

Self-regulation via neural simulation

Michael Gilead^{1,2†}, Chelsea Boccagno¹, Melanie Silverman¹, Ran Hassin³, Jochen Weber¹, Kevin Ochsner^{1†}

¹Department of Psychology, Columbia University, New York, NY ²Department of Psychology, Ben-Gurion University of the Negev, Israel ³Department of Psychology, the Hebrew University of Jerusalem, Israel †To whom correspondence should be addressed.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Can taking the perspective of other people modify our own affective responses to stimuli? To address this question, we examined the neurobiological mechanisms supporting the ability to take another person's perspective and thereby emotionally experience the world as they would. We measured participants' neural activity as they attempted to predict the emotional responses of two individuals that differed in terms of their proneness to experience negative affect. Results showed that behavioral and neural signatures of negative affect (amygdala activity, and a distributed multi-voxel pattern reflecting affective negativity) "simulated" the presumed affective state of the target person. Furthermore, the anterior mPFC—a region implicated in mental state inference—exhibited a perspective-dependent pattern of connectivity with the amygdala, and the multi-voxel pattern of activity within the mPFC differentiated between the two targets. We discuss the implications of these findings for research on perspective-taking and self-regulation.

perspective-taking | emotion regulation | mPFC | simulation | amygdala

The ability to respond adaptively in the face of emotionally challenging situations is essential to mental and physical health. So much so, in fact, that emotion dysregulation is a core feature of virtually every form of psychopathology. Given this, it isn't surprising that the past decade has seen enormous growth in behavioral and brain research asking how we can effectively regulate our emotions. While this work has made many important advances (1, 2), it has focused almost entirely on cognitive regulatory strategies that involve controlling attention to and/or rethinking the meaning of stimuli and events. As such, this work has completely overlooked the way in which *social* cognitive processes can be used to regulate our emotions.

The use of social cognition to regulate emotion was suggested by classic works in social psychology (3), which noted that by simulating others' perspective on the world we could shape our own experience and behavior. It is exemplified by "(Stanislavski) method actors" who understand a role by attempting to generate within themselves the presumed thoughts and feelings of a character, thereby allowing themselves to go beyond the written words in the script and respond as their character would (4). It is also present in everyday life when we seek guidance with respect to emotional dilemmas by asking ourselves how a friend, family member, mentor or religious figure (e.g., "What would Jesus do?") would respond in that situation.

In the current research we asked whether and how taking the perspective of *other people* can modify our *own* affective responses to stimuli. For example, by thinking of how someone more brave than ourselves would respond to a situation, we might down-regulate negative emotions, decrease aggression, and calm frazzled nerves. Alternatively, by thinking of how someone more sensitive and anxious would respond to the situation, we might enhance vigilance and increase reactivity to threatening situations.

To address these possibilities we conducted a neuroimaging experiment investigating whether seeing the world through the eyes of a "tough" vs. a "sensitive" person can up-regulate or down-regulate affective responding, respectively. Furthermore, we sought to delineate the neural mechanisms by which such perspective-dependent regulatory consequences transpire.

While no prior work has addressed these questions, *per se*, the literatures on emotion regulation (1, 5-11) and perspective-taking (12-18) can be integrated to generate testable hypotheses. On one hand, research on emotion regulation has shown that activity in lateral prefrontal cortex (i.e., dlPFC and vlPFC) and middle medial prefrontal cortex (i.e., pre-SMA, avMCC and adMCC; (19)) supports the use of cognitive strategies to modulate activity in (largely) subcortical systems for triggering affective responses, such as the amygdala, thereby altering individuals' emotional responses (20). On the other hand, research on perspective-taking has shown that drawing inferences about the mental states of others (also known as "mentalizing")—as would be involved in simulating their perspective on an event—is supported by a network of regions centered on the anterior medial frontal cortex, specifically, the pgACC and the dorsomedial prefrontal cortex (dmPFC; (13), (21), (19)).

Based on these literatures, we formulated two novel hypotheses. First, we predicted that by taking the perspective of a target person, an individual could change behavioral and brain markers of affective responding, thereby providing evidence that one is emotionally experiencing the world the way the target would. Second, we predicted that these regulatory effects would be supported not by lateral prefrontal regions implicated in attentional and cognitive control, but rather, by dorsomedial prefrontal regions involved in perspective-taking. Put another way, we predicted that perspective-taking related activity in the anterior mPFC would regulate activity in neural systems for affective responding.

To test these hypotheses we collected whole-brain fMRI data while participants attempted to predict the affective responses of other individuals. Before scanning, participants were presented with descriptions of two people, who they were led to believe

Significance

As Harper Lee tells us in her novel *To Kill a Mockingbird*, "You never really understand a person until you consider things from his point of view, until you climb in his skin and walk around in it". Classic theories in social psychology argue that this purported process of "social simulation" provides the foundations for self-regulation. In light of this, we investigated the neural processes whereby humans may regulate their affective responses to an event by simulating the way others would respond to it. Our results suggest that during perspective-taking, behavioral and neural signatures of negative affect indeed mimic the presumed affective state of others. Furthermore, the anterior mPFC—a region implicated in mental state inference—may orchestrate this affective simulation process.

