Pupil mimicry promotes trust through the theory-of-mind network

Prochazkova, E.; Prochazkova, L.; Giffin, M.R.; Scholte, H.S.; De Dreu, C.K.W.; Kret, M.E.

DOI
10.1073/pnas.1803916115

Publication date
2018

Document Version
Final published version

Published in
Proceedings of the National Academy of Sciences of the United States of America

Pupil mimicry promotes trust through the theory-of-mind network

Eliska Prochazkova[a,b], Luisa Prochazkova[a,b], Michael Rojek Giffin[b,c], H. Steven Scholte[d], Carsten K. W. De Dreu[b,c,e], and Mariska E. Kret[a,b,1]

[a]Cognitive Psychology Unit, Institute of Psychology, Leiden University, 2333 AK Leiden, The Netherlands; [b]Leiden Institute for Brain and Cognition, 2300 RC Leiden, The Netherlands; [c]Department of Social Psychology, Institute of Psychology, Leiden University, 2333 AK Leiden, The Netherlands; [d]Department of Psychology, University of Amsterdam, 1018 WS, Amsterdam, The Netherlands; and [e]Center for Experimental Economics and Political Decision Making, University of Amsterdam, 1001 NK Amsterdam, The Netherlands

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved June 19, 2018 (received for review March 6, 2018)

The human eye can provide powerful insights into the emotions and intentions of others; however, how pupillary changes influence observers' behavior remains largely unknown. The present fMRI–pupillometry study revealed that when the pupils of interacting partners synchronously dilate, trust is promoted, which suggests that pupil mimicry affiliates people. Here we provide evidence that pupil mimicry modulates trust decisions through the activation of the theory-of-mind network (precuneus, temporo-parietal junction, superior temporal sulcus, and medial prefrontal cortex). This network was recruited during pupil-dilation mimicry compared with interactions without mimicry or compared with pupil-constriction mimicry. Furthermore, the level of theory-of-mind engagement was proportional to individual's susceptibility to pupil-dilation mimicry. These data reveal a fundamental mechanism by which an individual's pupils trigger neurophysiological responses within an observer: when interacting partners synchronously dilate their pupils, humans come to feel reflections of the inner states of others, which fosters trust formation.

trust game | physiological linkage | neuroimaging | social cognition | affect

The propensity to trust is essential for individuals to cooperate and for societies to prosper (1, 2). Nevertheless, individuals also need to be equipped with decoding machinery in the brain, which allows them to quickly detect signals of danger (3), refrain from cooperation, and withhold trust (4). Among the many implicit cues that may inform assessments of someone's trustworthiness, the human eye region stands out as particularly salient and powerful. By contracting the muscles around their eyes and pupils, people communicate messages with affective meanings, such as friendliness or threat (5–8). Intriguingly, in our earlier research we observed that if partners' pupils synchronously dilate, trust is promoted (9–11). Apart from human adults, pupil mimicry has been reported in chimpanzees (12) and young infants (13, 14), which suggests that pupil mimicry may have evolved as a social mechanism to promote empathic bonding with kin and kith. Nevertheless, how pupil mimicry works on a mechanistic level and how it influences decisions of trust remains unclear. Revealing the mechanisms will clarify how pupil mimicry modulates brainwide neural interactions involved in trust formation.

In the literature, two core mechanisms have been proposed that facilitate pupil mimicry. One view suggests that pupil mimicry is controlled by a general "low-level" subcortical mechanism, possibly a direct amygdala–brainstem physiological response that can help people to quickly recognize socially arousing or threatening situations (15–17). In support of this hypothesis, observed pupil sizes are often processed nonconsciously (16, 18, 19), and perceived pupil dilation has been associated with increased amygdala activity (17, 19). Pupil mimicry might also recruit higher-level mechanisms (15). Previous fMRI research in humans and studies in monkeys indicate that the norepinephrine/acetylcholine systems associated with changes in own pupil size extend beyond functions exclusively mediated by the autonomic nervous system to emotion-processing areas (20, 21). Furthermore, neural regions supporting social cognition have been reported to be associated with the mimicry of affective cues (22), including subtle changes in pupil size (16, 18). The intertwined neural circuitry between social cognition and pupillary processes implies that pupil mimicry might serve a social function that extends beyond physiological responses, such as arousal (15). During pupil mimicry, the feedback from the visceral afferent fibers mapped hierarchically in the brain possibly influence cortical areas engaged in subjective feelings and social decisions (21). Another possibility therefore, is that the mimicry of pupil size shapes trust decisions via activation of the theory-of-mind (ToM) network, (expanding above subcortical circuits), which is implicated in prosocial behavior and trust formation. However, to date there is no evidence that directly investigates the engagement of either of these neural mechanisms during pupil mimicry and trust formation.

The present study investigates the neurocognitive link between pupil mimicry and trust; we performed a combined fMRI and pupillometry study during which participants made trust decisions (Fig. 1). During each trial, the pupils of virtual partners dilated, constricted, or remained static over stimulus presentation time, while subjects decided how much money they wanted to invest in their partner, whose eye region was shown. Based on our earlier research (9–11), we predicted that observed pupil dilation would increase participants' (*i*) trust and (*ii*) pupil size and that (*iii*) pupil mimicry would modulate the effect of the partner's pupil on trust. Crucially, we hypothesized that if pupil

Significance

Trusting others is central for cooperative endeavors to succeed. To decide whether to trust or not, people generally make eye contact. As pupils of interaction partners align, mimicking pupil size helps them to make well-informed trust decisions. How the brain integrates information from the partner and from their own bodily feedback to make such decisions was unknown because previous research investigated these processes separately. Herein, we take a multimethod approach and demonstrate that pupil mimicry is regulated by the theory-of-mind network, and informs decisions of trust by activating the precuneus. This evolutionary ancient neurophysiological mechanism that is active in human adults, infants, and chimpanzees promotes affiliation, bonding, and trust through mimicry.

Author contributions: C.K.W.D.D. and M.E.K. designed research; M.E.K. performed research; E.P., L.P., M.R.G., H.S.S., C.K.W.D.D., and M.E.K. contributed new reagents/analytic tools; E.P., L.P., and M.R.G. analyzed data; and E.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data have been deposited in NeuroVault, https://neurovault.org/collections/3965/.

[1]To whom correspondence should be addressed. Email: m.e.kret@fsw.leidenuniv.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1803916115/-/DCSupplemental.

Published online July 16, 2018.

