling a sensation of touch at her own finger as well, we created new videos in which the touch was applied by a plastic stick similar to the tactile device to reduce the covariant factors. Across all participants, 5 synesthetes completed the experiment with the initial fingertip videos and no differences were observed in their results.

Tactile device. The tactile stimuli were administered via an electrical device that was made in-house (Fig. S1B). The touch was delivered via two plastic sticks (one at each side of the face) with round edges made in a way that they resembled the feeling of a fingertip. Each one of the plastic sticks was attached to a flexible plastic arm that allowed for positioning of the device at the face. The plastic arms were attached to a surface supported by a microphone supporting rod. White noise was played via earplugs during the entire experiment, so participants could not determine the location of the real touch due to mechanical noise of the device. To prevent participants from moving (to ensure the device would touch them at the same location during the entire experiment) a chin rest was used.

Procedure. Before the actual experiment, participants familiarised themselves with the task through 12 practice trials. After the end of each block, participants had a break where they could move freely and the white noise was temporarily switched off. During the break after the first block, the second saliva measurement for cortisol level assessment took place. The experiment took 40 minutes in total.

Data-analysis. Error rates and RT data were analysed using in-house routines in Matlab 2013a (MathWorks). Raw data were cleaned by rejecting RTs of incorrect trials and trials with RTs that were above or below 2 SD from the subject and condition mean. Error rates and reaction times were compared for the different stimulus conditions at the group level (congruent, incongruent, and no touch stimuli for both object and person stimuli) in a repeated measures ANOVA. At the individual participant level, error rates were compared for the congruent and incongruent person stimuli using two samples T-tests.

Arousing pictures experiment

Eye movement procedure

Fixations and time spent within Regions-of-Interest (ROIs) in the images were calculated. ROIs were defined as area(s) of the image in which the physical touch was apparent (e.g. the area of physical contact of a needle with the skin). Two researchers marked the ROIs for each picture independently and the respective overlap of both selections served as the actual ROI for that picture that went into further analyses. A fixation was considered, if it would last at least 40 ms and would have a maximum dispersion of 100 pixels compared to the previous fixation. Viewing times for each ROI were calculated as the fraction of the total time spent observing the picture (i.e. the sum of all fixation durations within the ROI divided by the total fixation durations over the whole screen). Number of fixations and relative viewing time were normalised to account for the size of the given ROIs relative to the entire screen. Due to varying ROI sizes for each stimulus, results for both dependent measures were discounted for larger and weighted higher for smaller ROI sizes.

Cortisol sample collection

Saliva was collected using the commercially available saliva collection kit of Salivette (Sarstedt). Participants chewed gently on a cotton swab for 1 minute until it became humid, and placed it back in the Salivette tube. In order to minimise differences in baseline cortisol levels, participants were instructed not to brush their teeth, eat or drink anything but water for 1 hour prior to arriving at the laboratory, not to use any recreational drugs for 3 days and to refrain from drinking alcohol, exercising, and smoking for 12 hours prior to sample collection as instructed by the manufacturer´s information (https://www.sarstedt.com/en/products/diagnostic/salivasputum/). After collection, the samples were centrifuged at 1000 rpm for 2 minutes and then stored in a freezer at -20°C until analysed. The samples were sent to Dresden LabService GmbH (http://www.labservice-dresden.de/) for cortisol determination in saliva.

During statistical analysis the phase of the menstrual cycle was included as a covariate as cortisol levels vary with menstrual cycle [3]. Participants reported the day of their menstrual cycle and/or use of hormonal contraceptives in the exit questionnaire. Menstrual cycles were classified into one of five categories: menstrual phase (day 1-6), follicular phase (day 7-14), luteal phase (day 15-30), hormonal contraceptives (oral or intra-uterine device), or menopause.

Questionnaire completion

Empathy assessments. In order to assess empathic behaviour we used the Empathy Quotient [4] and the Interpersonal Reactivity Index [5] questionnaires. The Empathy Quotient (EQ) consists of 60 questions and it is designed to measure empathy in adults. From these questions, 40 are clinically relevant and 20 are there to distract participants. There are three main subscales: for Cognitive Empathy, for Emotional Reactivity and Social Skills. Each statement has to be rated on the scale of: strongly disagree, slightly disagree, slightly agree, strongly agree. The minimum score is 0 and corresponds to least empathetic behaviour possible and the maximum score is 80 and corresponds to the most empathetic behaviour possible.

The Interpersonal Reactivity Index (IRI) consists of 28 statements and has four main scales assessing cognitive and affective aspects of empathy: Empathic Concern (EC), Personal Distress (PD), Fantasy (FS) and Perspective Taking (PT). The EC measures feelings of sympathy and compassion for others in distress, the PD self-oriented feelings of anxiety and distress in response to tense interpersonal situations, FS scale measures the tendency to project oneself into fictional situations and finally the PT scale measures the tendency to adopt the psychological point of view of others. Each item is rated on a scale ranging from "does not describe me well" to "describes me very well". For each one of the subscales, a minimum score of 0 and maximum score of 28 is possible. For PD, FS and EC subscales higher scores indicate enhanced empathy. For the PD subscale higher scores are translated to self-oriented emotional reactivity.

