Minds of Monsters: Scary Imbalances between Cognition and Emotion

Ivan Hernandez

Virgina Tech

Ryan S. Ritter

Facebook

&

Jesse L. Preston

University of Warwick

Abstract

Four studies investigate a *fear of imbalanced minds* hypothesis: that threatening agents perceived to be relatively mismatched in capacities for *cognition* (e.g., self-control and reasoning) and *emotion* (e.g., sensations and emotions) will be rated as scarier and more dangerous by observers. In ratings of fictional monsters (e.g., zombies, vampires) targets seen as more imbalanced between capacities for cognition and emotion (high cognition-low emotion or low cognition-high emotion) were rated as scarier compared to those with equally matched levels of cognition and emotion (Studies 1 & 2). Similar effects were observed using ratings of scary animals (e.g., tigers, sharks; Studies 2 & 3), and infected humans (Study 4). Moreover, these effects are explained through diminished perceived control/predictability over the target agent. These findings highlight the role of balance between cognition and emotion in appraisal of threatening agents, in part because those agents are seen as more chaotic and uncontrollable.

Keywords: Mind Perception, Fear, Monster perception, Animal perception

Minds of Monsters: Scary Imbalances between Cognition and Emotion

Can a mind be scary? The last time you saw a horror movie, what was it about the psychopath serial killer, the walking dead, or the supernatural being that kept you on the edge of your seat and lingered in your imagination? There are many people and creatures— real and fictional—that scare us. One reason some agents are scarier are the obvious elements of physical threat; they could be especially strong, have sharp claws, or yield menacing weapons, for example. But other cues may be internal to a threatening agent: that is, there is something one perceives about the mind of the threatening agent that makes them scarier.

In this work we examine these scary minds – the monsters and villains we fear most (be they animal, human, or supernatural) and the characteristics of their mind that make them especially terrifying. We propose an *imbalanced mind hypothesis*, that agents appear scarier as they seem to have a mismatch in their relative cognition (e.g., reasoning and self-control) and emotions (e.g., emotions/sensations). Relative imbalances of mind may mean possessing either high cognition but low emotion, or high emotion but low cognition. Compared to targets perceived to have equal levels of cognition and emotion, those with an imbalanced mind are expected to elicit greater fear in part as they are seen as more unpredictable and uncontrollable.

Mind Perception and Scary Minds

Reasoning about threatening agents is a part of more general mind perception (Waytz et al., 2010), or "theory of mind" (Premack & Woodruff, 1978), the set of intuitive cognitive skills which allows people to readily perceive and reason about other minds. Theory of mind has clear advantages for survival: being able to "read" the minds of others makes it easier to draw inferences how others may act to help us –or hurt us (Guthrie, 1993). Some agents seem to possess more mind than others however (Haslam, 2006; Morewedge et al., 2007), and the degree

of perceived mind can have social implications for our future interactions with these agents (Waytz et al., 2010).

These judgments of mind generally fall along two dimensions: one representing an intellectual capacity for thinking and self-control, and the other reflecting an emotional capacity for sensations and feeling. These dimensions have been described as *agency* and *experience* (Gray et al., 2007), but similar distinctions are drawn between *competence* and *warmth* in the person perception literature (Fiske et al., 2002). But for our purposes we simply use the term "cognition*"* to refer to the mental capacities relating to self-control and thinking and "emotion" to refer to the affective capacity for feeling and sensation.[1] Just as agents can vary in degree of mind, agents vary along dimensions of cognition or emotion, with implications for social cognition. Targets that seem to lack capacity for emotion are judged to be cold, cruel, and robotic (Fiske et al., 2002; Loughnan & Haslam, 2007), and those agents subsequently receive greater harm when the opportunity arises (Fiske et al., 2007). On the other end, those agents perceived as lacking competence, civility and cognition are seen as mentally subhuman, licensing people to impose their own will towards them and deny them human rights (Haslam, 2006). The strength of these relationships often intensifies, when examining agent perceptions in the context of helping versus harming behaviors (Gray & Wegner, 2009).