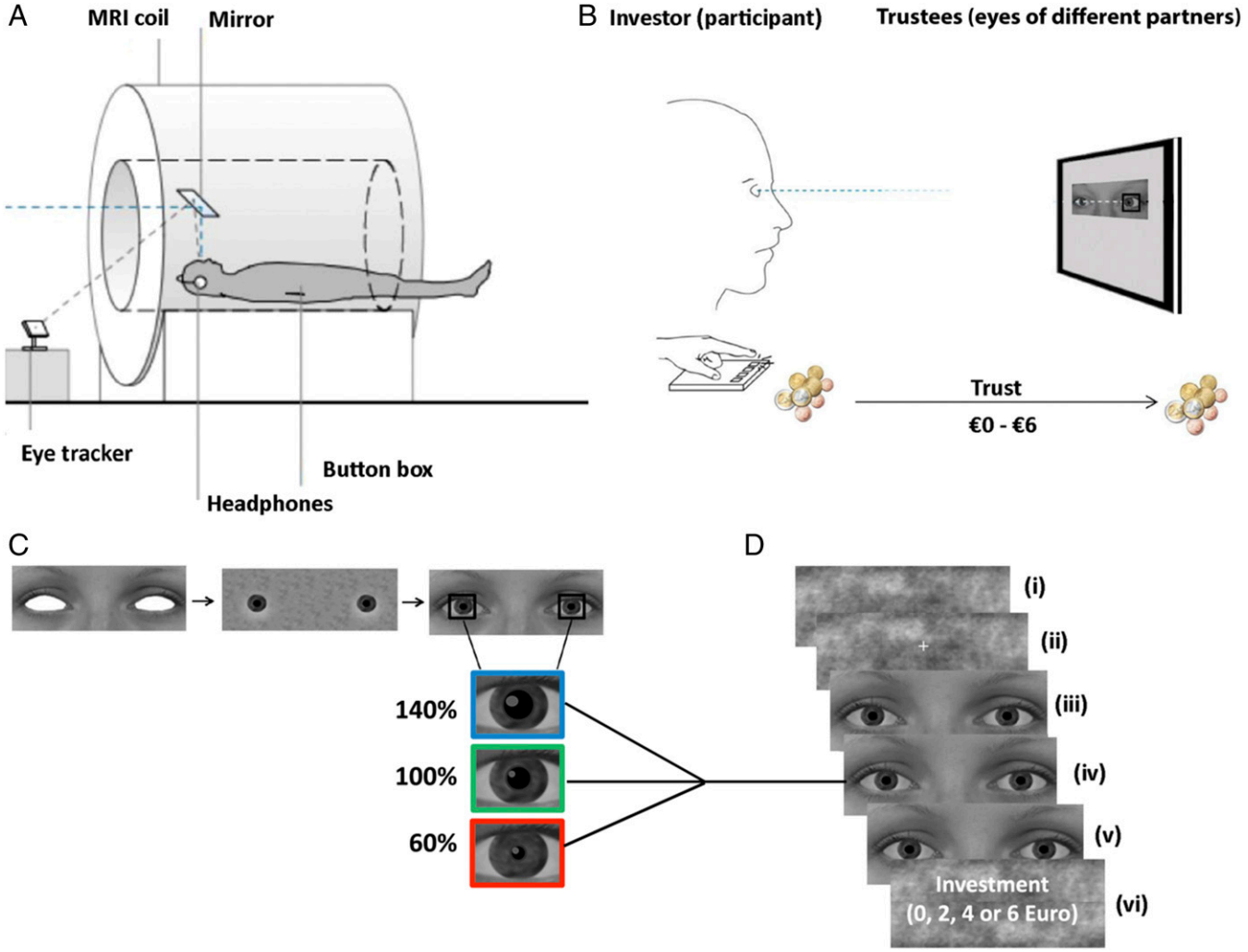

Fig. 1. Experimental set-up, stimuli, and task. (*A*) Inside the MRI scanner, the participants played one-player trust-games while their investment-decisions and pupil diameter were measured with a button box and eye-tracker, respectively. (*B*) Subjects (investors) watched short video clips showing the eye region of different virtual partners (trustees) whose pupils were manipulated to change in size. In each trial, subjects were asked to transfer between €0 and €6 to their partner. Investments were then tripled and the virtual trustee was asked to transfer between 0% and 100% of the tripled amount back to the investor. No feedback was provided so that subject's investments (indicating trust) were based on information from the partner's eye region only. (*C*) The stimulus material consisted of 18 photos with neutral expressions (nine males). The eyes were then filled with eye whites and irises, and an artificial pupil was added. The partner's pupil dilated (140% of the original diameter), constricted (60%), or remained static (range of 3–7 mm). (*D*) Stimuli presentation. (*i*) A Fourierscrambled image was presented for 4,000 ms; (*ii*) fixation followed for 500 ms; (*iii*) the eye stimulus remained static for the first 1,500 ms; then (*iv*) in the dilation and constriction conditions, the pupils gradually changed in size over 1,500 ms; and (*v*) remained static at that size during the final 1,000 ms (in the static condition, pupils remained at the same size throughout the trial). (*vi*) A screen appeared asking participants to make an investment decision.

mimicry activates a "threat-related" mechanism, it should engage the amygdala, the frontal pole, and the brainstem nuclei, which orient behavior toward basic survival needs (3). In contrast, if pupil mimicry operates a function similar to more overt emotional expressions, such as body postures or facial expressions, pupil mimicry should activate ToM areas involved in social cognition [precuneus, superior temporal sulcus (STS), temporoparietal junction (TPJ), and medial prefrontal cortex (MFPC)] (23–26). Accordingly, to disentangle these two possible neural pathways underlying pupil mimicry and trust formation, we included two independent localizer tasks to map threat-related and ToM-related neural networks and compared these to the pupil-mimicry pattern. In region-of-interest (ROI) analyses we investigated how pupil-dilation mimicry and pupil-constriction mimicry independently modulate ToM activity and tested which parts of the ToM network most closely associated with participants' level of trust.

Results

Behavioral and Pupillary Results. First, we conducted a series of multilevel models (*Methods*) to test our behavioral predictions. In the first model, we replicated previous findings (9–11), by showing that partners with dilating pupils were trusted more than partners with static pupils [$\beta = 0.19$, SE = 0.05, CI (0.08, 0.30), $P < 0.001$] and partners with constricting pupils were trusted less than partners with static pupils [$\beta = -0.28$, SE = 0.05, CI (−0.38, −0.17), $P < 0.001$], [$F(2, 5{,}933) = 37.897$, $P < 0.001$] (Fig. 2*A* and *SI Appendix*, Table S1). Second, we found support for pupil mimicry. Fig. 2*B* shows that participants' pupil sizes dilated fastest when observing partners' pupils that dilated compared with partners' pupils that constricted or remained static: linear trend × partner pupil size [$F(2, 153{,}987) = 8{,}276$, $P < 0.001$]. Specifically, during trials where partners' pupils dilated, participants' pupils dilated faster compared with trials when partners' pupils remained static [$\beta = -0.55$, SE = 0.02, CI (−0.01, −0.02), $P = 0.005$] or constricted [$\beta = -0.77$, SE = 0.02,

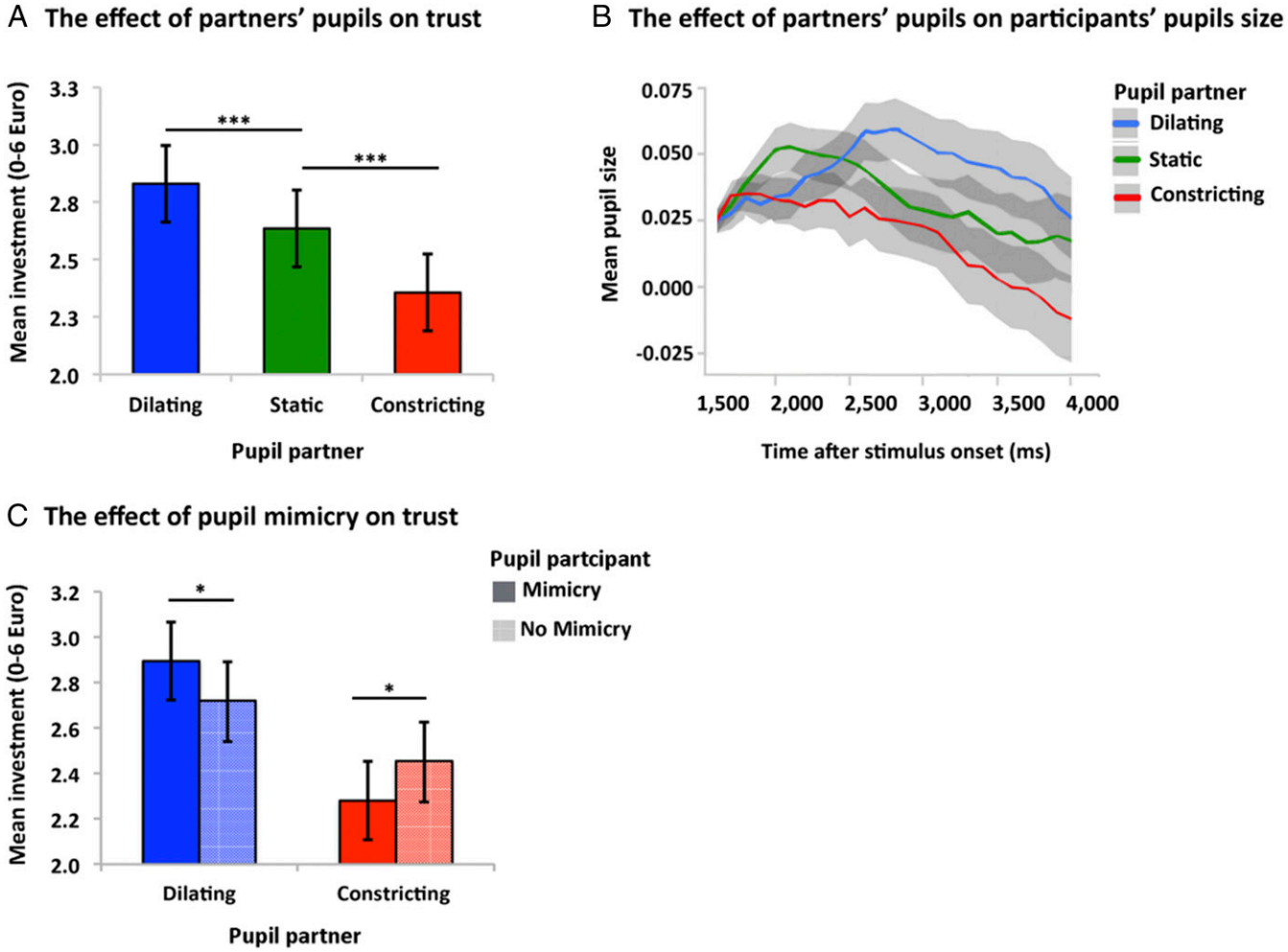

Fig. 2. Behavioral and pupillometry results. (*A*) The bar plot shows that mean trust-related investments (€) increased in response to partners' dilating pupil size (n = 40 participants). Error bars indicate ±1 SE. ***P < 0.001 for each factor, pairwise contrasts: dilating pupils vs. static pupils [B = 0.19, CI (0.08, 0.30)] and constricting pupils vs. static pupils [B = −0.28, CI (−0.38, −0.17)]. (*B*) Participants mimicked partner's pupil sizes: the curves correspond to participants' mean pupil response from baseline over the remaining of stimulus presentation time (ms), in response to partner's dilating, static, and constricting pupils. Mean pupil size is depicted in arbitrary values. Shaded areas indicate the 99% CI. (*C*) The bar plot shows mean investments (€) as a function of partners' and participants' pupil size. Error bars indicate ±1 SE. *P < 0.01. Mean investment increases when a participant's own pupils dilate in response to their partner's dilating pupils. Pairwise contrast: pupil-dilation mimicry vs. no pupil-dilation mimicry [B = 0.175, CI (0.02, 0.33)]. Mean investment decreases when participants' pupil constricts in response to their partners' constricting pupils. Pairwise contrast: pupil-constriction mimicry vs. no pupil-constriction mimicry [B = −0.173, CI (−0.33, −0.02)].