Theory of mind. The Reading the Mind in the Eyes test (Baron-Cohen et al., 2001) is an advanced theory of mind test where the participants' ability to put themselves into the mental state of others can be examined. Participants were presented with 25 photographs of the eye-region of the face of different actors of both sexes, and were asked to choose one out of four words that best described what the individual in the photograph was thinking or feeling. Higher scores demonstrate enhanced theory of mind.

Supplementary Results

Mirror-sensory synesthesia verification experiment

Data from 16 synesthetes and 17 controls were analysed. Two synesthetes did not follow the instructions correctly and were excluded, and their synesthesia verified by subjective report.

Group level. For the error rates (Fig. S6A), a repeated measures ANOVA with the within-subject factors Congruency (congruent, incongruent, no-touch) and Stimulus type (person, object), and the between-subjects factor Group (synesthetes, controls), revealed a significant interaction of Congruency x Group (marginal with frequentist statistics, $F(2,62)=2.59$, $p=.083$, $\eta_p^2=0.07$, but BF=7.73), as well as a main effect of Congruency ($F(2,62)=8.85$, $p<.001$, $\eta_p^2=0.22$) and a main effect of Group ($F(1,31) = 9.37$, $p=.005$, $\eta_p^2=0.23$). Stimulus type did not influence the error rates (all $p>.17$, all BF<.3). Follow up tests revealed that the main effect of group (and the group x congruency interaction) was driven by more errors for synesthetes than controls in the incongruent ($F(1,31)=7.14$, $p=.012$, $\eta_p^2=.19$) and no-touch ($F(1,31)=8.11$, $p=.008$, $\eta_p^2=.21$) conditions, with no difference in the error rates for the congruent condition ($F(1,31)=.626$, $p=.44$, $\eta_p^2=.020$). Overall, more errors were made in the incongruent than in the congruent and no-touch conditions (10.3 vs 4.6 and 5.3%, respectively). The results concur with earlier reports of more 'mirror touch errors' at the group level in the incongruent and no-touch conditions of this set-up [1].

The reaction times (Fig. S6B) were analysed with a similar ANOVA as the error rates. While no significant interaction of Congruency x Group ($F(2,62)=759$, $p=.47$, $\eta_p^2=.024$, inconclusive Bayes statistics with BF=.56) was found, a main effect of Congruency ($F(2,62)=73.7$, $p<.001$, $\eta_p^2=.70$), a main effect of Stimulus type ($F(2,62)=13.0$, $p<.001$, $\eta_p^2=.30$), a Congruency x Stimulus type interaction ($F(2,62)=9.96$, $p<.001$, $\eta_p^2=.24$), and a main effect of Group ($F(1,31)=12.1$, $p<.001$, $\eta_p^2=.29$) were present. The strong effect of Congruency was mainly driven by longer reaction times in the no-touch condition (see Fig. S6B), in which people did not receive a touch to the face but did observe touch in the video. This was anticipated as participants waited to ensure no touch would occur in the video before responding. The main effect of stimulus type was driven by faster reaction times for the object stimulus overall (1077 ms vs 1182 ms for persons), which was especially pronounced for the no-touch condition (1338 ms vs 1525 ms). Synesthetes' overall delay (1340 vs 918 ms) that was observed in all experimental conditions could be due to an overall higher level of difficulty for synesthetes during this experiment.

Individual level. In the original study where the vision-touch interference paradigm was developed [1], analyses were performed only at the group level and not at an individual level. Here, we also assess the data at the individual participant level to give more insight in the verification procedure and type of participants in the study (a more strict approach than in the original study). We took into account 1) the error rates for the congruent and incongruent person stimuli during the vision-touch interference paradigm, compared with a t-test for each individual; and 2) the self-report ratings of synesthetic sensations of pain and touch that the affective picture task elicited. It appeared that due to the nature of the interference paradigm, many controls also displayed a congruency effect in the reaction times; hence reaction times were not used to distinguish between synesthetes and controls. Detailed individual data are summarized below in Table S1.1 (synesthetes) and Table S1.2 (controls).

Comparisons between the error rates for the congruent and incongruent person stimuli for each participant individually yielded significant differences for errors for 8 out of the 16 synesthetes who completed the experiment according to instructions (see main text for details). Two control participants (C6 and C7) demonstrated a congruency effect in the verification experiment (Table S1.2), even though they did not report any synesthetic experiences in the initial email screening. One of these controls (C6) did not report any conscious synesthetic experiences in the exit questionnaire, in spite of the significant congruency effect, and was retained in the control sample. The other control (C7) additionally reported synesthetic experiences in the exit questionnaire (Table S1.2). Note that this participant was already excluded from the analyses of the picture ratings and from the Dictator's game, because she did not understand the use of the SAM scale correctly in the first experiment and did not follow the instructions of anonymity in the Dictator's game. Thus, the outcome of our main conclusions concerning these two tasks is not affected by the status of this participant. Further note that we still have included this participant in the heart rate and pupil dilation analyses; if, in fact, she is a synesthete, this would only increase the noise in our data and not enhance any effects. To verify that this individual participant was not causing the null-effects in our data we repeated the analyses of the heart rate and pupil dilation data with C7 in the synesthete group instead of the control group. It turned out that the outcomes did not change at all; hence, this one control participant did not determine the absence of significant group effects in the physiological data.