Judgments of cognition and emotion can also play an important role in detecting threatening agents. We argue that targets who are already perceived as threatening may have their threat intensified when there is an imbalance between their relative capacities of cognition and emotion. For example, consider two scary characters from popular movies: a cyborg assassin

[1] This dichotomization of mental capacities only refers to only how people naturally perceive minds, and is not intended to indicate that mental capacities actually have a two dimension structure or that the dimensions of cognition and emotion are themselve homogenous.

(*Terminator 2*; Cameron, 1991) and a demon-possessed child (*The Exorcist*; Friedkin, 1973). The cyborg is programmed to destroy humans, and so certainly poses some threat. But being made only of liquid metal, the cyborg is relatively incapable of experiencing emotion and is completely insensitive to feelings of pain (i.e., low emotion). Yet at the same time the cyborg is intelligent, calculating, and highly capable of reasoning and problem solving (i.e., high cognition). The demon-possessed child, on the other hand, is irrational and lacking in self-control (i.e., low cognition). At the same time, she is also filled with insatiable rage, and is extremely irritable and easily agitated (i.e., high emotion).

Granted, these two movie monsters represent extreme and fictional examples of imbalanced mind in scary agents, but the same principles of mind perception apply to real-world agents. A highly emotional person could be susceptible to mood changes, be hypersensitive to triggers and sensation, and respond to emotions with erratic behavior. But if the tendency for emotion is matched by cognition, that person then has the ability to control behavior and temper emotions with rationality. Likewise, a person high in capacity for cognition may be intelligent and execute complex plans of action. Such a person could be highly dangerous if they put their mind to malicious intentions, and who's to stop them? But if their capacity for thought and self-control is matched with the capacity for feeling, they may feel greater empathy for others and be less likely to manipulate or harm others for personal gain. Greater emotional experience also introduces more vulnerability in an agent, because they can experience pain – both physical and social. Therefore, being perceived as threatening is an important condition for this phenomenon. We would not expect an imbalance between cognition and emotional capabilities to make non-threatening targets scarier because the intentions and emotions would not be directed toward causing harm.

Why should a perceived imbalance in mind elicit fear? Fear, as an emotion, is rooted in a predatory defense mechanism characterized behaviorally by the fight-or-flight response (Öhman, 1986). Agents that stick out from the norm may be attributed special attention and differential categories, for example, entities whose movement is too slow or fast speed are attributed less mind and seem less human (Morewedge et al., 2007). Uncanniness or near-human attributes can also feel creepy and unsettling (Gray & Wegner, 2012), and supernatural agents are more memorable if they are "minimally" counterintuitive, that is if they violate just a couple of categories, such as being invisible (Norenzayan et al., 2006). But beyond simply seeming abnormal or unusual, we suggest that imbalance between cognition and emotion is scary because it suggests a decreased sense of control and predictability over that agent. Fear is experienced most when there is also a lack of control or a sense of uncertainty, i.e., situations which are unpredictable (Ellsworth & Scherer, 2003; Smith & Ellsworth, 1985). Motivation to feel in control also applies to our social interactions as we seek the cause of others' behavior (Kelley, 1973), and relying on stable traits to make others' actions more predictable (Ross, 1977). Predictability is also key to mind perception and social cognition, indeed one of the basic functions of theory of mind is to predict others' future actions through their desires and intentions. But agents perceived to have greater capacities for emotion than for cognition seem to lack capacities for basic self-control and planned action, making their future actions completely unpredictable. Likewise, targets who possess greater cognitive capacity than emotional capacity seem uncontrollable in that they are highly capable, but lack empathy or emotional vulnerability. Agents that seem imbalanced in cognition and emotion may therefore trigger a fight or flight response because they seem to be out of control and unpredictable.

The Present Research

In four studies we tested the fear of imbalanced minds hypothesis, where other agents become scarier as they seem to have more asymmetry in their relative capacities for cognition and emotion. We predict that targets perceived to have an imbalanced mind (defined as mismatched capacities for cognition and emotion) will elicit more fear than targets perceived to have a relatively more balanced mind (i.e., matched capacities for cognition and emotion). This imbalance includes those with high capacity for cognitive control but low emotion (e.g., the Terminator) and those with low capacity for cognitive control but high emotion (e.g., possessed child). Further, we predict that this effect is explained by decreased feelings of predictability and control over agents with imbalanced minds.