CI (−0.12, −0.04), P < 0.001] (*SI Appendix*, Table S2). Third, consistent with prior evidence (9, 11), pupil mimicry modulated trust. Specifically, the interaction between partner pupil size × participant's own pupil size had a significant effect on trust [F (2, 5,750) = 5.847, P = 0.003]. Pair-wise post hoc comparisons confirmed that when participants mimicked dilating pupils, they trusted their partner more compared with when they did not mimic their partner [β = 0.175, SE = 0.08, CI (0.02, 0.33), P = 0.027]. Conversely, pupil-constriction mimicry decreased trust levels compared with when constricting pupils were not mimicked [β = −0.173, SE = 0.08, CI (−0.33, −0.02), P = 0.028] (10, 11). Importantly, there was no significant difference in trust when participants' own pupils dilated compared with constricted during trials where partners' pupils remained static (P > 0.05) (Fig. 2*C* and *SI Appendix*, Table S3). Collectively, these behavioral analyses demonstrate that pupil mimicry enhances the effect that partners' pupils have on trust and support the notion that trust decisions are evaluated through integrating information from partners' pupils combined with own pupillary responses.

fMRI Results

Neural Correlates of Pupil Mimicry. Having established that the mimicry of subtle affective cues, such as pupil size, influences in part subjective evaluations of others' trustworthiness, we set out to investigate the neural regions that play a role in pupil mimicry, using the general linear model (GLM). The neural data were extracted 3,000 ms after the onset of the stimuli. This was the time point at which partners' pupils were maximally dilated or constricted (unless they had remained static) and participants' own pupils had had sufficient time to adjust to the presentation of the stimulus (9–11). The aim of the first analysis was to detect regions that are highly active during pupil mimicry. To test this, partners' and participants' pupillary responses were used as explanatory variables, resulting in the following conditions: pupil-dilation mimicry, pupil-constriction mimicry, no pupil-dilation mimicry, and no pupil-constriction mimicry (*Methods*). Of key interest was the whole-brain contrast comparing mimicry versus no mimicry trials.

Consistent with our hypothesis, the results showed that during pupil mimicry, participants displayed enhanced activation in all key regions of the ToM network: bilateral TPJ [60, −54, 18/−58, −54,

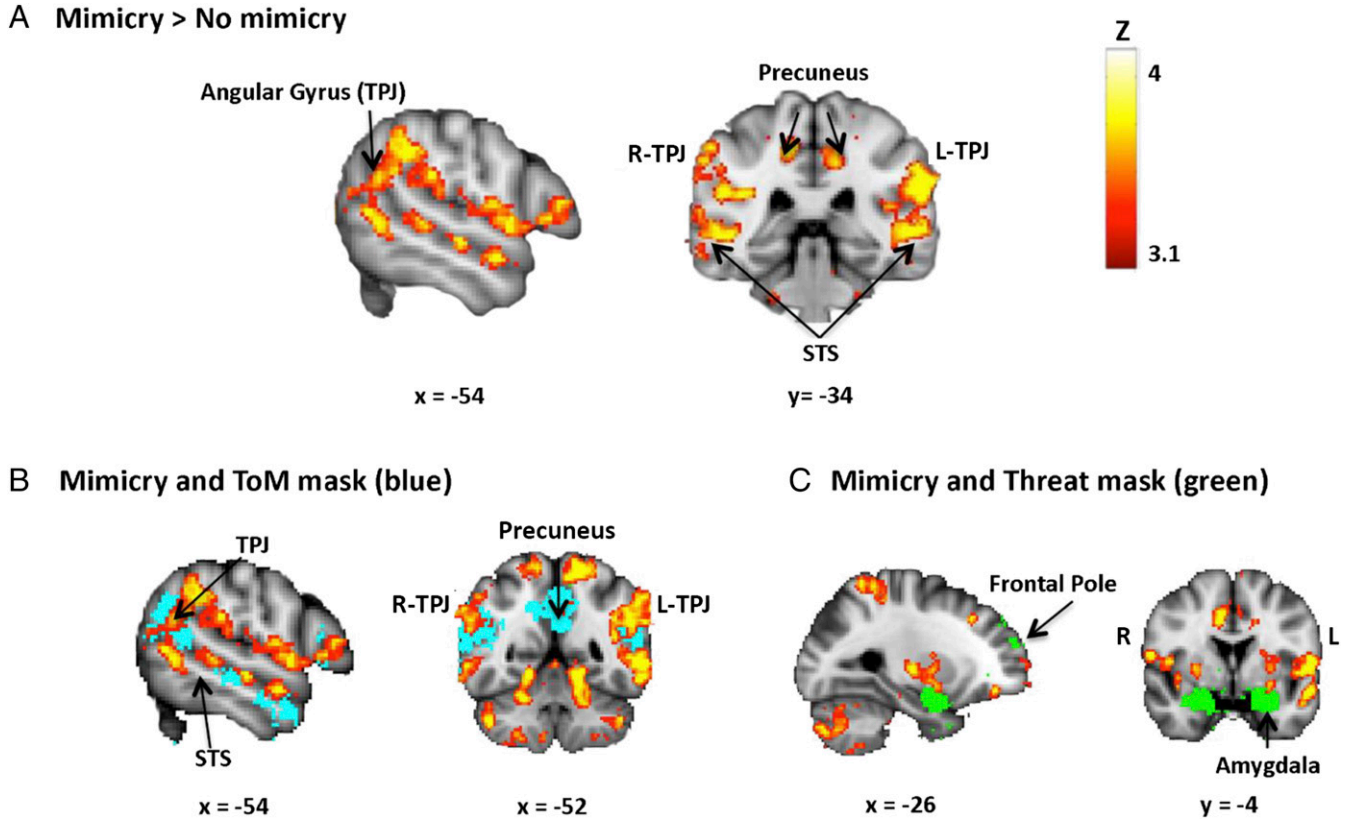

Fig. 3. Neural correlates of pupil mimicry. (*A*) During mimicry, subjects displayed enhanced ToM activation. Peak voxels MNI *x*, *y*, *z* coordinates TPJ [60, −54, 18/−58, −54, 18], bilateral STS [52, −34, 2/−52, −34, 2], right MPFC [6, 46, 8] (not displayed in the image), precuneus cortex [8, −40, 48/−8, −40, 48]; threshold at $P < 0.05$ [cluster-level FWE correction with multiple comparisons at 2.3, ($n = 34$ participants)]. For visualization, the threshold was set at $z = 3.1$–4. (*B* and *C*) The image shows additional overlaps between pupil-mimicry pattern and ToM network (blue) and threat network (green). The background images reflect MNI 2-mm template (0.05-voxel size smoothing kernel); the right side of the image corresponds to the left side of the brain. Location coordinates are in stereotactic MNI space with a 2 × 2 × 2-voxel size. The source of anatomical labels is FSL Atlas tools.

18], bilateral STS [52, −34, 2/−52, −34, 2], right MPFC [6, 46, 8] and bilateral precuneus cortex [8, −40, 48/−8, −40, 48] (Fig. 3*A* and *SI Appendix*, Table S4). Overlapping activation patterns during constriction mimicry and dilation mimicry were observed in the right lateral occipital cortex and in the precentral gyrus (*SI Appendix*, Tables S5 and S6). For a closer examination of the pupil mimicry fMRI pattern, Fig. 3 *B* and *C* depicts the neural overlap between the pupil mimicry-activation and ToM- and threat-related brain-activation masks, which we obtained by conducting a meta-analysis via Neurosynth (27) (*SI Appendix*, Table S7). Specifically, the masks are derived from a metaanalysis we conducted on previous studies displaying brain regions that are consistently active in studies that include the name "theory of mind" or "threat" in the abstract ($n = 140$ and $n = 170$, respectively). These neural overlaps clearly indicate that pupil mimicry extends beyond the threat-related areas to neocortical regions involved in "mindreading" (right temporoparietal junction) (25) and social judgment formation (medial prefrontal cortex) (28). These results provide supporting evidence for the relationship between pupil mimicry and higher-level ToM processes.

The whole-brain analysis confirmed that the brain decodes pupil size in a similar manner as more overt facial expressions of emotion (15, 21, 29, 30). Assuming that morphological expressions of primates evolved as biological adaptations to transfer social information, changes in pupil size are likely used as social cues by observers. What remains unknown is whether—and to what extent—pupil mimicry is required for the brain to detect pupillary cues as socially relevant. To answer this question, one alternative needed to be ruled out. Participants' pupils may employ ToM activation, regardless of mimicry. For example, observed partners' pupillary changes may result in a similar ToM activation pattern as seen during pupil mimicry. If true, this would suggest that pupil mimicry is not a prerequisite for enhanced level of ToM activity. We have ruled this out in a control analysis. In the control analysis we compared participants' neural activity when they saw partners' pupils dilate or constrict as opposed to staying static. The control analysis showed that without accounting for mimicry, the change in partners' pupil sizes (both dilation and constriction) was associated with enhanced activity in brain areas known to be involved in biological motion [movement of the eyes, mouth, or hand (31)] and face processing (32), but not in ToM processes (*SI Appendix*, Fig. S1 and Tables S8–S10). This shows that changes in a partner's pupil size, as a subtle form of autonomic expression, do not directly govern neural regions involved in implicit social evaluations. Instead, pupil mimicry is conditional for the engagement of social networks.

ToM ROI Selection. For a more quantitative examination of the effect of pupil mimicry on ToM activation, we incorporated a well-established independent ToM localizer task into our study's design (25, 33, 34). The task consisted of 20 stories of two different types presented in two different blocks (for examples, see *SI Appendix*, Table S12). This functional localizer helped us to identify brain regions involved in ToM in individual participants. One subject was excluded from the ROI analysis due to excessive head motion during the ToM localizer task ($n = 33$ participants). After we defined ToM regions in the individual space (with the use of the ToM localizer), we standardized each participant's functional ToM image by subtracting the dichotomized masks from the average activation ToM mask that we obtained by conducting a

metaanalysis on previous ToM/fMRI studies via Neurosynth (27) (*SI Appendix*, Table S7). Therefore, the final standardized ToM masks included only those voxels that were activated in the subjects of the present study as well as in the subjects that took part in the previous studies that were included in our metaanalysis (see Fig. 4*A* for an example of one subject's ToM mask).