Arousing pictures experiment

Eye movements. Data from 14 controls and 15 synesthetes were included in the analysis, as for the other subjects the amount of eye-movement data was insufficient for inclusion (insufficient eye movement data were recorded (2 synesthetes) or trials had no fixations). Normalised average fixation rates and viewing times were entered into separate repeated measures ANOVAs with the within-subject factor Physical Context (negative, positive, neutral pictures) and the between-subjects factor Group (synesthetes, controls). No interactions with Group or Group effects were found (all F<1, n.s.; all .7< BF_{10}<3, inconclusive), see Figure S5. However, for both dependent variables a main effect of Physical Context was present (fixations: $F(2,54)=35.8$, $p<.001$, $\eta_p^2=0.57$; viewing time: $F(2,54)=32.1$, $p<.001$, $\eta_p^2=0.54$) which was driven by more fixations and longer looking times in the negative and positive picture condition compared to the neutral condition (for details see Fig. S5).

Questionnaire results

Online questionnaires: Exploratory correlations between online questionnaire scores and experimental outcomes. In these correlations, controls are included as well. Because the N is low and the number of comparisons is high, these analyses are considered exploratory.

Subjective ratings of the arousing pictures experiment. Participants who scored higher on the cognitive empathy subscale of the Empathy Quotient rated positive emotional faces as more calming ($r(15)=.590$, $p=.026$). Participants who scored higher on the fantasy subscale of the IRI rated unpleasant images ($r(14)=.535$, $p<.05$) and negative emotional faces ($r(14)=.604$, $p=<.05$) as more arousing; participants who scored higher on the total IRI scale rated unpleasant images as more unpleasant ($r(12)=-.663$, $p<.05$) and positive emotional faces as more pleasant ($r(12)=.616$, $p<.05$). With Bayesian statistics, all the non-significant frequentist results were inconclusive (.32< BF_{10}<1.6).

Physiological responses in the arousing pictures experiment. Participants who scored higher on the perspective taking subscale of the IRI had narrower pupils in the unpleasant and neutral images conditions ($r(14)=-.560$, $p<.05$ and $r(14)=-.598$, $p<.05$, respectively). Participants who scored higher on the emotional contagion subscale of the IRI also had more narrow pupils in the unpleasant images condition ($r(14)=-.559$, $p<.05$). These results are unexpected since wider, not more narrow, pupils are associated with a stronger stress response and stronger empathic skills were hypothesised to lead to stronger physical reactions to the unpleasant images, which is opposite to what we find. No effects were found for heart rate (all $p>.33$, all $.32< BF_{10}<.45$, inconclusive).

Reaction times and error rates in the verification experiment. There were no significant correlations of the online questionnaire scores with the outcomes of the verification experiment (all $p>.11$, all $.31< BF_{10}<.6$, inconclusive).

Amount donated in the Dictator's game. There were no significant correlations of the online questionnaire scores with amount of money donated in the Dictator's game (all $p>.13$, all $.32< BF_{10}<.94$ inconclusive).

Exit questionnaire: Correlations between strength of synesthesia and experimental outcomes. Because of the high number of comparisons in these analyses, these analyses should be regarded as exploratory.

Subjective ratings of the arousing pictures experiment. Synesthetes who indicated (in the exit questionnaire) to experience stronger synesthetic touch during the arousing pictures experiment rated unpleasant images as more unpleasant ($r(16)=-.546$, $p<.05$). Synesthetes experiencing stronger synesthetic pain during the arousing pictures experiment rated pleasant images as more calming ($r(16)=-0.528$, $p<.05$) and unpleasant images as more arousing, although this was only a nonsignificant trend with frequentist statistics ($r(16)=.438$, $p=.089$) and the Bayesian results were inconclusive ($BF_{10}=.8$). Finally, synesthetes who indicated to experience stronger synesthesia for touch in the verification experiment rated unpleasant images of the arousing pictures experiment as more unpleasant ($r(16)=-.522$, $p<.05$). The results are in the predicted direction, illustrating that the stronger the subjective strength of synesthetic experience, the more extreme the synesthetes' subjective valence and arousal ratings were.

Physiological responses in the arousing pictures experiment. There were no significant correlations (all p>.13, all .29< BF_{10}<.35 inconclusive).

Reaction times and error rates in the verification experiment. Synesthetes who indicated that they experienced stronger actual touch from the tactile device in the verification experiment made less errors in the congruent touch condition ($r(16)$=-.533, p=.033). Synesthetes who reported stronger experiences of pain during the arousing pictures experiment showed trends towards making less errors in the incongruent and no-touch conditions of the verification experiment ($r(16)$=-.468, p=.067, and $r(16)$=-.460, p=.073) with frequentist statistics while the Bayesian equivalent results were inconclusive (BF_{10}=.59 and BF_{10}=.35).