We also note that an agent must at least be perceived as a potential threat to elicit fear, and we would not expect imbalance in mind to make an agent scarier if there is no potential physical threat posed. We are specifically interested in the minds of agents who may act to harm us in some way (the monsters and villains) and not concerned here with the benign and benevolent agents who are fundamentally unscary (e.g., puppies and babies). But within agents who pose a potential threat, the scariest agents are expected to be those with the most "imbalanced mind" that is, asymmetric capacities for cognition and emotion.

These predictions were tested using ratings of target agents, including fictional monsters (Studies 1 & 2), and dangerous animals (2 & 3). Study 3 extends findings beyond self-report to include linguistic associations found within natural discourse amongst descriptions of animals. Study 4 manipulates descriptions of cognition and emotion among a series of human targets with a fictional brain infection that affect either emotion and/or cognition. Across all studies, we

predicted that both between participants (i.e., ratings within scary targets) and between scary targets, imbalanced mind perception would be positively associated with fear.

STUDY 1: MONSTER PERCEPTION

As an initial investigation of the fear of imbalanced minds, Study 1 used ratings of various fictional monsters and creatures depicted in popular stories, books, and movies (e.g., *Frankenstein's monster, vampires, zombies)*. Targets were rated on their perceived cognitive and emotional capacity, as well as how scary they were. Our primary prediction was that greater perceived imbalance between capacities for cognition and emotion in a target would predict greater fear of that target. We report all manipulations, measures, and exclusions in these studies. All materials, codebooks, analyses/outputs, power calculations, and data can be found in the paper's associated Open Science Framework repository.[2] No study analysis pre-registration exists for this project.

Method

Participants

One hundred nine undergraduate students (57 men, 52 women; mean age = 20 years) from a large Midwestern university completed the survey for partial course credit. We collected across the entirety of a semester to obtain the sample. To justify the feasibility of the sample size to test the hypotheses, we used the *simr* package in *R* to conduct power-analysis for multilevel studies. We found that our sample size had 99% power to detect the presumed between-agent effect, and 81% power to detect the between participant-effect.[3]

---

[2] https://osf.io/x3s29/?view_only=5adfb2cdc9c848d8be5bb2576d40fde6

[3] As associations between perceptions of mind dimensions and perceptions of unease are highly variable (Gray & Wegner, 2012), we assumed the relationship at the between-agent level would be similar to the median effect size in social psychology (r = .21, Richard et al., 2003). Between-participant effect sizes are likely smaller, due to disaggregation, and we assumed an effect size roughly half of the between-agent effect (r = .10).

Procedure

Participants were seated in a private lab room in front of a computer and asked to think about a series of scary movie characters. The instructions asked them to imagine that the characters are real when making their judgments. After viewing a picture and description of each of the targets, participants rated the extent that each one exhibits capacities for cognition and emotion. Participants also rated how afraid they are of each target. At the end of the survey, participants provided demographic information.

Materials and Measures

Targets

The study presented participants with eighteen[4] movie monsters: *an alien, Bigfoot, The Blob, a "classic" zombie, a demon, a flying monkey, Frankenstein's monster, a ghost, Godzilla, an "infected" zombie, King Kong, a mummy, a demon-possessed human, Predator, a troll, a vampire, a werewolf,* and *a Xenomorph.* A representative picture and brief description were provided to introduce each target.

Target Ratings

Participants answered five questions for each target using 7-point scales ranging from 1 (*not at all*) to 7 (*very much*). Capacities for cognition and emotion were assessed with the items "How capable is the character of thoughts and intentions?" and "How capable is the character of emotions and feeling?" respectively. Fear was assessed with the single item "How afraid would you be of this character?" And finally we also measured perceived "strangeness" of each agent.

[4] We chose the large enough sample of agents that we felt had broad enough familiarity to a wide-population, and were culturally prevalent.