Reserved for Publication Footnotes

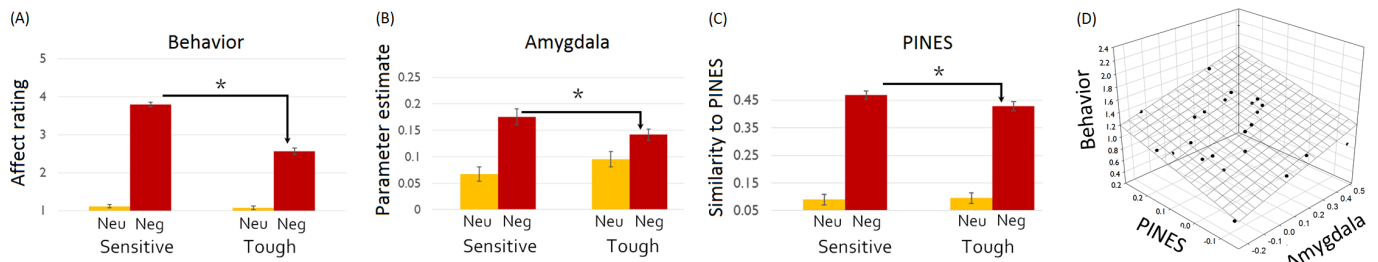


Fig. 1. (A) Behavioral ratings of negative affect in response to negative images were higher for Sensitive (vs. Tough) targets. (B) Right amygdala response to Negative images was higher when adopting the Sensitive (vs. the Tough) perspective. (C) When participants adopted the Sensitive (vs. Tough) perspective, their neural response to negative images reflected higher levels of negative affect, measured as the level of similarity to the PINES pattern. Error bars denote within-participant standard errors. (D) Participants who exhibited a greater difference in amygdala activity and PINES expression for the Tough vs. Sensitive target subsequently estimated greater differences in predicted negative affect for these targets.

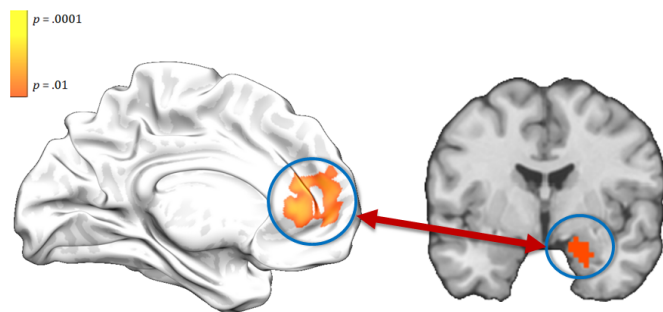


Fig. 2. The Right amygdala cluster identified by contrasting the processing of negative and neutral images from the self's perspective (on right), and the anterior mPFC region that was implicated in perspective-based regulation of amygdala activity (on the left). The results suggest that the anterior mPFC up- or down-regulated amygdala activity as a function of the perspective (Sensitive vs. Tough) that participants adopted.

had previously participated in the experiment. These descriptions suggested that one person was likely to be emotionally sensitive and squeamish, while the other was likely to be rugged and tough. Next, participants viewed neutral- and negative-affect-inducing images and evaluated the images from either their own or the tough or sensitive targets' perspective.

We examined the effect of perspective-taking on multiple behavioral and brain markers of affective responding, including reports of the target's predicted affective reactions to stimuli, activation in the amygdala (which is the brain region most strongly associated with detecting, encoding, and promoting responses to affectively relevant and especially potentially threatening stimuli (22, 23)), and finally, a recently identified picture-induced negative emotion signature (PINES; (24)). PINES is a distributed, whole-brain multi-voxel activation pattern developed using machine learning techniques that can reliably predict levels of negative affect elicited by aversive images. Because this signature is not affected by general arousal and is not reducible to activity in the amygdala, it provides a "neural marker" of negative affect independent of participants' own self-reports. We predicted that both neural measures of negative affective responding (amygdala and PINES) would "simulate," the presumed affective state of the target person; namely, negative affect-related activity would be up- vs. down-regulated for the sensitive (vs. tough) perspective.

To address the prefrontal systems that might support perspective-taking and regulate affective responding, we used a combination of connectivity and multi-voxel pattern analyses to identify a brain region whose activity was associated with amygdala up-regulation when adopting the sensitive perspective and/or down-regulation for the tough perspective—and whose distributed pattern of activity provided evidence that it differen-

tially represented the two perspectives. As noted, we predicted this region to be located in the anterior mPFC.

Results

Does perspective-taking modulate affective processing?