Amount of money donated in the Dictator's game. There were no significant correlations (all p>.31 and all .3< BF10<.36 inconclusive).

Online questionnaires. There was a significant correlation of the subjective strength of experienced synesthetic pain in the arousing pictures experiment with IRI Total score ($r(10)$=805, p=.005). There were no other correlations of the subjective strength of synesthesia with the EQ or IRI or their subscales, nor with the Reading the Mind in the Eyes test with frequentist statistics, while the Bayesian statistics were inconclusive (all .37< BF_{10}<1.35).

References

[1] Banissy, M. J. & Ward, J. 2007 Mirror-touch synesthesia is linked with empathy. *Nat Neurosci.* **10**, 815-816. (DOI:10.1038/nn1926).

[2] Banissy, M. J., Kadosh, R. C., Maus, G. W., Walsh, V. & Ward, J. 2009 Prevalence, characteristics and a neurocognitive model of mirror-touch synaesthesia. *Exp. Brain Res.* **198**, 261-272. (DOI:10.1007/s00221-009-1810-9).

[3] McCormick, C. M. & Teillon, S. M. 2001 Menstrual Cycle Variation in Spatial Ability: Relation to Salivary Cortisol Levels. *Horm Behav.* **39**, 29-38. (DOI:10.1006/hbeh.2000.1636).

[4] Baron-Cohen, S. & Wheelwright, S. 2004 The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *J. Autism Dev. Disord.* **34**, 163-175. (DOI:10.1023/b:jadd.0000022607.19833.00).

[5] Davis, M. H. 1983 Measuring individual differences in empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol.* **44**, 113-126. (DOI:10.1037/0022-3514.44.1.113).