Pupil Mimicry and ToM Activation. Growing evidence suggests that social signals, such as emotional expressions and gaze direction, are automatically encoded in social brain networks (30, 35). Here we build upon these findings by investigating whether pupil mimicry modulates ToM activity, which further impacts on trust. To test this hypothesis, we extracted the parameter estimates of the neural activation from the individualized ToM masks and averaged ToM activation across subjects. A one-way repeated-measures ANOVA revealed that there was a significant difference in ToM activation across our four experimental conditions [$F(1, 33) = 9.821$, $P = 0.004$] (Fig. 4*B*). As expected, the activation was higher when subjects mimicked partners' pupil size compared with when they did not. Interestingly, this was only the case when subjects mimicked dilating pupils. Follow-up pairwise comparison tests (with Bonferroni correction) revealed that the mimicry of dilating pupils was associated with a significantly greater activation in ToM regions compared with all of the other conditions, including pupil-constriction mimicry [mean difference = −0.098, SE = 0.03, CI (0.02, 0.18), $P = 0.005$]. This result suggests that the increase of activity in the ToM network observed on the whole-brain level (Fig. 3) was driven mainly by the mimicry of dilating pupils. These results imply that the ToM activation might be selectively sensitive to a partner's pupil dilation compared with pupil constriction.

One could think of this as supporting the view that pupil dilation correlates with physiological arousal (36). Considering that pupil dilation is paired with norepinephrine release, accompanied by heightened activity in other brainstem areas (20), it is possible that a participant's own pupil dilation explained the heightened activity in ToM areas, regardless of whether participants mimicked their partner or not. However, our analysis ruled out this alternative interpretation. Instead, we show that during trials where participants' own pupils dilated but their partners' constricted, ToM activity was significantly lower compared with when participants' mimicked partners' dilating pupils [mean difference = −0.113, SE = 0.03, CI (0.03, 0.19), $P = 0.005$]. Together, these results imply that the mirroring response is conditional for pupil dilation to activate ToM regions.

The second potential issue is regarding individuals' neural differences in social processing. That is, although ToM has been identified as a key system underlying social cognition (23–25),

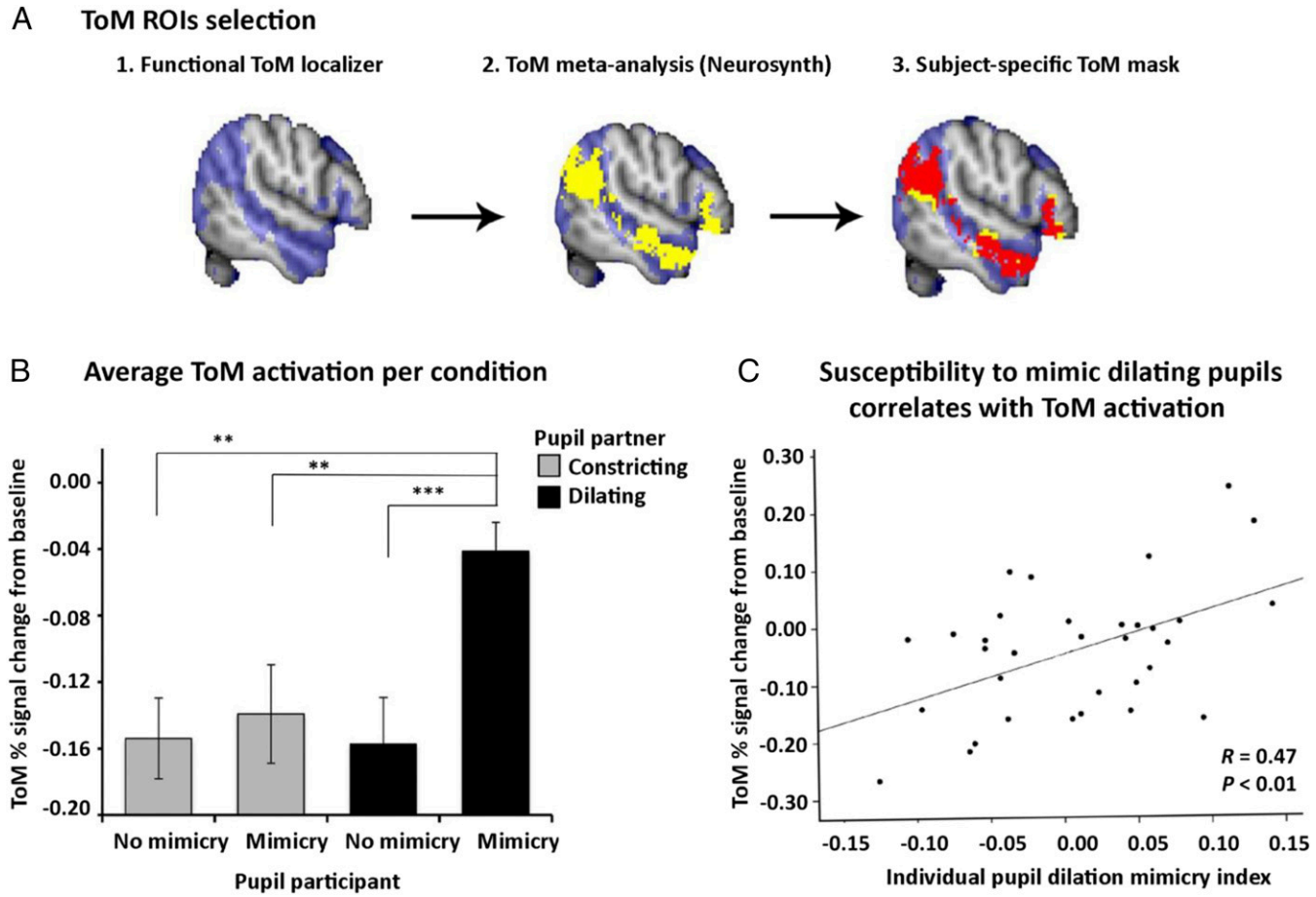

Fig. 4. ROI analyses. (*A*) Example of one subject's ToM network mask selection. The ROIs were classified by an independent TOM localizer (blue) as well as by the additional inclusion of ToM masks (yellow), derived from our metaanalyses on previous studies. The overlapping voxels were used as the final mask (red). (*B*) The bar plot displays the mean parameter estimates averaged across all subjects ($n = 33$ participants) of the neural activation extracted from the individual ToM masks during four experimental conditions. During pupil-dilation mimicry the ToM percentage signal increase was significantly greater compared with all of the other conditions, including pupil-constriction mimicry [mean difference = −0.098, SE = 0.03, CI (0.02, 0.18), $P = 0.005$], no-constricting mimicry [mean difference = −0.113, SE = 0.03, CI (0.03, 0.19), $P = 0.005$], and no-dilation mimicry conditions [mean difference = −0.116, SE = 0.02, CI (0.05, 0.18), $P < 0.001$]. $^{**}P < 0.01$, $^{***}P < 0.001$. Error bars indicate ±1 SE. (*C*) The scatter plot shows that the same subjects that displayed larger pupil-dilation mimicry index also displayed the greatest increase in ToM percentage signal during pupil-dilation mimicry ($R = 0.47$, $P < 0.01$).

whether all of our subjects' engaged ToM areas during mentalizing about other's intentions is unclear. Therefore, we conducted an additional ROI analysis, indeed showing that pupil mimicry with dilating and constricting pupils modulates subject-specific ToM areas. Finally, a ROI analysis was conducted to investigate whether these social areas relate to trust. We elaborate on these analyses in the sections below.

Individual Differences in Pupil-Dilation Mimicry Correlate with ToM Activation. Given that the group analysis revealed that the mimicry of a partner's dilating pupils is associated with greater ToM activation, we next determined whether a similar relationship was evident across all individuals. That is, we investigated whether an individual's susceptibility to mimic their partner's dilating pupils correlated with the average ToM percentage signal increase during pupil-dilation trials. We subtracted each subject's average pupil size on trials showing a partner with dilating pupils from his/her mean pupil size during trials where their partner's pupils remained static (Fig. 4*C*). This difference represented an individual index of pupil-dilation mimicry susceptibility. As expected, the result shows a positive relationship between the pupil-dilation mimicry index and increases in the ToM signal during those trials (Pearson's $R = 0.473$, $P = 0.005$). Importantly, there was no direct association found between an individual's average pupil size (regardless of partner's pupil size) and activity in this network (Pearson's $R = 0.029$, $P = 0.870$). This suggests that the susceptibility to mimic a partner's dilating pupils discriminates people on the basis of their social network engagement.

Trust and ToM Network. The final goal of the fMRI experiment was to determine which neural mechanisms were engaged in the pupil mimicry–trust linkage. To that extent, we investigated whether ToM regions that were associated with pupil mimicry were also modulated by trust decisions. To test this prediction, on each trial we used the participant's level of investment as a regression parameter convolved with the hemodynamic response function to identify the ToM voxels that were most closely correlated with trust (the level of the investment). The higher-level analysis and group-level analysis were performed by averaging the mean activation within and between subjects, without any additional contrasts. We compared the ToM signal against baseline and tested for significance with permutation testing (*Methods*) (36). As predicted, the results revealed that the level of trust modulated ToM activation, with peak activation in the precuneus [−2, −68, 38] threshold-free cluster enhancement $P = 0.04$ (Fig. 5), confirming the involvement of the ToM network in the development of trust. This analysis supported the hypothesis that a partner's pupil dilation drives trust through pupil mimicry and associated neural activation in brain areas related to social cognition.