Behavioral ratings: A manipulation check showed that participants reported more negative affect in response to Negative than to Neutral images, $F(1,23) = 572.56, p < .001$. We conducted an ANOVA to see whether the perspective manipulation indeed altered participants predicted affective response. The results showed a significant interaction, $F(1,23) = 202.08, p < .001$, such that affect ratings were lower when participants viewed Negative images from the perspective of the Tough ($M = 2.564, SD = 0.118$) vs. the Sensitive target ($M = 3.793, SD = 0.103$), $t(23) = 12.60, p < .001$. There was no significant difference in ratings for Neutral images from the perspective of the Tough ($M = 1.071, SD = 0.021$) and Sensitive ($M = 1.117, SD = 0.028$) targets, $t(23) = 1.58, p = .126$. (Figure 1). There were also no significant differences in response latencies for the Sensitive ($M = 921.57, SD = 129.19$) and Tough ($M = 941.90, SD = 120.97$) perspectives, $t(23) = 0.8, n.s.$

For Negative images, affect ratings from the Self perspective ($M = 3.140, SD = 0.496$) were higher than those for the Tough target, $t(23) = 5.08, p < .001$, and lower than those for the Sensitive target, $t(23) = 8.10, p < .001$. For Neutral images, affect ratings from the Self perspective ($M = 1.058, SD = 0.104$) did not differ from the Tough perspective, $t(23) = 0.48, n.s.$, and were lower than those for the Sensitive perspective, $t(23) = 2.63, p = .015$.

Based on participants' affect ratings for the Tough, Sensitive, and Self perspectives, we calculated for each participant a measure of "Similarity to Sensitive/Tough Target" that indexed the extent to which affect ratings from the self-perspective were more similar to one target or the other. This measure, alongside with other neural and self-report measures of self-other similarity, indicated that overall, participants did not identify more with one perspective or another, and that the level of self-other similarity did not modulate our key measures (see SI for details of these analyses).

Figure 1. (A) Behavioral ratings of negative affect in response to negative images were higher for Sensitive (vs. Tough) targets. (B) Right amygdala response to Negative images was higher when adopting the Sensitive (vs. the Tough) perspective. (C) When participants adopted the Sensitive (vs. Tough) perspective, their neural response to negative images reflected higher levels of negative affect, measured as the level of similarity to the PINES pattern. Error bars denote within-participant standard errors. (D) Participants who exhibited a greater difference in amygdala activity and PINES expression for the Tough vs. Sensitive target subsequently estimated greater differences in predicted negative affect for these targets.

273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340

Amygdala analysis: As a first step in examining whether perspective-taking modulates affective processing, we defined the right and left amygdala as anatomical regions-of-interest based on the Harvard-Oxford probabilistic atlas, (using voxels with a 50% or higher probability of being labeled as the amygdala) and extracted parameter estimates for the six conditions (Negative/Neutral x Sensitive/Tough/Self). As predicted, in the left amygdala, when participants observed the images from their own perspective, activation was higher for Negative ($M = 0.177$, $SD = 0.167$) than for Neutral ($M = 0.086$, $SD = 0.148$) images, $t(23) = 2.77$, $p = .005$; Likewise, in the right amygdala, activation was higher for Negative ($M = 0.149$, $SD = 0.149$) than for Neutral ($M = 0.087$, $SD = 0.112$) images, $t(23) = 2.15$, $p = .020$.

After establishing that amygdala activity is responsive to the presentation of aversive images when viewing them from one's own perspective, we asked whether the amygdala was modulated when taking a tough or sensitive perspective. To do so, we conducted a 2x2 ANOVA with Perspective (Sensitive/Tough) and Valence (Negative/Neutral) as within-participant factors. As predicted, the results showed an interaction of Perspective and Valence in both the right $F(1,23) = 6.77$, $p = .007$, partial eta-squared = 0.227, and the left amygdala, $F(1,23) = 2.96$, $p = .049$, partial eta-squared = 0.114. In the right hemisphere, amygdala activation was lower when viewing Negative images from the perspective of the Tough ($M = 0.141$, $SD = 0.098$) vs. the Sensitive target ($M = 0.175$, $SD = 0.120$), $t(23) = 2.02$, $p = .027$; there was no significant difference in activation for Neutral images from the perspective of the Tough ($M = 0.095$, $SD = 0.092$) and Sensitive ($M = 0.067$, $SD = 0.089$) targets, $t(23) = 1.38$, $p = .180$. In the left hemisphere, there was a marginally significant effect wherein amygdala activation was lower when viewing Negative images from the perspective of the Tough ($M = 0.175$, $SD = 0.130$) vs. Sensitive target ($M = 0.200$, $SD = 0.129$), $t(23) = 1.45$, $p = .079$. There was no significant difference in activation for Neutral images from the perspective of the Tough ($M = 0.100$, $SD = 0.117$) and Sensitive ($M = 0.078$, $SD = 0.105$) targets, $t(23) = 1.04$, $p = .306$. For both Negative and Neutral images, the Self perspective did not differ from the Sensitive perspective in neither the right nor left amygdala (p 's > .27); likewise, the Self perspective did not differ from the Tough perspective in the neither the right amygdala (p 's > .64)—suggesting that the Perspective x Valence interaction was not driven solely by either the Tough or Sensitive perspective.

Although the amygdala was more active for Negative vs. Neutral images when viewed from the self's perspective, it could be argued that different sub-regions of the amygdala may be differentially engaged under the self and other conditions. In order to address this concern, we conducted a whole-brain search based on the Self Negative > Self Neutral contrast. This contrast yielded significant activation across several brain regions, including the left and right amygdala, which we then masked with anatomically-defined amygdala regions based on the Harvard-Oxford probabilistic atlas. The interaction of Perspective and Valence remained significant in the right amygdala cluster (47 voxels, peak coordinate, $x = 18$, $y = -3$, $z = -18$), $F(1,23) = 5.25$, $p = .015$; however, the interaction in the left amygdala (102 voxels, peak MNI coordinate, $x = -12$, $y = -6$, $z = -18$) did not attain significance, $F(1,23) = 1.25$, $p = .136$. In light of this, we limited our subsequent analyses to the right amygdala cluster.