Supplementary Figures

Figure S1

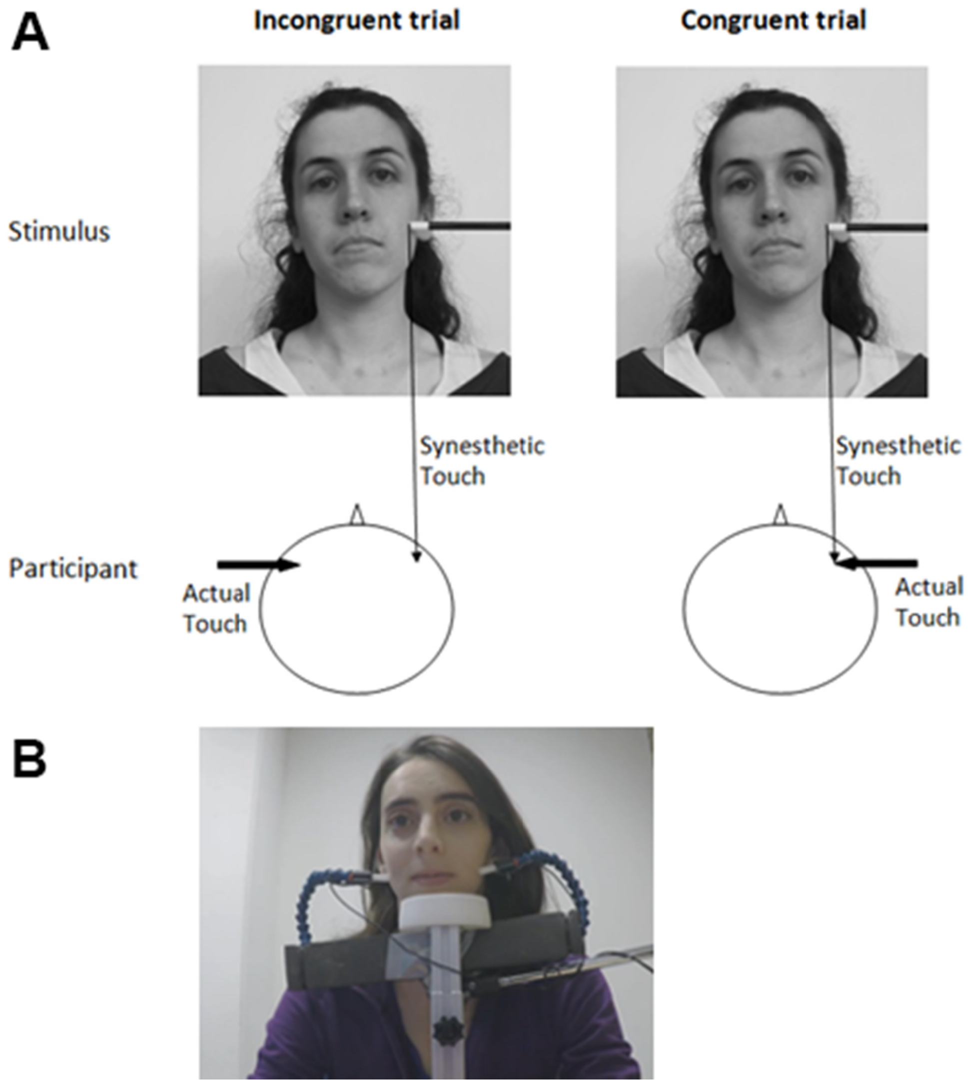

Figure S1. Interference paradigm of the mirror-sensory synesthesia verification experiment
The paradigm was adapted from Banissy & Ward (2007) [2]. A. The stimuli (top) consisted of movies of touch to the face, which elicited synesthetic touch for the synesthetes. Simultaneously with the touch in the videos, the participants were also touched on the cheek (schematic at the bottom). Actual touch could be incongruent – conflicting – with the synesthetic touch (depicted on the left, i.e. actual touch on the left cheek and synesthetic touch on the right) or congruent – matching with the synesthesia (on the right). Note that for some synesthetes congruence/incongruence was reversed because the synesthesia was not experienced as mirrored but in anatomical correspondence to their own bodies. B. Device applying actual touch, mounted on a microphone rod. The participants used a chin rest.

Figure S2

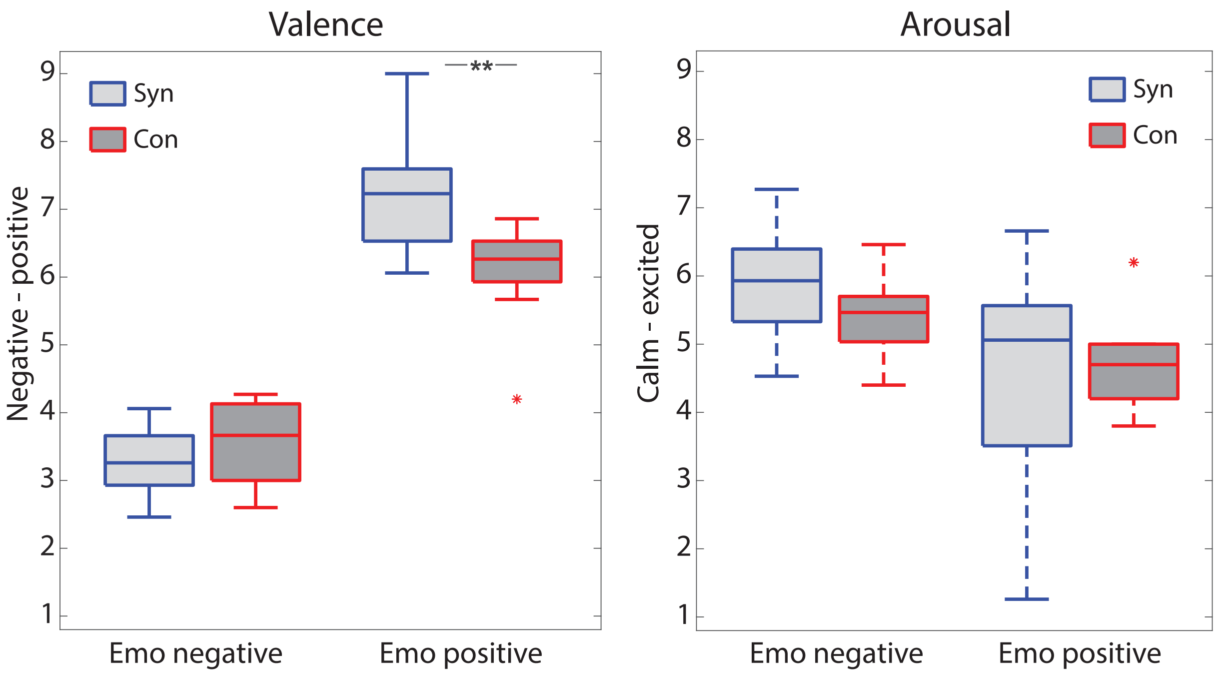

Figure S2. Ratings for negative and positive emotional faces pictures (arousing pictures experiment). Box plots summarizing the ratings results for both valence (left) and arousal (right) for synesthetes (Syn – in light grey boxes) and control participants (Con– in dark grey boxes). $^{**}p<.01$.