Results

Mean target ratings are presented in Table 1 and visualized in Figure 1. Of the 1,962 possible observations in the data set (109 participants x 18 scary targets), 21 were missing as a result of some participants failing to respond to all the targets, leaving 1,941 observations included in the final analysis. Perceived strangeness of the agent was uncorrelated with fear (r = .01), thus we do not believe the results could be attributable to people interpreting the "fear" question as synonymous with "unusual."

Imbalanced Mind

We generated an *imbalanced mind* score from the cognition and emotion ratings by calculating an absolute deviation score (i.e., the absolute value of the difference between the cognition and emotion items). Higher values on this scale thus indicate that the target was perceived to have greater imbalance in mind. To test the hypothesis that imbalanced mind perception is associated with increased fear both as a within-subject and between-subject effect, we fit a mixed-effects model (using restricted maximum likelihood estimation and Kenward-Roger approximation for the degrees of freedom), treating both stimuli and participants as random effects:

$$\text{Fear}_{ij} = \alpha_0 + \alpha_1\text{Imbalanced}_{ij} + \alpha_2\text{Imbalanced}_{j} + \mu_{0i} + \mu_{0j} + \varepsilon_{ij}$$

where i refers to the specific participant and j refers to the specific scary target. This model thus allows us to investigate the effect of imbalanced mind perception on fear between participants (within scary targets; α_1) and the effect of imbalanced mind perception on fear between scary targets (α_2). The between participants predictor was created by centering around each scary target mean, and the between target predictor was created by centering around the grand mean (Enders & Tofighi, 2007).

Consistent with our predictions, participants who perceived a scary target to have a more imbalanced mind also reported greater fear of that target (b = 0.192, β =.135, t_{1883} = 7.686, p < .0001, 95% CI= [0.100 0.169]). Moreover, scary targets perceived to have a more imbalanced mind overall elicited greater fear overall (b = 1.299, β =.290, t_{16} = 3.241, p = .005, 95 % CI = [0.114, .467]).

Main Effects of Cognition and Emotion

We also tested for independent effects of cognition and emotion on fear by entering the predictors representing cognition and emotion alone into a mixed-effects model. Cognition was positively associated with fear between participants (b = 0.246, β = 0.227, t_{1885} = 10.363, p < .0001, 95% CI = [0.184, .270]) and marginally so between scary targets (b = 0.669, β = .407, t_{15} = 1.974, p = .067, 95% CI = [0.015, .0.798]). Emotion was negatively associated with fear between participants (b = -0.094, t_{1890} = -3.758, p = .0002, β= -0.083, 95% CI = [-0.126, 0.040]), but there was no reliable association between scary targets (b = -0.582, β = -.334, t_{15} = -1.620, p = .126, 955 CI=[-.725, 0.058]). However indices of imbalanced mind remained statistically significant predictors of fear ($b_{Between\ Participants}$ = 0.082, β = .058, t_{1865} = 2.547, p = .011, 95% CI = [0.017, 0.133]; $b_{Between\ Targets}$ = 1.797, β = 0.402, t_{14} = 2.271, p = .039, 95% CI =[0.196, 1.975]) even after statistically controlling for the independent effects of cognition ($b_{Between\ Participants}$ = 0.197, β = 0.181, $t_{1867.92}$ = 6.425, p < .0001, 95% CI =[0.097, .0.226]; $b_{Between\ Targets}$ = -0.419, β = -.254, t_{14} = -0.741, p = .471, 95% CI =[-1.345, 0.440]) and emotion ($b_{Between\ Participants}$ = -0.050, β = -.044, $t_{1867.92}$ = -1.653, p = .098, 95% CI = [-0.025 , 0.005]; $b_{Between\ Targets}$ = 0.412, β = .24, t_{14} = 0.763, p = .458, 95% CI = [-0.354, 0.315]),

Effect of Overall Mind

For agents to have an imbalanced mind, there must be some base-level of mind–one dimension must be high and another must be low. Therefore, an alternative explanation is that

agents who have more mental capacity overall may be scarier, which imbalance is simply confounded with. This property is a common limitation expressed with difference scores (for a review see Griffin et al., 1999). Therefore we sought to demonstrate that the effect remains, even after controlling for total mind. Using the sum of the Gray et al. (2007) items has previously been used as a measure of total mind (Lefkeli et al., 2021, Morewedge et al., 2007). Therefore, higher values on this scale thus reflect that a target was perceived to possess more mental capacity overall.