PINES analysis: Another concern is that although amygdala activation is strongly associated with the processing of negatively valenced stimuli, it is sometimes activated when processing positive stimuli (25), which may reflect a more general role for the amygdala in detecting and encoding of goal-relevant stimuli (26-28). This suggests that the amygdala's role in negative affect be indirect, which complicates attempts to rely on its activation as a neural marker of negative affective responses.

In light of this, we sought to strengthen our claim that emotional perspective-taking modulates negative affective processing by utilizing a recently identified Picture-Induced Negative Emotion Signature (PINES; (24)). The PINES is a whole brain activation pattern developed using machine learning techniques that can reliably predict self-reported emotional responses to aversive images. As noted, prior work (24) has shown that this signature is not affected by general arousal, and is not reducible to patterns of activity in the amygdala. Thus, it provides an independently validated "neural marker" of experienced affective negativity.

We first validated the PINES method in the current dataset by showing that the PINES expression score was significantly higher when observing Negative ($M = 0.490$, $SD = 0.192$) vs. Neutral ($M = 0.069$, $SD = 0.152$) images from the perspective of the self, $t(23) = 11.24$, $p < .001$. Furthermore, regardless of perspective, the PINES score was higher for Negative ($M = 0.462$, $SD = 0.175$) vs. Neutral ($M = 0.094$, $SD = 0.156$) images viewing conditions, $F(1,23) = 349.18$, $p < .001$. Having established that the PINES pattern differentiates images as a function of their negativity in our dataset, we investigated the effect of Perspective (Sensitive/Tough) and Valence (Negative/Neutral) on the degree of affective negativity, as gauged by the PINES expression score. As predicted, there was an interaction between Perspective and Valence, mirroring the effect in the amygdala, $F(1,23) = 7.72$, $p = .005$. The PINES expression score was lower when viewing Negative images from the perspective of the Tough ($M = 0.428$, $SD = 0.173$) vs. Sensitive target ($M = 0.468$, $SD = 0.188$), $t(23) = 1.83$, $p = .039$; there was no significant difference in expression for Neutral images from the perspective of the Tough ($M = 0.095$, $SD = 0.092$) vs. the Sensitive ($M = 0.067$, $SD = 0.089$) targets, $t(23) = 1.28$, $p = .210$. The PINES calculation was done on a trial-based model which we used for multivariate analyses (PINES, MVPA and pattern similarity analysis, later described). The results of the analysis are identical when using an aggregated-trial model. Valence x Perspective interaction: $F(1,23) = 7.72$, $p = .005$. (See SI for further details concerning PINES method).

Brain-behavior correlation: Having established that the different measures of affective response are each impacted by the perspective taken, an important next question was whether and how they are related. In particular, it is important to demonstrate that one or both of the neural measures—amygdala activity and/or PINES score—predict self-reports of negative affective experience, as such correlations would support the idea that the neural regions supporting simulation of the Tough vs. Sensitive target's emotions lead to changes in a reports of affective experience.

To address this issue, we calculated for each participant a measure of each type of effect on a measure of affective response (i.e. the Behavioral Effect, PINES Effect, and Amygdala Effect), as the difference between Negative and Neutral conditions for the Sensitive vs. Tough Perspectives (i.e. Sensitive[Negative-Neutral] - Tough[Negative-Neutral]). As predicted, the results showed that participants who exhibited a greater difference in amygdala activity for the Tough vs. Sensitive target subsequently exhibited a greater difference in their behavioral evaluations of the affective states of these targets, $r = .38$, $p = .033$. Likewise, participants who exhibited a greater difference in PINES scores for the Tough vs. Sensitive target subsequently exhibited a greater difference in their behavioral evaluations of the affective states of these targets, $r = .39$, $p = .014$. Interestingly, there was no correlation between the PINES Effect and Amygdala Effect, $r = .01$, as would be expected based on prior work establishing the PINES that suggested they could be independent predictors of negative affect (24). A multiple regression with both the PINES Effect and Amygdala Effect as predictors and the Behavioral Effect as the dependent variable showed a significant effect for the PINES ($b = 1.193$, $SE = 0.562$, [95% CI: 0.023, 2.364], $p = .046$), and a marginally significant effect for the Amygdala ($b = 0.742$, $SE =$

341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408

0.363, [95% CI: -0.103, 1.499], $p = .053$, R-squared = .295. Thus, our results suggest that each of the two patterns made a unique contribution to changing reports of affective responding.

What are the neural systems that support the perspective-based modulation of affective processing?

PPI: To identify regions that may play a key role in the perspective-taking-based regulation of amygdala activity, we conducted a psychophysiological interaction (PPI) analysis (29). This was done by creating regressors for each of the experimental conditions, the amygdala time series, and interaction terms for the amygdala time series and the experimental conditions. The difference of the relevant PPI-term regression coefficient, i.e., [(amygdala time series) \times (Sensitive Negative)] $>$ [(amygdala time series) \times (Tough Negative)] was then subjected to a second-level random effects analysis, which also included a between-participants covariate coding for the average difference in affect rating across conditions.