Figure S3

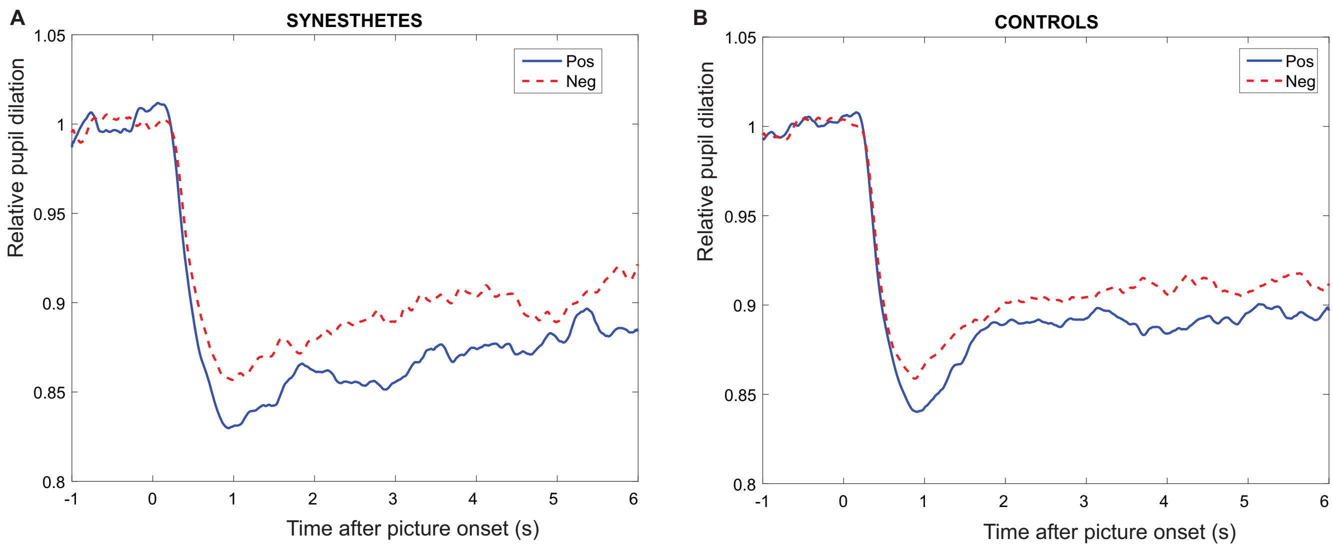

Figure S3. Pupil dilation for negative and positive emotional faces pictures (arousing pictures experiment). Pupil dilation relative to picture onset is plotted for positive (Pos) and negative (Neg) emotional faces images. For results see main text.

Figure S4

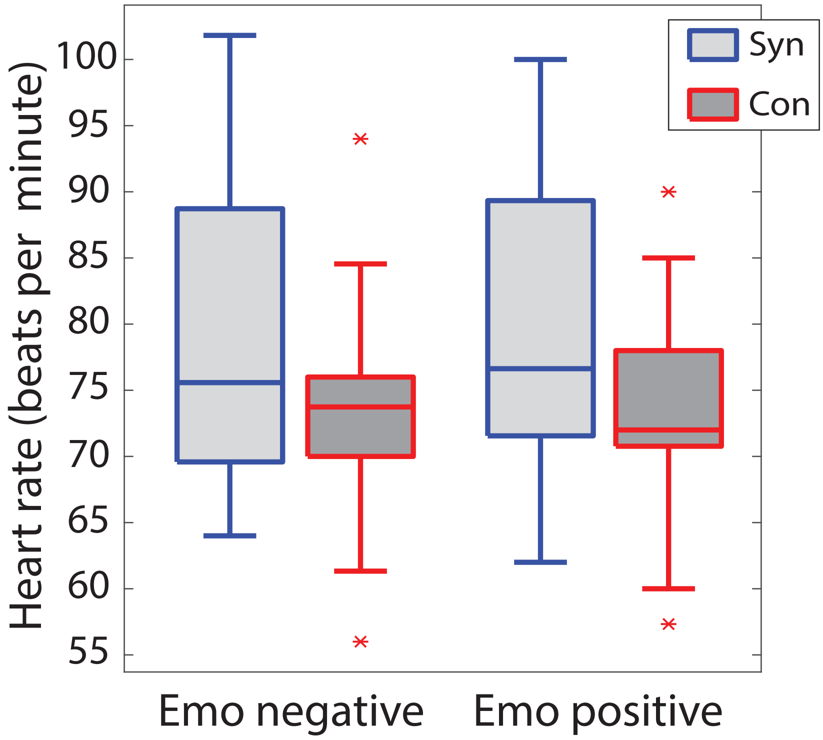

Figure S4. Heart rate for negative and positive emotional faces pictures (arousing pictures experiment). Box plots summarizing the heart rates for synesthetes (Syn – in light grey boxes) and control participants (Con– in dark grey boxes). For results see main text.

Figure S5

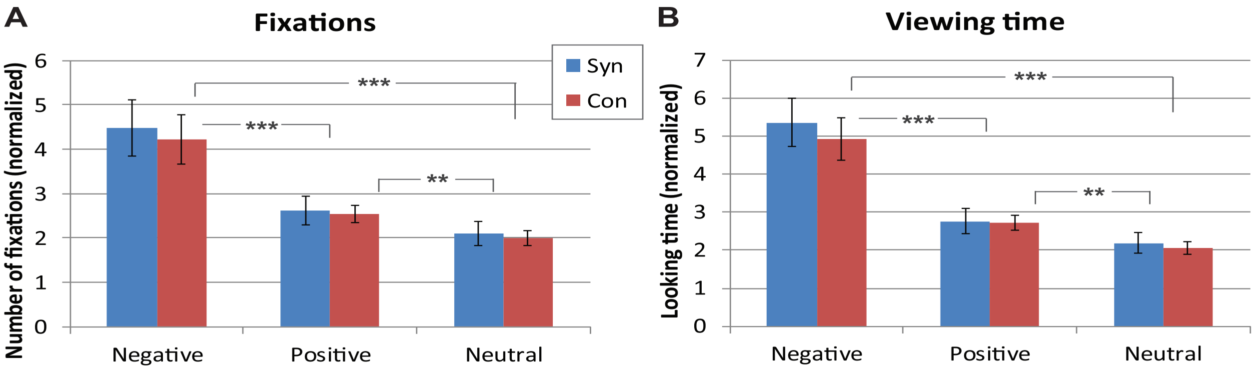

Figure S5. Eye movement parameters for negative, positive, and neutral pictures (arousing pictures experiment). Normalised number of fixations (A) and viewing times (B) for three different valence conditions. Asterisks indicate significant main effects of valence condition. For detailed results see Supplementary Results.