When entered into a mixed-effects model alone, the composite measure of mind was positively associated with increased fear between participants (b = 0.163, β = .132, t_{1898} = 7.307, p < .0001, 95% CI = [0.097, 0.168]), but there was no reliable association between scary targets (b = 0.127, β = .072, t_{16} = 0.632, p = .536, 95% CI = [-0.152, 0.296]). But importantly, the effects of imbalanced mind perception remained statistically significant ($b_{Between\ Participants}$ = 0.171, β = .121, t_{1882} = 6.954, p < .0001, 95% CI =[0.087, 0.156]; $b_{Between\ Targets}$ = 1.293, β = .289, t_{15} = 3.042, p = .008, 95% CI =[0.108, 0.470]), after statistically controlling for the effect of the absolute perception of mind between participants (b = 0.145, β = .118, t_{1897} = 6.535, p < .0001, 95% CI = [0.082, 0.153]) and between scary targets (b = 0.01, β = .0127, t_{15} = 0.076, p = .940, 95% CI =[-.174, 0.188]),

Study 1 Discussion

In Study 1, fictional monsters from popular culture were rated as scarier if they seemed to have more imbalance in relative cognition and emotion (e.g., demon, werewolf) compared to monsters more "balanced" in mind (e.g., the Blob, Bigfoot). The effect was also found within-targets: among ratings for the same target monster, those who perceived more imbalance also reported stronger fear. These effects were not explained by perceived "strangeness" of the target and the imbalance between cognition and emotion was a better predictor of fear than absolute

perceptions of mind, or the independent effects of cognition or emotion. These findings illustrate the role of mind perception in fear, and that perceived imbalances between cognition and emotion are perceived as scary.

STUDIES 2A & 2B: MONSTERS AND BEASTS

Studies 2A and 2B extended results of Study 1 to include imbalanced minds in both fictional monsters and threatening animals. Social fears towards animals are considered part of the same general evolutionary threat management system; therefore, we expect the effect of imbalance to generalize across threatening domains (Öhman, 1986). Study 2A adopts a within-subject design with subjects rating multiple targets, and Study 2B used a between-subject design where subjects were randomly assigned to rate one of 11 monsters or 11 animals. In both kinds of designs, we predicted that perceptions of an imbalanced mind would be positively associated with fear.

Method

Participants

For Study 2A, eighty participants (34 men, 46 women; mean age = 33 years) were recruited online from the United States through Amazon Mechanical Turk (MTurk). No participants were excluded from the analysis. This sample was determined by available funding, however, we conducted a power analysis using the effect sizes (r = .29) found in Study 1 as the presumed effect for a multilevel study with 80 participants evaluating 11 agents each. The power analysis finds 99% power for detecting the between-target effect.

For Study 2B, one hundred thirty-five undergraduate students (57 men, 78 women; mean age = 19 years) were recruited from a large Midwestern university, and 186 participants (69 men, 117 women; mean age = 34 years) were recruited online from the United States through MTurk.

No data were excluded from the analysis for a total sample size of 321. We conducted a power analysis that assumed the same effect sizes used in the Study 2A power analysis, finding 84% power to detect the between-target effect.

Procedure

As in Study 1, participants in Study 2 were asked to rate targets on their mental capacities, and that they should rate each one as though it were real. In Study 2A, Participants were randomly assigned to view and rate either 11 supernatural targets (n = 35) or 11 animal targets (n = 45). After viewing a picture and description of each of the targets as an introduction, participants rated the extent to which each one exhibits capacities for cognition and emotion, and how afraid they are of each target. The order of presentation of the 11 targets was randomized across participants.