This analytical strategy allowed us to identify regions that during Sensitive Negative trials were more positively correlated with right amygdala activation and/or during Tough Negative trials we more negatively correlated with amygdala activity, and exhibited this pattern more so for participants that displayed greater perspective-related modulation of affective response (i.e., greater Behavioral Effect). The resulting analysis yielded a cluster of 203 voxels in the anterior mPFC (specifically, pgACC and dmPFC; peak MNI coordinate, $x = -9$, $y = 54$, $z = 15$); Figure 2), which survived the $p < .05$, whole-brain corrected significance threshold determined by AlphaSim. Masking out this anterior mPFC cluster did not alter the results of the PINES analysis.

In other words, participants who showed the greatest perspective-dependent modulation of affective experience also showed the greatest perspective-dependent modulation of the anterior mPFC-amygdala pathway. More specifically, our results showed that participants who showed the greatest perspective-dependent modulation of affect ratings showed a negative co-activation pattern between the anterior mPFC and the amygdala when adopting the Tough perspective.

MVPA: If the anterior mPFC cluster identified in the PPI analysis is indeed responsible for the perspective-based modulation of amygdala activity, then the multi-voxel pattern of activity in this region during image viewing could be expected contain information that can discriminate whether participants were taking the perspective of the Tough or Sensitive target. To test this, we conducted a Multi-Voxel Pattern Analysis (MVPA) examining classification accuracy in the anterior mPFC cluster. As predicted, the classifier was able to predict the perspective participants were taking with a mean accuracy of 54.60% ($SD = 6.79$), which significantly differed from chance performance, $t(23) = 3.32$, $p = .001$. There was no difference in overall average levels of activity in this cluster between the Sensitive and Tough conditions, $t(23) = 0.37$, $p = .709$. (See SI for further details concerning MVPA analyses).

Figure 2. The Right amygdala cluster identified by contrasting the processing of negative and neutral images from the self's perspective (on right), and the anterior mPFC region that was implicated in perspective-based regulation of amygdala activity (on the left). The results suggest that the anterior mPFC up- or down-regulated amygdala activity as a function of the perspective (Sensitive vs. Tough) that participants adopted.

Discussion

We sought to investigate whether (and how) taking the perspective of other people can modify our own affective responses to stimuli. We hypothesized that: (i) taking the perspective of others would regulate affective processing in neural mechanisms that subservise one's own affective experience, and (ii) that the neural system involved in regulating perspective-dependent affective

processing would be a region implicated in mental states inference, such as the anterior mPFC (i.e., the dmPFC and pgACC).

Consistent with our first hypothesis, whenever participants took the perspective of a sensitive (vs. tough) target, three neural indicators of negative affective processing converged to suggest that participants "simulated," the presumed affective state of the target individual. First, amygdala activity was up-regulated for the sensitive (vs. tough) perspective. Second, a multi-voxel, whole-brain pattern of activity that has been independently shown to accurately predict participants' affective state (PINES; (24)) indicated up-regulated negative affectivity when taking a sensitive (vs. tough) perspective. Third, participants who behaviorally predicted a greater difference in the affective responses of the sensitive and tough targets also exhibited a greater difference in their PINES and amygdala response when adopting the sensitive (vs. tough) perspectives.

That perspective-taking modulates amygdala activity provides initial support to the claim that perspective-taking modulates affective processing. However, because the amygdala responds to goal-relevant stimuli in general (23, 27, 28), it could be argued that its activation does not reflect negative affective intensity *per se*. Yet, the finding that perspective-taking modulated the PINES pattern—and that this modulation uniquely contributed to predictions of subsequent judgments of a targets affective response over and above amygdala activity—provides strong converging evidence to the claim that "seeing the world through another's eyes" really does change one's own affective processing.

Having provided support for that claim, we sought to delineate the cognitive and neural mechanisms by which such perspective-dependent regulatory consequences occur. Consistent with our second hypothesis, results suggested that the anterior mPFC may regulate, or exert top-down influence over, the affective simulation. Specifically, this brain region exhibited a pattern of perspective-dependent coupling with the amygdala that was dependent on the magnitude of perceived differences in the targets' affective response. Relatively speaking, when adopting a sensitive perspective, anterior mPFC activity was associated with increased amygdala activity; when adopting a sensitive perspective, anterior mPFC activity was associated with relatively decreased amygdala activity. Furthermore, an MVPA analysis showed that that the multi-voxel pattern of activity in this region during image viewing contained information that discriminated whether participants were taking the perspective of the tough or sensitive target.

Implications for the study of perspective-taking

The current research addressed an age-old question concerning the process of perspective-taking. It is often suggested that people are able to take the perspective of others through a process of *simulation* (note that the term simulation is polysemous: it can be used to discuss a cognitive *process* by which people may take the perspective of others, as well as a *consequence* of perspective taking. In this section we refer to the former). The philosopher Alvin Goldman described simulation as such: "First, the attributor creates in herself pretend states intended to match those of the target... The second step is to feed these initial pretend states into some mechanism of the attributor's own psychology ... and allow that mechanism ...to generate one or more new states (e.g., decisions)" (ref. (30), pp. 80–81). In other words, according to simulation theory, the path to understanding the emotions of others relies on a readout from the very same core emotional processes that generate the emotional response in the *self* (see (13, 31, 32) for similar suggestions).

