

Blocking facial mimicry affects recognition of facial and body expressions

Sara Borgomaneri, Corinna Bolloni, Paola Sessa & Alessio Avenanti

Supporting Information

- S1 Text. Pilot study 1 and 2.
- S2 Text. Analysis of RTs in Experiment 1 and 2.
- S1 Table. Mean ratings of stimuli selected in pilot study 1.
- S2 Table. Mean recognition accuracy of stimuli selected in pilot study 2.

S1 Text. Pilot study 1 and 2. Description of the procedures used to select visual stimuli for Experiment 1 and 2.

Pilot study 1: selection of visual stimuli for Experiment 1. In the first pilot study, 20 participants were presented with an initial pool of 214 body expressions taken from a validated database showing emotional and neutral expressions of the bodies of male actors whose faces were blanked out [58,76]. Moreover, participants were presented with 158 facial expressions that were created by modifying stimuli from the Nimstim database [77] in order to start with a sample of facial expressions showing different degrees of emotional intensity. For each actor, we created morphing transitions between the neutral face and the final expressive face using Fantamorph software (<http://www.fantamorph.com>). Participants were presented with the 372 pictures and asked to judge the intensity of happiness and fear conveyed by each expression using a 9-point Likert scale ranging from 1 (no emotion) to 9 (maximal intensity of the emotion). To avoid building up artificial correlations between the different evaluations, each rating was collected separately during successive presentations of the whole set of stimuli. Facial and body stimuli were presented in separate blocks. Stimuli remained on the screen until the participants' response. Ratings were self-paced and we did not collect RTs. The order of ratings, blocks and stimuli was randomized. Based on the mean intensity ratings of each stimulus across the 20 participants, we selected three groups of emotional expressions for each face/body medium: happy expressions (mean ratings of happiness: $5 \geq h \geq 8$; mean ratings of fear: $f \leq 3$), fearful expressions (mean ratings of fear: $5 \geq f \geq 8$; mean ratings of happiness: $h \leq 3$) and neutral expressions (mean ratings of happiness and fear: h and $f \leq 3$). This procedure ensured that the selected expressions conveyed appropriate emotional information. In this way, we also discarded emotional stimuli with extremely high intensity ratings, to ensure recognition of the stimuli was not trivial. Then, an experimenter further selected 138 stimuli (23 stimuli for each combination of face/body medium and emotional expression), with the aim of achieving a balanced representation of 4 actors while maintaining similar ratings between facial and body stimuli. S1 Table shows subjective ratings of the final set of stimuli.

Pilot study 2: selection of visual stimuli for Experiment 2. In the second pilot study, participants were presented with an initial pool of stimuli belonging to the same databases used in Experiment 1 [58,76,77]. The initial pool comprised 138 body stimuli and 140 face stimuli, equally distributed across four stimulus categories, namely happy females, happy males, fearful females and fearful males. Participants performed two tasks presented in two separate sessions, whose order was counterbalanced across participants. In each task/session, participants were presented with 5 blocks during which the whole set of stimuli was presented on the screen in a randomized order. They were asked to report either the gender (gender discrimination task) or the emotion expressed by the model (emotion discrimination task). Each picture was presented for 500 ms and the order of the tasks and blocks was randomized. Based on participants' performance, we selected a pool of 40 faces and 40 bodies.

S2 Text. Analysis of RTs in Experiment 1 and 2. Description of the statistical analyses performed on RTs data in the main experiments.

Analysis of RTs data in Experiment 1. In Experiment 1, the Medium x Facial Manipulation x Expression ANOVA carried out on RTs showed a main effect of the factor Expression ($F_{2,46} = 4.69$; $p = 0.01$; $\eta_p^2 = 0.17$) and a Medium x Expression interaction ($F_{2,46} = 16.78$; $p < 0.001$; $\eta_p^2 = 0.42$). For facial stimuli, RTs to neutral expressions (660 ± 109 ms) were faster than RTs to happy expressions (711 ± 111 ms; $p < 0.001$; Cohen's $d = 0.80$) and fearful expressions (717 ± 115 ms; $p < 0.001$; Cohen's $d = 0.94$), which in turn did not differ from one another ($p = 0.91$). In contrast, for body stimuli, RTs to fearful expressions (673 ± 105 ms) were faster compared to happy expressions (716 ± 126 ms; $p = 0.003$; Cohen's $d = 0.82$) and neutral expressions (712 ± 127 ms; $p = 0.004$; Cohen's $d = 0.60$), which in turn did not differ from one another ($p = 0.89$). These findings can be accounted for by the visual distinctiveness of our pool of facial and body stimuli: fearful bodies were the only ones showing crouched postures and neutral faces were the only ones showing a static/relaxed face. These features could have made recognition of those stimuli faster relative to the other facial and body categories. No other main effects or interactions were significant (all $F \leq 1.81$, all $p \geq 0.17$), including the Facial manipulation x Expression interaction ($F_{4,92} = 0.28$; $p = 0.89$). Thus, the reduction in accuracy observed in Experiment 1 for happy expressions when participants bit a pen cannot be due to any speed-accuracy trade-off.

Analysis of RTs data in Experiment 2. In Experiment 2, the Medium x Task x Condition ANOVA carried out on RTs showed a significant main effect of the factor Medium ($F_{1,23} = 6.77$; $p = 0.02$; $\eta_p^2 = 0.23$), which was qualified by a Medium x Task interaction ($F_{1,23} = 53.85$; $p < 0.0001$; $\eta_p^2 = 0.70$). This interaction showed that participants were faster at discriminating emotions from bodies than from faces (551 ± 95 ms vs. 575 ± 112 ms; $p = 0.02$; Cohen's $d = 0.41$) and faster at discriminating gender from faces than from bodies (543 ± 105 ms vs. 568 ± 112 ms; $p < 0.001$; Cohen's $d = 0.50$). No other main effects or interactions approached significance (all $F \leq 3.22$, all $p \geq 0.08$), including the critical Task x

Condition interaction ($F_{3,693} = 0.19$; $p = 0.90$). Thus, the reduction in accuracy observed in Experiment 2 for happy expressions when participants bit a pen cannot be due to any speed-accuracy trade-off.

S1 Table. Mean ratings of stimuli selected in pilot study 1. Mean \pm S.D. ratings of happiness and fear reported on a 9-point Likert scale ranging from 1 (no emotion) to 9 (maximal intensity of the emotion).

		Happy expression	Neutral expression	Fearful expression
Ratings of Happiness	Face	6.2 \pm 1.1	2.2 \pm 0.6	1.3 \pm 0.3
	Body	6.2 \pm 0.7	2.5 \pm 0.9	1.4 \pm 0.2
Ratings of Fear	Face	1.3 \pm 0.2	1.6 \pm 0.5	6.2 \pm 0.7
	Body	1.4 \pm 0.3	1.6 \pm 0.3	6.2 \pm 0.8

S2 Table. Mean recognition accuracy of stimuli selected in pilot study 2. Mean \pm S.D. of emotion and gender recognition accuracy (% of correct response).

		Female models		Male models	
		Happy expression	Fearful expression	Happy expression	Fearful expression
Emotion recognition	Face	92 \pm 4	95 \pm 3	88 \pm 6	94 \pm 4
	Body	92 \pm 3	92 \pm 9	91 \pm 5	93 \pm 5
Gender recognition	Face	95 \pm 3	94 \pm 4	94 \pm 4	94 \pm 5
	Body	91 \pm 6	88 \pm 5	94 \pm 4	92 \pm 7