The procedure for Study 2B was similar to Study 2A with three important exceptions. First, participants were introduced to only a single target: one of 11 supernatural targets or one of 11 animals, randomly assigned. Second, participants rated the target's capacity for cognition and emotion using the 18 different mental capacities found by Gray et al. (2007). Finally, we also included an additional item in Study 2B to rate how dangerous the target is perceived, diversifying the outcomes measured.

Materials and Measures

Targets

In Study 2A, participants were randomly assigned to rate either monsters or animals. In the monster condition, participants rated 11 targets: *an alien, The Blob, a classic zombie, the devil, Frankenstein's monster, a ghost, an infected zombie, a mummy, a demon-possessed human, a vampire,* and a *werewolf.* In the animal condition, the 11 targets were a: *bat, bear,*

cobra, crocodile, gorilla, lion, great white shark, tarantula, tiger, swarm of wasps, and *wolf*. In Study 2B, participants were randomly assigned to view and rate only one of the 11 supernatural targets or one of the 11 animals.

Target Ratings

In this and all subsequent studies, participants rated each target on multiple measures of cognition and emotion (informed by Gray et al., 2007), using a relative 7-point scale with a normal adult human as the reference point (-3 = *much less than the average human*, 3 = *much more than the average human*).

In Study 2A, capacities for cognition were assessed with two items: "How capable is the character of thinking and planning actions?" and "How capable is the character of self-control?" Capacities for emotion were also assessed with two items: "How capable is the character of experiencing emotions?" and "How capable is the character of feeling pleasure and pain?" Fear was assessed with the single item "How afraid of this character would you be, if you encountered it when by yourself?" on a scale ranging from 1 (*not at all*) to 7 (*very much*).

In Study 2B, capacities for cognition were assessed with 7 items: *self-control, morality, memory, emotion recognition, planning, communication*, and *thought*. Capacities for emotion were assessed with 11 items: *hunger, fear, pain, pleasure, rage, desire, personality, consciousness, pride, embarrassment*, and *joy*. Fear was assessed using two items: "If you encountered them, when by yourself, how afraid of [*target's name*] would you be?" and "How dangerous is [*target's name*]?" on respective scales ranging from 1 (*not at all*) to 7 (*extremely*). The measures of fear and danger were highly correlated, $r(319) = .54$, $p < .001$, and we computed their mean to form a single composite fear dependent variable.

Study 2A Results

Mean target ratings are presented in Table 2 and visualized in Figure 2. Of the 880 possible data points in the dataset, 16 were missing, leaving 864 observations included in the final analysis.

Effect of Imbalanced Mind

To test our main hypothesis, we generated an *imbalanced mind* score by computing the mean of the two cognition items (mean within-target correlation = .66, *SD* = .17), the mean of the two emotion items (mean within-target correlation = .72, *SD* = .12), and then calculating the absolute value of the difference between these measures of cognition and emotion.

As in Study 1, we fit a mixed-effects model to test the hypothesis that perceptions of an imbalanced mind would be positively associated with fear both between participants (within scary targets) and between scary targets. As predicted, participants that perceived a scary target to have a more imbalanced mind also reported greater fear of that target (b = 0.064, β = 0.173, t_{814} = 2.332, p = .020, 95% CI =[0.010, 0.118])[3], and scary targets perceived to have a more imbalanced mind overall elicited greater fear overall (b = 1.212, β = 0.173, $t_{20.8}$ = 2.060, p = .052, 95% CI = [0.007, 0.338]). Adding a scary target category (natural = -.5, supernatural = .5) revealed no main effect (p = .45) or interactions with between target imbalance (p = .27) or with between participant imbalance (p =0.86), suggesting that the effect of imbalanced mind was consistent across ratings of both natural and supernatural targets.

Independent Effects of Cognition and Emotion

When entering all within and between subjects simultaneously, cognition was positively associated with fear between participants (b = 0.101, β = 0.084, t_{838} = 2.217, p = .027, 95% CI=

[0.009, 0.159]), but there was no reliable association between targets (b = -0.119, β = -0.058, $t_{19.3}$ = -0.354, p = .727, 95% CI= [-0.374, 0.258]). Emotion was not associated with fear ($b_{Between\ Participants}$ = -0.018, β = -0.015, t_{828} = -0.412, p = .680, 95% CI= [-0.087, 0.057]; $b_{Between\ Targets}$ = 0.417, β = .198, $t_{19.2}$ = 1.173, p = .255, 95% CI= [-0.126, 0.521]).