The current study allowed us to investigate the process of simulation with converging measures of affective processing. We showed that participants indeed exhibited greater affect negativity when they took the perspective of the sensitive (vs. tough)

target. Importantly, participants who exhibited greater difference in amygdala activity/PINES expression for the tough vs. sensitive target subsequently exhibited greater difference in their evaluations of the affective state of these targets. Together, these findings present perhaps the most direct evidence, to date, for the viability of simulation theory.

The existence of shared mechanisms for both self- and other-focused processing is a prerequisite for simulation theory. However, it does not suffice to explain the process of perspective-taking. As acknowledged in some of the earliest discussions of simulation theory, if people were to simply copy their own experience and project it onto others, attempts at perspective-taking would be ineffective (15, 33). Thus, for perspective-taking to succeed, individuals must accommodate their simulation on the basis of a conceptual model of the target (e.g., "This guy is neurotic, he must be distressed by cockroaches"). This process is unlikely to rely on the amygdala alone, which is a phylogenetically ancient brain system that is unlikely to subserve the type of symbolic thought involved in conceptually-mediated perspective-taking (18). Therefore, we predicted that amygdala activity should be modulated through an interaction with a brain system that subserves such model-based, conceptual capacities.

As noted earlier, our results suggest that that this system involves the anterior mPFC. This region is widely-implicated in conceptual thought in general (34-36) and social cognition in particular (17, 37). To give one example, recent work shows that multi-voxel patterns of activity in the anterior mPFC can be used to predict which one of two individuals a participant is thinking about (21). The current research dovetails and builds on this prior work by showing that anterior mPFC doesn't just support inferences about others states and traits, but supports simulation of their perspective on world, thereby changing the way that we appraise the affective significance of events and subsequently respond to them.

Implications for models of the self-regulation of emotion

An important implication of the current findings is the suggestion that perspective-taking could have emotion regulatory benefits. In the current study, participants did not have the explicit goal of up- or down-regulating their emotions, and yet, merely trying to understand the emotions of tough vs. sensitive others modulated the activity in a brain system involved in the generation of negative affect. Thus, our research suggests that the attempt to "walk in the shoes" of an emotionally resilient individual may cause people to feel less unpleasant in the face of adversity.

Accordingly, it may be possible to harness the type of emotional perspective-taking studied here as an emotion regulation strategy, aimed at helping individuals cope with emotional distress. Extant research within the field of emotion regulation has shown that people can effectively down-regulate negative affect by using top-down cognitive control (20). However, a limitation of many cognitive emotion regulation strategies is that they depend upon attentional, linguistic and working memory systems supported by lateral prefrontal regions. Lateral prefrontal regions are not fully-developed until late adolescence (38) and can be disrupted under severe stress (39). Thus, the finding that perspective-based regulation of the amygdala relies on anterior medial rather than lateral prefrontal regions may suggest a new pathway for effective emotion regulation.

Specifically, a simulation-based emotion regulation strategy may be important in populations for which strategies dependent on lateral PFC may be problematic because lateral frontal functionality is compromised or yet-to-develop (40). For example, future studies could investigate whether young children may especially benefit from being taught how to regulate their emotions using simulative pretend-play ("imagine that you are a big boy/girl").

More broadly, the current findings highlight that there may be a plurality of computations and neural pathways by which emotion-regulatory consequences can occur. In this way, the current findings contribute to our growing understanding of the complexity of neural interactions that subserve important behavioral outcomes. Hopefully, future research extending the findings described herein could shed further light on strategies that support adaptive socioemotional functioning.

Experimental Procedures

Participants

Twenty-four right-handed participants (12 females, average age 20.5, $SD = 2.577$, range 18-28) participated in the experiment for monetary compensation. All were native-level English speakers, all had normal or corrected vision, and none had a history of neurological or psychiatric disorders and. Sample size was determined a-priori, based on previous neuroimaging studies showing regulation-related modulation of amygdala activity (20). Three additional participants were excluded from the final analysis (one for missing data and two for failing to comply with task instructions, as evident by deviation of more than 3 SDs from the mean affect rating in at least one task condition). Participants gave written consent prior to taking part in the experiment. The study was approved by the Institutional Review Board of Columbia University.

Materials

Target description questionnaires: The descriptions of the tough and sensitive targets were given in the form of printed questionnaires that were ostensibly filled out by two previous participants. At the top of each questionnaire, a name appeared in hand-written text. Both names were matched to each participants' gender. The questionnaire contained demographic details (e.g. place of birth) and responses to personal questions (e.g. music preferences, hobbies). The key differences between the two types of targets arose from the way each one had supposedly responded to particular questions. In actuality, the answers had been pre-tested to elicit perceptions that one target was "tough" and the other "sensitive". For example, the tough character worked as an EMT, enjoyed action and horror movies, and loud music. By contrast, the sensitive character worked as a graphic designer, liked classical music and romantic comedies. Furthermore, in one of the free response items the tough target described him/herself as being relatively resilient and the sensitive character described him/herself as being relatively sensitive. These characteristics were embedded within more mundane details in order to bolster the believability of the experiment.