Discussion

In this study, the combination of psychophysiological, behavioral, and neuroimaging data allowed us to disentangle the elusive link between pupil mimicry and trust. Behaviorally, we replicated previous findings by showing that trust is increased when looking into the eyes of a partner with dilating pupils and lowered when looking into the eyes of a partner with constricting pupils (9–11). We further demonstrated that peoples' pupil sizes mimic those of observed partners and if pupil-dilation mimicry occurs, trust is promoted (9–11). These unprecedented findings already suggested a fundamental link between autonomic pupil mimicry and social cohesion. However, whereas pupil mimicry has now been well-documented (9–14, 16, 18, 19), its function and underlying mechanisms remained largely unknown (15). To test which neural organization facilitates this autonomic form of mimicry, we compared participants' neural activity when they mimicked versus did not mimic their partner's pupil size. A whole-brain analysis revealed that the mimicry of pupillary changes was associated with increased activation in the precuneus, TPJ, STS, and MPFC, all of which are key regions of the ToM network (24, 25). Neuroimaging research in humans has identified the ToM network as the basic system that facilitates social understanding (23–25).

ToM areas can be well dissociated from similar, but not entirely overlapping areas involved in empathy (37). Kanske et al. (37) define empathy and ToM as being "affective" and "cognitive" routes to understanding others. The central distinction between empathy and ToM is that empathy refers to the sharing of a sensory, affective, or bodily state (38), while ToM involves both, affective states and the cognitive reasoning about others. On the neural level, empathy and ToM networks largely overlap (e.g., in the precuneus, STS, left TPJ) (36). Nevertheless, empathy is more closely related to activity in the anterior insula and middle anterior cingulate cortex (39, 40), while ToM is associated with activity in the TPJ (37).

It is important to note that areas involved in ToM appear often in the literature on mirror neurons or more broadly on motor theories of social cognition (41, 42). For example, the

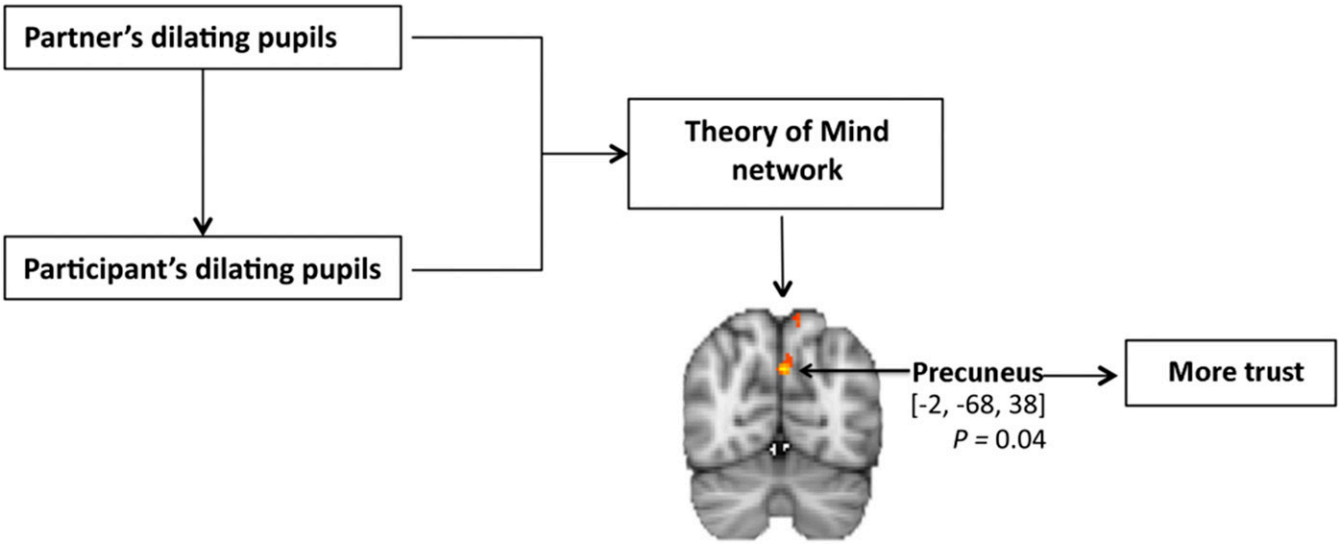

Fig. 5. ToM and trust. The diagram shows that the ToM network is modulated by pupil-dilation mimicry. Within the ToM network, investment rates predicted precuneus BOLD signal changes, confirming ToM involvement in trust decisions. Peak voxel MNI *x*, *y*, *z* coordinates [−2, −68, 38], corrected for ToM network with threshold free cluster enhancement (42) (threshold $z = 3.1$, $P < 0.05$, FWE-corrected *P* value from the minimum voxels in the cluster).

inferior frontal gyrus, parietal areas, and STS are active during grasping and the observation of grasping (43). The current fMRI findings further contribute to this body of research by showing that pupil mimicry induces a neural pattern similar to those found during more explicit forms of mirroring (e.g., motor movement). According to its location in the brain, this evidence fosters the view that mimicry, even on a subtle autonomic level, may enable a route to interpret others' behavior. An additional control analysis revealed that the ToM areas were active only when participants mimicked partners' pupil sizes, but not in response to partners' pupillary changes alone. This finding supports the hypothesis that pupil mimicry is involved in higher-level, social functions as opposed to being a lower-level mechanism restricted to subcortical structures. Although we did not find strong evidence for the involvement of subcortical structures during mimicry, this does not mean that these structures are not involved at all (16). Instead, the current data provide supporting evidence for the view that pupil mimicry extends beyond physiological responses, such as arousal (15, 16, 21), as it also triggers higher neural areas involved in social processing.

To provide more direct evidence for the "social" hypothesis of pupil mimicry, we have built upon previous findings by including a ToM localizer task. This allowed us to map subject-specific ToM regions involved in social processing (the ability to attribute mental states to others). Within subject-specific ToM masks, we examined how pupillary responses to partners' pupils modulated these neural areas. Intriguingly, our ROI analysis demonstrated that the ToM network was significantly more active during pupil-dilation mimicry, compared with pupil-constriction mimicry. This effect was also evident on the individual level, whereby individuals' susceptibility to mimic dilating pupils was proportional to participants' ToM level of engagement. The fact that the ToM network is recruited significantly more during pupil-dilation mimicry compared with pupil-constriction mimicry suggests that the mimicry of dilating pupils is more socially relevant, at least in a relatively nonthreatening interaction, as in the current study. According to Tylén et al. (44), the ToM network activation represents an adaptive neural system for rapid alerting in response to mutual social interest. If we consider that pupil dilation is tied to sympathetic nervous system activation, another person's pupil dilation as a form of affective expression is likely to trigger a spontaneous attribution of mental states (e.g., "interest in me") in the observer's brain. When a partner's pupils constrict, an observer can infer that the partner is not—or no longer—motivated to pursue social exchange and that such mental state attribution may be reduced or be absent during pupil-constriction mimicry. From this perspective, it could be argued that social brain activation during pupil dilation is more socially/evolutionary relevant. Importantly, our data demonstrate that ToM activation could not be fully explained by participants' or partners' pupil dilation alone. Instead, the mirroring response was conditional for pupillary cues to become reflected in a ToM signal (Fig. 4). Taken together, these fMRI data illustrate that only seeing a partner's pupils dilate, as a sign of arousal, is not sufficient for the brain to recognize this information as socially meaningful. In other words, there is need for mimicry—the autonomic alignment between interacting partners—to render another's autonomic expression as socially relevant.

Our findings go beyond previous studies (16, 19) by linking the engagement of neural mechanisms during pupil mimicry to social behavior. Here we demonstrate that within the ToM network, activation within the precuneus significantly correlated with participants' level of trust. The precuneus involvement in trust has been documented in previous studies (45, 46), yet the present study reveals that a partner's pupils' dilation drives trust through pupil mimicry and associated neural activation in brain areas related to social cognition. In parallel to this evidence, the empirical literature has implied that others' actions can be decoded by activating our own somatic and autonomic systems (47). For example, Harrison et al. (18) found that individual sensitivity to another's pupil size predicted scores of emotional empathy. Furthermore, previous behavioral research has shown that pupil mimicry occurs within two members of the same species (human–human and chimpanzee–chimpanzee) but not across species (human–chimpanzee) (12). Kret et al. (9, 11) found that the pupil-dilation mimicry–trust linkage is bound to interactions between members of the same ethnical group and breaks in cross-ethnical group interactions. These studies are in line with other work showing increased trust with partners that are more familiar compared with partners who are unfamiliar, creating an in-group bias (48). Although we did not manipulate familiarity in the present study, the neural mechanism that we observed during pupil mimicry, especially the activations in the temporal areas, suggest this factor might be of importance (49, 50). Future studies might therefore want to investigate whether pupil mimicry is strengthened between closely bonded partners, such as parents and their children (51). Such evidence would further support our view that pupil mimicry is a social phenomenon, which possibly evolved in and because of group life.

To conclude, by examining the neural mechanisms of pupil mimicry in the context of an economic game, the present study provides support for the social hypothesis of pupil mimicry. We demonstrated that the neural regions involved in social decision-making are modulated by the subtle expression of pupil size and that mimicry is the target mechanism underlying this process. This is important because by knowing that pupil mimicry is involved in healthy social cognition, these data reveal a fundamental mechanism by which an individual's pupils trigger neurophysiological responses within an observer. We propose that pupil-dilation mimicry seems to bring interacting partners' neural activity into mutual alignment, creating a joint-pupillary state that may facilitate communicative success. In the future, pupil mimicry might be an especially useful measure for early social deficits because autonomic cues are not likely to be influenced by learning, social norms, or conscious control compared with facial expressions and other overt affective signals. Given that fMRI measures are of a correlational nature, further research using real-life interactions and pupil-mimicry manipulations will be highly valuable to determine the putative causal link between pupil mimicry and trust formation.