After statistically controlling for the effects of cognition ($b_{Between\ Participants}$ = 0.097, β = 0.081, t_{836} = 2.145, p = .032, 95% CI= [0.004, 0.158]; $b_{Between\ Targets}$ = -0.181, β = -0.088, $t_{18.3}$ = -0.542, p = .595, 95% CI= [-0.115, 0.210]) and emotion ($b_{Between\ Participants}$ = -0.024, β = -0.020, t_{826} = -0.553, p = .581, 95% CI= [-0.090, 0.035]; $b_{Between\ Targets}$ = 0.293, β = .139, $t_{18.2}$ = 0.813, p = .427, 95% CI= [-0.186, 0.138]), the effect of imbalanced mind perception remained statistically significant between participants (b = 0.106, β = 0.061, t_{808} = 2.191, p = .0287, 95% CI= [0.010, 0.167]) but not between scary targets (b = 1.00, β = 0.142, t_{20} = 1.312, p = .205, 95% CI= [-0.201, 1.178]).

Effect of Overall Mind

We generated a *composite mind* score as in Study 1. When entered into a mixed-effects model alone, there was no statistically reliable association between total mind perception and fear ($b_{Between\ Participants}$ = 0.077, β = 0.059, t_{832} = 1.784, p = .074, 95% CI = [-0.006, 0.123]; $b_{Between\ Targets}$ = 0.295, β = .137, $t_{20.2}$ = 1.551, p = .137, 95% CI= [-0.037, 0.312]). After statistically controlling for the effect of the overall perception of mind ($b_{Between\ Participants}$ = 0.069, β = 0.052, t_{832} = 1.578, p = .1149, 95% CI=[-0.013, 0.117]; $b_{Between\ Targets}$ = 0.096, β = .045, $t_{19.3}$ = 0.415, p = .682, 95% CI=[-0.162, 0.251]), the effect of imbalanced mind perception remained statistically significant between participants (b = 0.105, β = 0.060, t_{810} = 2.165, p = .030, 95% CI=[0.006, 0.114]) but not between scary targets (b = 1.027, β = 0.146, t_{20} = 1.374, p = .185, 95% CI=[-0.058, 0.351]).

Results: Study 2B

Mean target ratings are presented in Table 3 and visualized in Figure 3. The number of participants rating each target ranged from 7 to 21.

Effect of Imbalanced Mind

We created a composite cognition score using the mean of the 7 cognition items (Cronbach's α = .86), and 11 emotion items to create a composite emotion score (Cronbach's α = .78). An *imbalanced mind* score was generated as in previous studies.

Consistent with our prediction, a mixed-effects model showed that imbalanced mind perception was associated with increased fear both between participants (b = 0.268, β = .124, t_{298} = 2.50, p = .013, 95% CI = [0.061, 0.424]), and between scary targets (b = 1.064, β = .23, $t_{19.7}$ = 2.23, p = .037, 95% CI = [0.061, 0.424]). Adding a scary target category (natural = -.5, supernatural = .5) revealed no statistically significant main effects or interactions (all p-values > .300), suggesting that the effect of imbalanced mind perception on fear was similar across different categories of targets.

Independent Effects of Cognition and Emotion

Perceived cognition was not associated with fear between participants (b = -0.120, β = -.084, t_{298} = -1.291, p = .198, 95% CI = [-0.120, 0.035]), and negatively associated with fear between scary targets (b = -0.616, β = -.333, t_{20} = -3.040, p = .006, 95% CI = [-0.701, -0.092]). Perceived emotion was also not associated with fear between participants (b = -0.017, β = -.010, t_{298} = -0.153, p = .879, 95% CI = [-0.130, 0.144]), and positively associated with fear between scary targets (b = 0.917, β = .365, $t_{19.5}$ = 3.27, p = .004, 95% CI = [0.218, 0.785]).