Figure S6

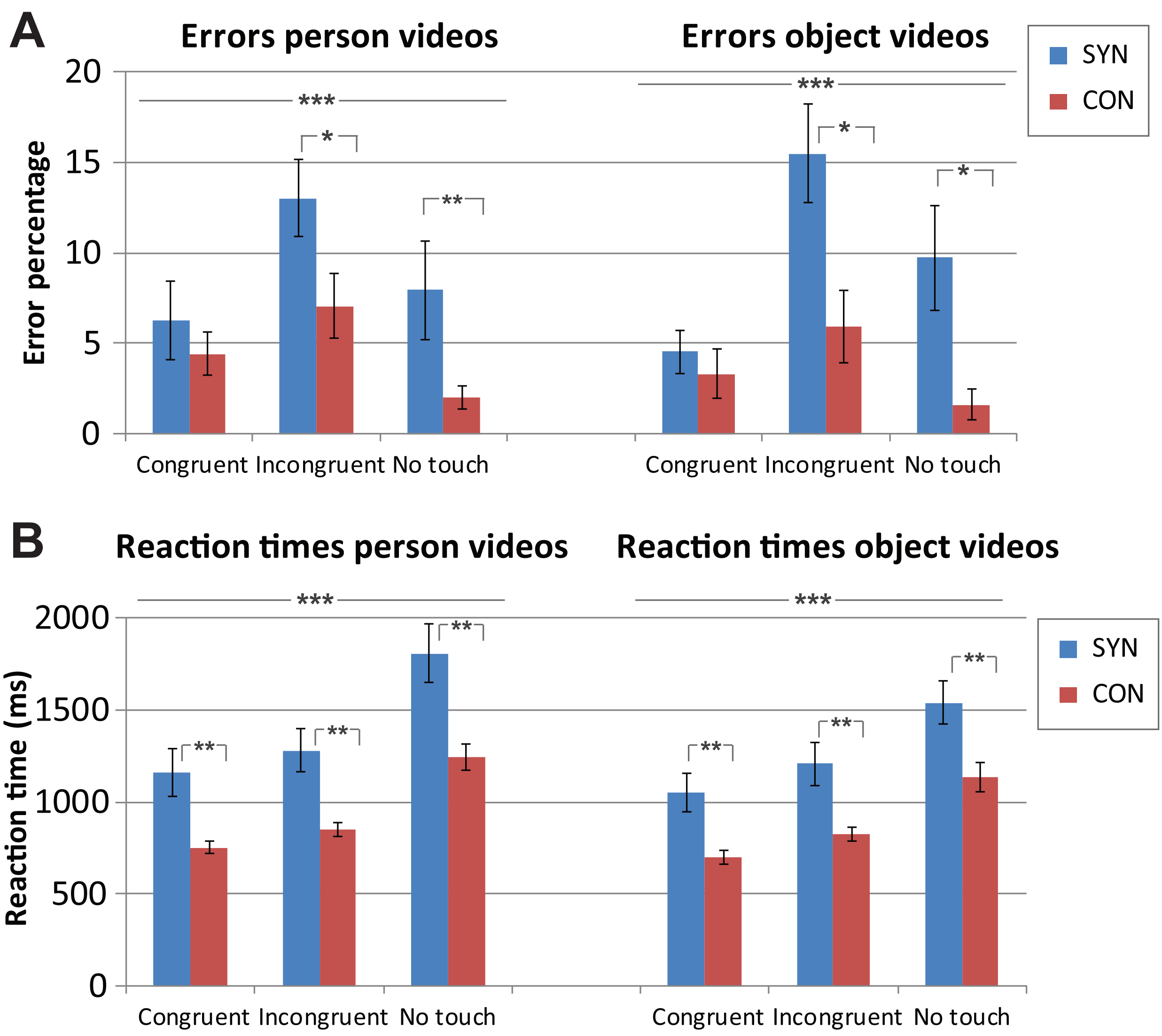

Figure S6. Results of the mirror-sensory synesthesia verification experiment
Percentage of errors (A) and reaction times (B) obtained in the verification experiment for the congruent, incongruent, and no touch conditions for both person and object stimulus types for synesthetes (blue) and controls (red). Not all significant effects are depicted in the figure, for detailed results please see the Supplementary Results. Errors bars depict ± the standard error of the mean. *p<.05, **p<.01, ***p<.001

Table S1.1 and S1.2. Errors and synesthesia ratings for individual synesthetes and controls, respectively, in the verification experiment.