Methods

Participants. Forty-one healthy, right-handed, Dutch participants without a neurological or psychiatric history and normal or corrected-to-normal vision were recruited for the present experiment. One participant had symptoms of mild depression. We have excluded this participant following our previous work, which suggested that depressed patients process observed pupils differently than healthy controls (11). With excluding this subject, we had a total of 40 participants for behavioral analysis [mean age (±SD) 23.40 (±2.91) y, 21 females, range: 19.5–32.7]. Six participants (three males and three females) were excluded from the fMRI data analysis due to excessive head movements (more than 1.5-mm displacement), leaving 34 subjects for the fMRI data analyses [mean age (±SD) 23.5 (±2.78) y, 18 females, range: 19.5–32.7]. For two participants, activation was averaged over two instead of three runs because of an insufficient number of eye-tracking data to measure mimicry. Our sample size was motivated by those used in previous studies (9, 16, 19). The experimental procedures were in accordance with the Declaration of Helsinki and approved by the Ethical Committee of the Faculty of Behavioral and Social Sciences of the University of Amsterdam. All participants provided informed consent.

Stimuli. The stimulus material consisted of nine female and nine male photos with neutral expressions derived from the validated Amsterdam Dynamic Facial Expression Set (52). Pictures were standardized in Adobe Photoshop (Adobe Systems), converted to gray scale, and cropped to reveal only the eye region. Average luminance and contrast were calculated for each picture and then adjusted to the mean. The eyes were then filled with new eye whites and irises, and an artificial pupil was added in Adobe After Effects. After a static presentation of 1,500 ms, the partner's pupil increased (140% of the original size), decreased (60% of the original size), or remained static within the physiological range of 3–7 mm. In all stimuli, the pupils were 5 mm during the first 1,500 ms and then dilated to 7 mm, constricted to 3 mm, or stayed the same (static: 5 mm). In the last second of the stimulus presentation, the pupils were static again. This way, 54 unique stimuli (3 pupil types × 18 eye regions) were created. In addition, in Matlab R2013b, Fourier-scrambled images were created from the first frame of each video. These images contained the same low-level features including contrast and

luminance of the original ones and were presented before the stimulus to reduce the light reflex. Stimuli were viewed on a back-projection screen via a mirror system attached to the MRI head coil.

Trust-Game Task. The trust game was first practiced outside of the scanner. When participants correctly answered three practice questions, we moved on to the real experiment. The trust-game experiment used a randomized event-related design. In each of the three runs, all 54 videos were presented in random order. The pupils inside the eyes of the virtual partners dilated, constricted, or remained static over stimulus presentation time; these were the three experimental conditions. A scrambled picture appeared for 4,000 ms and then a fixation cross was presented on top for 500 ms, after which a video showing eyes appeared. One video showed one eye pair with dilating, constricting, or static pupils. After observing each stimulus, the participant was prompted to make an "investment decision (€0 or €6)." The participants then had 2,000 ms to choose 0, 2, 4, or 6 euros; no feedback was provided. The intertrial interval (lasting between 9,300–12,300 ms) was sufficient for the hemodynamic response to return to the baseline.

Localizer Tasks. Two localizer tasks were performed using a randomized block design to map ToM and threat-related networks. Both localizers lasted 8.6 min and their order was counterbalanced across participants.

The ToM localizer task was taken from a widely used task to identify brain regions involved in social cognition (for more details, see refs. 33, 34, and 53). The task consisted of 20 stories, in which 10 of them described a situation in which someone held a false belief (for examples, see *SI Appendix*, Table S12). Participants had to indicate whether the suggestions referring to the stories were true or false. The other 10 stories were false-photograph stories, which described situations with a false or outdated representation of the world. The false-belief localizer was presented in a block design and counterbalanced, starting with either the false-belief or the false-photograph story. These types of stories required the participant to deal with incorrect representations about the world and were therefore matched in their difficulty, logical complexity, and inhibitory demands, but differed in the need to think about someone's thoughts. Crucially, they differ in building a representation of someone else's mental state. The amount of words was matched over the two conditions. One story started with a 12-s fixation cross, followed by a 10-s story. After the presentation of the story, the participant had 4 s to decide whether the story was true or false.

The threat-localizer task was designed to be structurally as similar as possible to the ToM localizer and also included 20 stories; however, it presented 10 threatening versus 10 neutral sentences (for examples, see *SI Appendix*, Table S12). The threat localizer was presented in a block design counterbalanced starting with either the threatening or the neutral story. One story started with a 12-s fixation cross, followed by a 10-s long story. After presentation of the story, the participant had 4 s to make the decision whether the story was true or false. The threat and ToM localizers were matched in terms of the number of words they contained. The threat localizer has been validated before usage so that only stories that were very threatening or very neutral were included. The stories were rated for threat sensation on a scale from 0 to 10 (0 being nonthreatening, 10 being very threatening) and selected by 14 people out of a list of 15 threating and 15 nonthreatening stories. Furthermore, only situations that participants could imagine or were rated as probable were selected. The 10 most threatening stories had an average threatening value of 8.75 $\pm$ 0.73 with a probability of 6.04 $\pm$ 0.083. The 10 nonthreatening stories had an average of 0.86 $\pm$ 0.68, with a probability of 6.94 $\pm$ 0.57.

Procedure. The participants were instructed about the procedure, practiced the trust game, and completed the medical screening 2 d before scanning. Participants filled out a series of questionnaires as a control to ascertain that our sample did not deviate from the normal population on following scales: Interpersonal Reactivity Index measuring empathy (with empathic concern and perspective taking scales) (54) and the Liebowitz Social Anxiety Scale, to test social anxiety disorders (55) (*SI Appendix*, Table S11). After participants signed the informed consent, they were reminded about the rules of the games while inside a 3.0-T MRI scanner. Next, two electrodes were attached to participants' left ring and index fingers, assessing skin conductance. After entering the room, a pulse oxidation signal was recorded from the middle finger. Breathing rate was measured with a band around the participants' chest (Philips Achieva). Participants were instructed to watch short video clips showing the eye region of different partners and decide how much money they would want to invest in the partner of whom the eye region was shown. Presentation 16.4 was used to present stimuli and acquire behavioral responses. Participants viewed stimuli on the projector screen over a mirror, which was mounted on the MRI head coil. They responded via a button box held in the right hand. First, a sham scan was implemented to ensure that the magnetic field was homogeneous. Subsequently, we obtained the T1 anatomical scan during which the participants performed a nine-point calibration of the eye-tracking system. Between the runs, two localizer tasks were performed to map ToM and threat-related networks (*SI Appendix*, Table 12). The scan settings were the same as for the trust-game task (see *fMRI Data Acquisition*, below). The total scanning session lasted between 60 and 80 min. After the scan session, participants filled out the State-Trait Anxiety Inventory for mental disorders (56), rated the eyes they had seen in the scanner on attractiveness, trustworthiness, and arousal and performed the reading the mind in the eyes test (57), and were instructed to draw pupils in a happy and angry face (5). Two weeks after the scanning session, participants received the Beck Depression Inventory (58).

Eye-Tracking Data Acquisition. Pupil data acquisition was collected concurrently with the fMRI measurements and sampled at 1,000 Hz, with an average spatial resolution of 15- to 30-min arc. The MRI-compatible EyeLink 1000 Long Range Mount system (SR Research) was placed outside the scanner bore and subjects' pupils were tracked via the mirror attached to the head coil. For optimal measuring with the eye-tracker, participants did not wear eye make-up. The eye-tracker was calibrated before the start of each run. Pupil preprocessing was done in five steps. (*i*) Each participant's pupillary response were measured on a trial-by-trial basis; if a participant's pupil sizes across two time-samples exceeded 2 SD, the data were identified as outliers and removed from the analysis. (*ii*) The gaps smaller than 250 ms were interpolated. (*iii*) We smoothed the data with a 10th-order low-pass Butterworth filter. (*iv*) The average pupil size 500-ms prior to when the partner's pupils began to change (i.e., 1,000–1,500 ms after stimulus onset) served as the baseline and was subtracted from the pupil size during the remaining stimulus presentation (1,500–4,000 ms). Only the final 2.5 s of stimulus presentation was included in the analysis, as from that point on partners' pupils started to change in size (9–11). (*v*) Analyses of pupil-related measures included those participants who had less than 50% signal loss during less than half of the trials.

Behavioral and Pupil Analysis. Because of the nested structure of the data, multilevel modeling was the most appropriate method to analyze the data (9–11, 59). All behavioral and pupillary data were analyzed using generalized mixed multilevel models in IBM SPSS Statistics (v20). This allowed for the estimation of individual differences by modeling random slopes and intercepts. The multilevel structure was defined by trials (level 1), nested in runs (level 2), nested in participants (level 3). As is common, nonsignificant factors were dropped one by one, starting with the higher-order interactions. Via log-likelihood tests, we determined whether dropping nonsignificant factors improved model fit or significantly worsened it, in which case the nonsignificant factor was kept. After specifying the fixed effects, model building proceeded with statistical tests of the variances of the random effects.

Trust Investment Decisions. Trusting behavior was analyzed with a series of two-level models defined by the different trials that were nested in runs and within participants. To test the effect of partners' pupils on trust, the partners' pupil size coded as −1 (constrict), 0 (static), and 1 (dilate) was used as a fixed factor, with investment level being the dependent variable.

Pupil Mimicry. To investigate the effect of partners' pupils on the participant's own pupils, we kept the three-level structure but this time we added time (100-ms time slots) as a repeated factor with a first-order autoregressive covariance structure (AR1) to control for autocorrelation. The factors partners' pupil sizes in the three conditions—constricting, static, dilating (coded as −1, 0, and 1)—served as predictors. The participants' own baseline-corrected pupil sizes were used as target variables. Furthermore, three orthogonal polynomials were included to account for linear, quadratic, and cubic trends in the growth curves. A random intercept and random linear, quadratic, and cubic terms accounted for individual differences in the pupil response.