Affective stimuli: 54 negative images (mean normative valence = 2.76, mean normative arousal = 5.91, on a 1 to 9 scale) and 54 neutral images (mean normative valence = 5.32, mean normative arousal = 3.15) were taken from the International Affective Picture System (41). Both negative and neutral images were divided to three lists, matched for arousal and valence. An additional set of 6 similarly-valenced and arousing negative images were used during training.

Behavioral procedure

Pre-scanning: After providing consent, participants were asked to fill out a questionnaire describing various demographic and personal details about themselves. They were told that in the experiment they will be asked to predict the emotions of previous participants, and that we need their answers to the personal details questionnaire in order to use them for the next participant. In actuality, this questionnaire was only administered to bolster the believability of the experiment, and it was not subsequently used. Immediately after filling out the questionnaire, participants were given the "character description" questionnaires, which were in the same format as the one they filled out. They were asked to read the answers of each previous participant carefully and form an impression of them in their mind.

Participants then were instructed on the task they would perform inside the scanner. They were told that they will be presented with images, and that each image will be preceded either by a cue with the name of the participant whose perspective they should take or by a cue asking them to take their own perspective. Each image would be followed with a screen asking them to rate the affective response (either of themselves or the target individual) the image elicits. They were then told that they should rate the images based upon the perspective they were cued with, and that these answers would be compared with the previous participants' actual ratings. We told participants that trials wherein they gave the rating from their own perspective would be used for the next participants (in actuality, self-perspective trials were used to identify the neural substrates of spontaneous emotional response). Participants' goal was to predict the previous participants' responses as accurately as possible. To increase the incentive to do so, participants were told that if they were in the top 10% of participants in terms of accuracy, they will receive a \$100 bonus (in actuality, the bonus criteria was based on scanner movement). Participants then performed a short training on the task that involved completing sample trials guided by the experimenter.

Finally, as a pre-task manipulation check, participants were asked to recall the answers for each of the two previous participants' questionnaires. Whenever participants made a mistake, the questions were repeated later on until participants arrived at 100% recall accuracy.

681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748

Scanner task: The task consisted of 108 trials (18 negative images and 18 neutral images for each of the three perspectives) that were divided into three functional runs. Each run contained 36 trials (6 negative and 6 neutral for each of the 3 perspectives) and lasted 10 minutes and 48 seconds.

Stimuli were presented using E-Prime 2.0 (Psychology Software Tools, Inc.). Each experimental trial began with the presentation a cue with the name of the participant whose perspective they should take, or a cue asking them to take their own perspective, shown for 2 s. After a jittered fixation period (1-5 s), participants viewed the affective image for 6 s. The image was replaced by a screen that appeared for 3 s, asking them to rate the affective reaction to the image from the perspective they were asked to adopt (1 = neutral, 5 = very bad). The trial concluded with a second jittered fixation period (3-9 s). Stimuli were displayed in random order and the assignment of images to the three perspective conditions was counterbalanced across participants.

- Ochsner KN, Silvers JA, & Buhle JT (2012) Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences* 1251:E1-24.
- Buhle JT, et al. (2013) Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*.
- Mead GH (1934) Mind, self, and society: From the standpoint of a social behaviorist (Works of George Herbert Mead, Vol. 1).
- Bandelj N (2003) How method actors create character roles. *Sociol. Forum* 18(3):387-416.
- Gross JJ (1998) Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology* 74(1):224-237.
- McRae K, Ochsner KN, Gross JJ, & Gabrieli JDE (2002) Exploring the Neural Bases of Emotion Regulation: Comparing the Effects of Cognitive Reappraisal and Working Memory Load. *Proceedings of the Annual Meeting of the Society for Neuroscience*.
- Ochsner KN, Bunge SA, Gross JJ, & Gabrieli JDE (2002) Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience* 14(8):1215-1229.
- Ochsner KN & Gross JJ (2008) Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science* 17(2):153-158.
- Wager TD, Davidson ML, Hughes BL, Lindquist MA, & Ochsner KN (2008) Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* 59(6):1037-1050.
- Kross E, Davidson M, Weber J, & Ochsner K (2009) Coping with Emotions Past: The Neural Bases of Regulating Affect Associated with Negative Autobiographical Memories. *Biological Psychiatry* 65(5):361-366.
- Diekhof EK, Geier K, Falkai P, & Gruber O (2011) Fear is only as deep as the mind allows A coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *Neuroimage* 58(1):275-285.
- Frith CD & Frith U (2006) The neural basis of mentalizing. *Neuron* 50(4):531-534.
- Mitchell JP, Banaji MR, & Macrae CN (2005) The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience* 17(8):1306-1315.
- Mitchell JP, Macrae CN, & Banaji MR (2006) Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50(4):655-663.
- Tamir DI & Mitchell JP (2010) Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences of the United States of America* 107(24):10827-10832.
- Saxe R & Kanwisher N (2003) People thinking about thinking people - The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19(4):1835-1842.
- Skerry AE & Saxe R (2014) A common neural code for perceived and inferred emotion. *The Journal of Neuroscience* 34(48):15997-16008.
- Spunt RP, et al. (2015) Amygdala lesions do not compromise the cortical network for false-belief reasoning. *Proceedings of the National Academy of Sciences of the United States of America* 112(15):4827-4832.
- de la Vega A, Chang LJ, Banich MT, Wager TD, & Yarkoni T (2016) Large-Scale Meta-Analysis of Human Medial Frontal Cortex Reveals Tripartite Functional Organization. *The Journal of Neuroscience : the official journal of the Society for Neuroscience* 36(24):6553-6562.
- Buhle JT, et al. (2014) Cognitive Reappraisal of Emotion: A Meta-Analysis of Human Neuroimaging Studies. *Cerebral Cortex* 24(11):2981-2990.
- Hassabis D, et al. (2014) Imagine All the People: How the Brain Creates and Uses Personality Models to Predict Behavior. *Cerebral Cortex* 24(8):1979-1987.
- Phan KL, Wager T, Taylor SF, & Liberzon I (2002) Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16(2):331-348.