Table S1.1 Synesthetes

	Errors (%) congruent	Errors (%) incongruent	p-value	t-value		Strength of syn. touch*	Strength of syn. pain*
S1	66.67	47.41	0.02	2.27	misunderstood instr.	4	4
S2	3.39	15.38	0.02	-2.37	confirmed errors	1	2
S3	14.75	23.08	0.20	-1.29	non-significant	1	5
S4	65.00	42.31	0.00	2.85	misunderstood instr.	5	1
S5	1.67	2.65	0.68	-0.41	not confirmed	4	4
S6	0.00	5.26	0.07	-1.82	marginally significant	1	5
S7	6.67	31.86	0.00	-3.88	confirmed errors	1	1
S8	3.45	0.88	0.23	1.22	not confirmed	5	5
S9	1.69	19.47	0.00	-3.34	confirmed errors	3	5
S10	0.00	7.08	0.04	-2.11	confirmed errors	3	3
S11	2.04	17.89	0.01	-2.80	confirmed errors	0	4
S12	13.33	13.16	0.97	0.03	not confirmed	1	3
S13	16.33	19.35	0.65	-0.46	not confirmed	4	4
S14	6.90	7.83	0.83	-0.22	not confirmed	2	1
S15	1.75	27.43	0.00	-4.23	confirmed errors	0	5
S16	5.17	16.96	0.03	-2.19	confirmed errors	1	2
S17	1.72	28.32	0.00	-4.38	confirmed errors	3	4
S18	0.00	1.63	0.37	-0.90	not confirmed	2	4

*Strength of synesthesia for observed touch and pain as experienced during the pictures experiment, as indicated by self-report in the exit questionnaire. The majority of synesthetes for whom mirror-touch synesthesia was not confirmed in the verification experiment reported strong synesthetic experiences for pain, except for S14, who reported rather weak synesthetic experiences overall.

Table S1.2 Controls

	Errors (%) congruent	Errors (%) incongruent	p-value	t-value		Strength of syn. touch*	Strength of syn. pain*
C1	3.33	6.73	0.36	-0.92	no effect of congruency	2	0
C2							
C3	1.64	1.92	0.90	-0.13	no effect of congruency	0	0
C4	6.67	5.77	0.82	0.23	no effect of congruency	0	0
C5	11.67	3.81	0.05	1.96	no effect of congruency	0	0
C6	1.69	20.19	0.00	-3.42	congruency effect	0	0
C7	11.67	25.44	0.03	-2.15	congruency effect	4	1
C8	0.00	5.31	0.07	-1.82	no effect of congruency	1	2
C9	0.00	0.88	0.48	-0.71	no effect of congruency	0	0
C10	0.00	3.54	0.15	-1.46	no effect of congruency	0	0
C11	10.17	5.31	0.24	1.19	no effect of congruency	0	0
C12	5.00	11.61	0.16	-1.42	no effect of congruency	0	0
C13	13.33	17.54	0.48	-0.72	no effect of congruency	2	0
C14	1.67	3.54	0.49	-0.70	no effect of congruency	0	0
C15	0.00	0.00	NaN	NaN	no effect of congruency	4	0
C16	1.75	3.54	0.52	-0.65	no effect of congruency	0	0
C17	1.72	0.00	0.17	1.39	no effect of congruency	0	0
C18	0.00	0.00	NaN	NaN	no effect of congruency	0	0

*Strength of synesthesia for observed touch and pain as experienced during the pictures experiment, as indicated by self-report in the exit questionnaire.

Table S2. Correlations between subjective ratings and physiological responses in the arousing pictures experiment.

		Condition#	Arousal	Pupil dilation (relative)	Heart rate (bpm)
Synesthetes	Valence	Negative	**-.538****	-.030	.073
		Positive	**-.185***	**-.194***	-.141
		Neutral	**-.271****	-.077	.084
		Emo_neg	**-.411****	-.070	.048
		Emo_pos	.112	-.040	-.045
	Arousal	Negative		-.048	-.136
		Positive		.007	.030
		Neutral		-.007	**-.177***
		Emo_neg		-.130	-.187
		Emo_pos		.185	-.068
	Pupil dilation (relative)	Negative			.148
		Positive			**.207****
		Neutral			.066
		Emo_neg			.137
		Emo_pos			.179
Controls	Valence	Negative	**-.197***	-.059	-.095
		Positive	**.216***	-.053	.160
		Neutral	.151	-.097	.029
		Emo_neg	-.135	-.064	-.029
		Emo_pos	.173	.020	.015
	Arousal	Negative		.061	-.015
		Positive		.043	.129
		Neutral		-.145	-.187
		Emo_neg		.007	-.083
		Emo_pos		.106	.188
	Pupil dilation (relative)	Negative			-.102
		Positive			.059
		Neutral			.035
		Emo_neg			.137
		Emo_pos			-.105

#After Bonferroni correction for multiple comparisons across all conditions and tests per group (N=30) the correlations marked with ** remain significant.

*significant at the p<.05 level **significant at the p<.01 level