Defining Pupil Mimicry. To further investigate the source of trust, participants' pupil responses were separated based on a median split into dilation-mimicry trials or constriction-mimicry trials. For example, a trial was categorized as a "dilation-mimicry trial" or a "constriction-mimicry trial" when the mean pupil size of the trial was higher than the median pupil size of a participant when viewing a partner with dilating pupils or when viewing a partner with constricting pupils. This way, we had an approximately equal number of trials in each condition: (*i*) dilation mimicry, (*ii*) constriction mimicry, (*iii*) no dilation mimicry, (*iv*) no constriction mimicry, and (*v*) static pupils.

Pupil Mimicry and Trust. To test whether pupil mimicry modulates trust, we labeled each trial as a "mimicry trial" or a "no mimicry trial" depending on

the participants' pupillary behavior. The fixed effects were partners' pupil size (dilate, constrict, static), mimicry (mimicry, no mimicry), and partners' pupil size × mimicry. The target was the level of investment.

fMRI Data Acquisition. We collected the fMRI data on a 3.0-T Philips Achieva XT MRI scanner equipped with a standard 32-channel head coil. Structural images were obtained with a gradient echo-planar T1 sequence T1 turbo field echo, 240 × 188 mm^2 field-of-view (FOV), comprising a full brain volume of 220 slices (1-mm slice thickness). Volumes were acquired continuously with a repetition time (TR) of 2 s and an echo time (TE) of 3.73 ms [8° flip angle (FA), sagittal orientation]. Next, functional data were collected with T2*-weighted echo planar imaging sequence (2.0 s TR, 27.63 ms TE, 192 × 141.24 mm^2 FOV, 39 slices, 3.3-mm slice thickness, 76.1 °FA, sagittal orientation) covering the whole brain. The fMRI data were analyzed and preprocessed using the fMRI Expert Analysis Tool (FEAT) in FSL v6.0 (Oxford Centre for Functional MRI of the Brain Software Library) (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki) on a MacBook Pro (Retina, 15-inch, mid-2015; Mac OS X 10.11.6).

fMRI Data Preprocessing. Preprocessing steps were run ahead of the first-level analysis and included motion correction, spatial smoothing using a Gaussian kernel of full-width at half-maximum (5 mm), and high-pass temporal filtering with a cut-off of 100 s. Voxels belonging to brain tissue were extracted from nonbrain tissue voxels using the Brain Extraction Tool (BET). Data from all runs were realigned to the mean volume of the middle run using a least-squares approach with six-degree rigid spatial transformation.

fMRI Analysis. fMRI data were analyzed using the GLM for event-related designs in FEAT tool in FSL 5.6 (60). All fMRI data were prewhitened, slice-time–corrected, spatially smoothed, motion-corrected, and high-pass–filtered. In the first level analysis, the hemodynamic response to events of each condition was modeled as the main effect by the hemodynamic response function. To correct for motion artifacts, subject-specific realignment parameters were modeled as covariates of no interest. Linear contrasts of regression coefficients (β-values) were computed at the run level, averaged at the subject level, and taken to a group-level random-effect analysis, using one-sample *t* tests. The same steps applied for ROI analysis. Our primary goal was to determine if the ToM network is modulated by pupil mimicry. We analyzed the fMRI data using two main GLMs of blood oxygen-level dependent (BOLD) responses with first-order autoregression.

GLM 1: Pupil Mimicry. The GLM contained regressors including: (*i*) pupil-dilation mimicry, (*ii*) pupil-constriction mimicry, (*iii*) no pupil-dilation mimicry, (*iv*) no pupil-constriction mimicry, and (*v*) static pupil trials. For this GLM we calculated the following first-level single-subject contrasts: pupil-dilation mimicry vs. baseline, pupil-constriction mimicry vs. baseline, no pupil-dilation mimicry vs. baseline, no pupil-constriction mimicry vs. baseline, pupil mimicry > no mimicry (pooling over constriction and dilation mimicry), no mimicry > pupil mimicry.

GLM 2: Partner's Pupil Change. In a control analysis, we looked for regions that show significant increase in activation in response to changes in partner's pupil size irrespective of mimicry. The following repressors were used: (*i*) partner's pupils dilate, (*ii*) partner's pupils constrict, (*iii*) partner's pupils stay static. The analysis of main interest was in the whole-brain contrast comparing trials when a partner's pupil change (dilate/constrict) > stay static. In addition, the following contrasts were examined: partner's pupils dilate > partner's pupils stay static, partner's pupil constrict > partner's pupil stay static, partner's pupils dilate > partner's pupils constrict.

Whole-Brain Analysis. We pooled the fMRI data for each condition across three runs using a second-level (within-subject) fixed-effects analysis. A third-level (across-subject) analysis was performed within mixed-effects (FLAME 1 + 2) analysis in FSL, treating subjects as a random effect. The effect for each experimental condition was calculated with FEAT. Unless otherwise specified, all cortical regions with a height threshold of $z = 2.3$ and a cluster probability of $P < 0.05$ were reported. The resulting contrast images were linearly registered to the anatomical structure using FMRIB's Linear Registration Tool (FLIRT) with 7 degrees-of-freedom and the full search space, then spatially normalized to the T1-weighted Montreal Neurological Institute (MNI) -152 stereotaxic space template (2 mm) using FMRIB's Non-Linear Registration Tool (FNIRT) with 12 degrees-of-freedom and the full search space. Activation maps were overlaid on the MNI 2-mm brain and regions were determined using the Harvard–Oxford Cortical Structural Atlas that accompanies FSL.

ROI Analysis. An additional ROI analysis was performed using Featquery within FSL within subject-specific ToM masks (Fig. 6). First, we functionally defined ROIs. In the false-belief localizer, the GLM conditions were (*i*) the false belief (ToM) and (*ii*) false-photograph contrast (neutral). A group analysis was conducted on the false belief > false photograph contrast, providing threshold maps ($z = 2.3$, $P = 0.05$). By contrasting these conditions, we localized regions that subject's recruited when they were processing others' mental states. Within the threat-localizer, the GLM contained the following two regressors: (*i*) threat and (*ii*) neutral. A group analysis was conducted on the very threatening > nonthreatening contrast providing *t*-maps ($z = 2.3$, $P = 0.05$). Contrasting those conditions localizes regions that are recruited during threat. The final ROIs were determined by additional inclusion masks obtained via Neurosynth (27). The masks were derived from a metaanalysis of previous studies displaying brain regions that are consistently active in articles that include the name "theory of mind" and "threat" in the abstract. The final ToM regions were defined in the individual space by subtracting the binarized masks acquired by the localizers (liberal threshold $z = 1.5$, $P = 0.05$) from the average activation mask obtained via Neurosynth, in each participant separately. As a result, we created subject-specific ToM inclusion masks in the individual space, which were used for the further ROI analysis. The parameter estimates of the neural activation were extracted from the ROIs for each subject, and averaged across the following four experimental conditions: (*i*) pupil-dilation mimicry, (*ii*) pupil-constriction mimicry, (*iii*) no pupil-dilation mimicry, (*iv*) no pupil-constriction mimicry. A one-way repeated-measures ANOVA was conducted to compare the ToM activation across conditions with zero determined by the implicit baseline (i.e., whatever is not included in the model). This was followed by pairwise comparisons with Bonferoni correction for multiple comparisons.

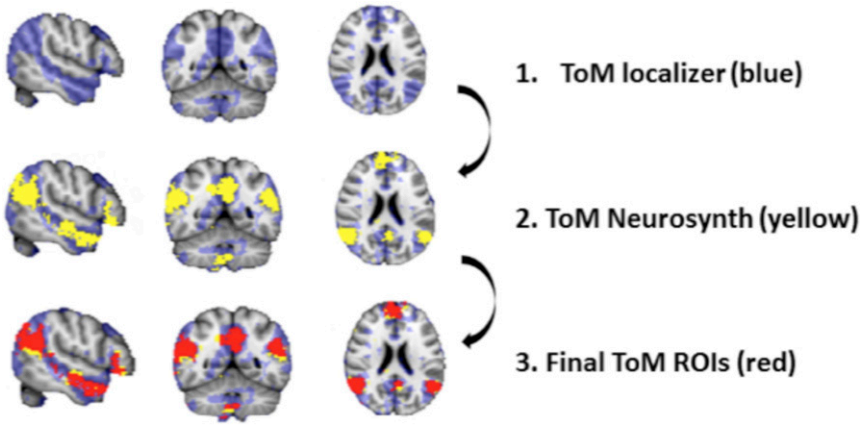

Fig. 6. ROIs selection example from one subject. The final ToM mask included MNI coordinates mentioned by Saxe and Kanwisher (63). These were [−54, −60, 21] for the left TPJ, [51, −54, 27] for the right TPJ, [−9, −51, 33] for the precuneus, [−57, −27, −12] for the left anterior STS, and [66, −18, −15] for the right anterior STS. All subjects shared activation in threat ROIs: amygdala [24, 2, −20/−22, 0, −22], frontal pole: [−24, 58, 16], brainstem [2, −24, −14]. The background image reflects MNI 2-mm template (0.05 voxel size smoothing kernel).