Post-scan: At the end of the study, participants completed standardized questionnaires assessing individual differences in affective responding (BDI; (42), STAI; (43)) and perspective-taking (IRI; (44)). None of these individual-difference measures were significantly correlated with our dependent variables of interest (PINES scores, amygdala activity, affect ratings) nor did they moderate the effect of Perspective (or the interaction of Perspective and Valence) on these DVs. In light of this, they are not discussed in results section.

Acknowledgements.

M. Gilead wishes to thank M. Perkas, L. Chang, B. Dore, C. Helion, J. Shu, B. Denny, and R. Martin for their assistance. M. Gilead is supported by fellowships from the Fulbright and Rothschild foundations. Completion of the manuscript was supported by grants AG043463 from NIA, HD069178 from NICHD and MH090964 from NIMH awarded to K. Ochsner.

- Stillman PE, Van Bavel JJ, & Cunningham WA (2015) Valence Asymmetries in the Human Amygdala: Task Relevance Modulates Amygdala Responses to Positive More than Negative Affective Cues. *Journal of Cognitive Neuroscience* 27(4):842-851.
- Chang LJ, Gianaros PJ, Manuck SB, Krishnan A, & Wager TD (2015) A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect. *PLoS Biol* 13(6):e1002180.
- Breiter HC, et al. (1996) Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17(5):875-887.
- Cunningham WA & Brosch T (2012) Motivational Salience: Amygdala Tuning From Traits, Needs, Values, and Goals. *Current Directions in Psychological Science* 21(1):54-59.
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, & Barrett LF (2012) The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences* 35(3):121-143.
- Whalen PJ & Phelps EA (2009) *The human amygdala* (Guilford Press).
- Friston KJ, et al. (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6(3):218-229.
- Goldman AI (2005) Imitation, Mind Reading, and Simulation. *Perspectives on Imitation: Imitation, human development, and culture*:79.
- Epley N, Keysar B, Van Boven L, & Gilovich T (2004) Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology* 87(3):327-339.
- Gordon R (1986) Folk psychology as simulation. *Mind and language* 1(2):158-171.
- Ryle G (1949) *The Concept of Mind*.
- Fairhall SL & Caramazza A (2013) Brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience* 33(25):10552-10558.
- Binder JR, Desai RH, Graves WW, & Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19(12):2767-2796.
- Honey CJ, Thompson CR, Lerner Y, & Hasson U (2012) Not Lost in Translation: Neural Responses Shared Across Languages. *Journal of Neuroscience* 32(44):15277-15283.
- Denny BT, Kober H, Wager TD, & Ochsner KN (2012) A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience* 24(8):1742-1752.
- Gogtay N, et al. (2004) Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America* 101(21):8174-8179.
- Raio CM, Orederu TA, Palazzolo L, Shurick AA, & Phelps EA (2013) Cognitive emotion regulation fails the stress test. *Proceedings of the National Academy of Sciences of the United States of America* 110(37):15139-15144.
- Coccaro EF, Sripada CS, Yanowitch RN, & Phan KL (2011) Corticolimbic Function in Impulsive Aggressive Behavior. *Biological Psychiatry* 69(12):1153-1159.
- Lang PJ, Bradley MM, & Cuthbert BN (1997) International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*:39-58.
- Beck AT, Steer RA, & Garbin MG (1988) PSYCHOMETRIC PROPERTIES OF THE BECK DEPRESSION INVENTORY - 25 YEARS OF EVALUATION. *Clinical Psychology Review* 8(1):77-100.
- Marteau TM & Bekker H (1992) THE DEVELOPMENT OF A 6-ITEM SHORT-FORM OF THE STATE SCALE OF THE SPIELBERGER STATE TRAIT ANXIETY INVENTORY (STAI). *British Journal of Clinical Psychology* 31:301-306.
- Davis MH (1983) MEASURING INDIVIDUAL-DIFFERENCES IN EMPATHY - EVIDENCE FOR A MULTIDIMENSIONAL APPROACH. *Journal of Personality and Social Psychology* 44(1):113-126.

749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816