ToM and Trust. To link ToM activation back to trust behavior, the level of investment on each trial was taken as a regression parameter convolved with the hemodynamic response function to identify the regions that most closely correlated with the level of investment. The higher-level analysis and group-level analyses were performed within the ToM mask by averaging the mean activation within and between subjects, without any additional contrasts. Based on recent work showing parametric neuroimaging analyses to be susceptible to inflated false-positive rates (61), we corrected for multiple comparisons with FSL's randomize threshold-free cluster enhancement (62) with nonparametric permutation testing (5,000 permutations) and a variance smoothing kernel of 5 mm. This method enhances the signal of contiguous voxels that form clusters, but returns voxel-wise *P* values family-wise error-corrected (FWE-corrected) for the multiple voxels within a ROI. The *P* values reported in the text are FWE-corrected *P* values from the minimum voxels in the cluster.

Overview of the Statistical Tests Used for the Analysis of the Neural Data in the Present Study. Parametric tests were used with the assumption of normality (the normality of the data were not formally tested). This approach is typical in the analysis approaches used for neuroimaging. It is worth noting that, for some key results, we also conducted permutation tests, which do not require normality assumptions regarding the data.

ACKNOWLEDGMENTS. We thank to Dr. W. Boekel and D. Lindh for insightful discussions, and Jasper Wijnen for programming the task. Funding Research was supported by the Netherlands Science Foundation VENI 016-155- 082 (to M.E.K.) and Talent Grant 406-15-026 from Nederlandse Organisatie voor Wetenschappelijk Onderzoek (to M.E.K. and E.P.), and an Affect Regulation grant (to C.K.W.D.D.). This study was further funded by the Brain and Cognition Research Priority Program of the University of Amsterdam.

1. Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415:137–140.
2. Rand DG, Greene JD, Nowak MA (2012) Spontaneous giving and calculated greed. *Nature* 489:427–430.
3. Tamietto M, de Gelder B (2010) Neural bases of the non-conscious perception of emotional signals. *Nat Rev Neurosci* 11:697–709.
4. De Dreu CKW, et al. (2016) In-group defense, out-group aggression, and coordination failures in intergroup conflict. *Proc Natl Acad Sci USA* 113:10524–10529.
5. Hess EH (1975) The role of pupil size in communication. *Sci Am* 233:110–112, 116–119.
6. Kret ME (2017) The role of pupil size in communication. Is there room for learning? *Cogn Emotion* 5:1–7.
7. Kleinke CL (1986) Gaze and eye contact: A research review. *Psychol Bull* 100:78–100.
8. Kobayashi H, Kohshima S (1997) Unique morphology of the human eye. *Nature* 387: 767–768.
9. Kret ME, Fischer AH, De Dreu CKW (2015) Pupil mimicry correlates with trust in in-group partners with dilating pupils. *Psychol Sci* 26:1401–1410.
10. Wehebrink KS, Koelkebeck K, Piest S, de Dreu CKW, Kret ME (2018) Pupil mimicry and trust—Implication for depression. *J Psychiatr Res* 97:70–76.
11. Kret ME, De Dreu CKW (2017) Pupil-mimicry conditions trust in partners: Moderation by oxytocin and group membership. *Proc Biol Sci* 284:1–10.
12. Kret ME, Tomonaga M, Matsuzawa T (2014) Chimpanzees and humans mimic pupil-size of conspecifics. *PLoS One* 9:e104886.
13. Fawcett C, Arslan M, Falck-Ytter T, Roeyers H, Gredebäck G (2017) Human eyes with dilated pupils induce pupillary contagion in infants. *Sci Rep* 7:9601.
14. Fawcett C, Wesevich V, Gredebäck G (2016) Pupillary contagion in infancy: Evidence for spontaneous transfer of arousal. *Psychol Sci* 27:997–1003.
15. Prochazkova E, Kret ME (2017) Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neurosci Biobehav Rev* 80:99–114.
16. Harrison NA, Singer T, Rotshtein P, Dolan RJ, Critchley HD (2006) Pupillary contagion: Central mechanisms engaged in sadness processing. *Soc Cogn Affect Neurosci* 1:5–17.
17. Demos KE, Kelley WM, Ryan SL, Davis FC, Whalen PJ (2008) Human amygdala sensitivity to the pupil size of others. *Cereb Cortex* 18:2729–2734.
18. Harrison NA, Wilson CE, Critchley HD (2007) Processing of observed pupil size modulates perception of sadness and predicts empathy. *Emotion* 7:724–729.
19. Harrison NA, Gray MA, Critchley HD (2009) Dynamic pupillary exchange engages brain regions encoding social salience. *Soc Neurosci* 4:233–243.
20. Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* 89:221–234.
21. Critchley HD (2009) Psychophysiology of neural, cognitive and affective integration: fMRI and autonomic indicants. *Int J Psychophysiol* 73:88–94.
22. Lee TW, Josephs O, Dolan RJ, Critchley HD (2006) Imitating expressions: Emotion-specific neural substrates in facial mimicry. *Soc Cogn Affect Neurosci* 1:122–135.
23. Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev* 42:9–34.
24. Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015) Deconstructing and reconstructing theory of mind. *Trends Cogn Sci* 19:65–72.
25. Saxe R, Wexler A (2005) Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia* 43:1391–1399.
26. Kret ME, Pichon S, Grèzes J, de Gelder B (2011) Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage* 54:1755–1762.
27. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8:665–670.
28. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
29. Critchley H, et al. (2000) Explicit and implicit neural mechanisms for processing of social information from facial expressions: A functional magnetic resonance imaging study. *Hum Brain Mapp* 9:93–105.
30. Mitchell RLC, Phillips LH (2015) The overlapping relationship between emotion perception and theory of mind. *Neuropsychologia* 70:1–10.
31. Pelphrey KA, Morris JP, Michelich CR, Allison T, McCarthy G (2005) Functional anatomy of biological motion perception in posterior temporal cortex: An fMRI study of eye, mouth and hand movements. *Cereb Cortex* 15:1866–1876.
32. Rossion B, et al. (2003) A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126:2381–2395.
33. Dufour N, et al. (2013) Similar brain activation during false belief tasks in a large sample of adults with and without autism. *PLoS One* 8:e75468.
34. Dodell-Feder D, Koster-Hale J, Bedny M, Saxe R (2010) fMRI item analysis in a theory of mind task. *Neuroimage* 55:705–712.
35. Senju A, Johnson MH (2009) The eye contact effect: Mechanisms and development. *Trends Cogn Sci* 13:127–134.
36. Reimer J, et al. (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat Commun* 7:13289.
37. Kanske P, Böckler A, Trautwein FM, Singer T (2015) Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind. *Neuroimage* 122:6–19.
38. Singer T (2006) The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neurosci Biobehav Rev* 30:855–863.
39. Fan Y, Duncan NW, de Greck M, Northoff G (2011) Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neurosci Biobehav Rev* 35:903–911.
40. Decety J, Lamm C (2006) Human empathy through the lens of social neuroscience. *Sci World J* 6:1146–1163.
41. Gallese V, Keysers C, Rizzolatti G (2004) A unifying view of the basis of social cognition. *Trends Cogn Sci* 8:396–403.
42. Rizzolatti G, Sinigaglia C (2016) The mirror mechanism: A basic principle of brain function. *Nat Rev Neurosci* 17:757–765.
43. Caspers S, Zilles K, Laird AR, Eickhoff SB (2010) ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50:1148–1167.
44. Tylén K, Allen M, Hunter BK, Roepstorff A (2012) Interaction vs. observation: Distinctive modes of social cognition in human brain and behavior? A combined fMRI and eye-tracking study. *Front Hum Neurosci* 6:331.
45. Emonds G, Declerck CH, Boone C, Seurinck R, Achten R (2014) Establishing cooperation in a mixed-motive social dilemma. An fMRI study investigating the role of social value orientation and dispositional trust. *Soc Neurosci* 9:10–22.
46. Krueger F, et al. (2007) Neural correlates of trust. *Proc Natl Acad Sci USA* 104:20084–20089.
47. Lamm C, Decety J, Singer T (2011) Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage* 54:2492–2502.
48. de Dreu CKW, Giffin MR (2017) Neuroendocrine pathways to in-group bounded trust and cooperation. *Trust in Social Dilemmas*, eds Van Lange PAM, Rockenbach B, Yamagishi T (Oxford Scholarship Online, New York), pp 57–76.
49. Negro E, et al. (2015) Neurofunctional signature of hyperfamiliarity for unknown faces. *PLoS One* 10:e0129970.
50. Weibert K, et al. (2016) An image-invariant neural response to familiar faces in the human medial temporal lobe. *Cortex* 84:34–42.
51. Levenson RW, Gottman JM (1983) Marital interaction: Physiological linkage and affective exchange. *J Pers Soc Psychol* 45:587–597.
52. van der Schalk J, Hawk ST, Fischer AH, Doosje B (2011) Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion* 11:907–920.
53. Saxe R, Carey S, Kanwisher N (2004) Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annu Rev Psychol* 55:87–124.
54. Davis MH (1983) Measuring individual differences in empathy: Evidence for a multidimensional approach. *J Pers Soc Psychol* 44:113–126.
55. Heimberg RG, et al. (1999) Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychol Med* 29:199–212.
56. Spielberger CD (2010) State-trait anxiety inventory. *The Corsini Encyclopedia of Psychology* (John Wiley & Sons, Inc., Hoboken, NJ).
57. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I (2001) The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry* 42:241–251.
58. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. *Arch Gen Psychiatry* 4:561–571.
59. van Breen JA, De Dreu CKW, Kret ME (2018) Pupil to pupil: The effect of a partner's pupil size on (dis)honest behavior. *J Exp Soc Psychol* 74:231–245.
60. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl 1):S208–S219.
61. Flandin G, Friston KJ (2016) Analysis of family-wise error rates in statistical parametric mapping using random field theory. *Proc Natl Acad Sci USA* 113:7900–7905.
62. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44:83–98.
63. Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19:1835–1842.