After statistically controlling for the effects of cognition ($b_{Between\ Participants}$ = 0.004, β = .003, t_{296} = 0.039, p = .969, 95% CI= [-0.079, 0.144]; $b_{Between\ Targets}$ = -0.638, β = -.345, $t_{19.9}$ = -2.20, p = .040, 95%

CI = [-0.543, -0.054]) and emotion ($b_{Between\ Participants}$ = -0.180, β = -.102, t_{296} = -1.395, p = .164, 95% CI = [-0.092, 0.041]; $b_{Between\ Targets}$ = 0.945, β = .376, $t_{19.7}$ = 2.418, p = .025, 95% CI = [0.153, 0.665]), the effect of imbalanced mind perception remained statistically significant between participants (b = 0.32, β = .150, t_{296} = 2.61, p = .01, 95% CI = [0.044, 0.262]) but not between scary targets (b = -0.07, β = -.014, $t_{19.4}$ = -0.107, p = .916, 95% CI = [-0.714, 0.237]).

Effect of Overall Mind

The mean of the cognition and emotion scores was computed to create an *overall mind* score. Perceiving more mind overall was marginally associated with *less fear* between participants (b = -0.146, β = -0.082, $t_{298.10}$ = -1.652, p = .099, 95% CI = [-0.129, 0.015]), and there was no association between scary targets (b = 0.077, β = 0.031, $t_{19.92}$ = .279, p = .783, 95% CI = [-0.264, 0.251]). After statistically controlling for the effect of the overall perception of mind ($b_{Between\ Participants}$ = -0.161, β = -0.091, $t_{297.03}$ = -1.837, p = .067, 95% CI = [-0.071, 0.005]; $b_{Between\ Targets}$ = 0.089, β = 0.036, $t_{18.82}$ = 0.348, p = .732, 95% CI = [-0.224, 0.122]), the effects of imbalanced mind perception remained statistically significant ($b_{Between\ Participants}$ = 0.281, β = 0.130, $t_{297.03}$ = 2.622, p = .0092, 95% CI = [0.029, 0.277]; $b_{Between\ Targets}$ = 1.066, β = .227, $t_{18.7}$ = 2.188, p = .042, 95% CI = [0.064, 0.806]).

Summary

Studies 2A and 2B examined fear of imbalanced minds with both fictional and real creatures as targets, with both within and between-subjects designs. As predicted, targets (whether animal or supernatural) who were perceived as having more imbalance between cognition and emotion elicited greater fear than targets with a balance between cognition and emotion. This extended the effect of Study 1 by demonstrating that both fictional monsters and real animals perceived to have imbalanced minds were perceived to be scarier than targets with relatively more balanced minds.

STUDY 3: IMBALANCED MIND IN NATURAL LANGUAGE

Study 3 extended our study of the imbalanced minds hypothesis with natural language models to detect linguistic evidence that animals perceived to have more imbalanced minds are perceived as more frightening. Recent research using natural language models has shown the ability to capture associations between concepts by studying the context of words within text (e.g., "King" and "Queen," shares a similar context because they are both surrounded by words such as "throne," and "royalty."). "King," however, also shares a masculine context, surrounded by words like "he" and "his," whereas "Queen" has a feminine context. Compared to simpler embedding models, Natural Language Inference (NLI) models (Geiger et al., 2020) more directly learn target-concept associations based on the implied properties of concepts (i.e., "A Queen is always a woman, but a woman is not always a Queen". Study 3 leverages these models, which are trained on natural discourse from internet sources to analyze the linguistic relationship between imbalanced minds (whether an animal is associated with imbalanced cognition and emotion) and fear of the animal.

Procedure

Data Sources

Data used to pre-train the NLI model were collected from the Book Corpus (Zhu et al., 2015), the 2019 version of English Wikipedia, the Giga5 news dataset, the ClueWeb 2012-B web crawl, and the Common Crawl, collectively representing 200 GB of text and over 33 billion words. To fine-tune the model to understand logical inferences the model was trained on the FEVER dataset of 145,449 pairs of hypotheses and premises extracted from Wikipedia (Thorne et al., 2018), and the Adversarial Natural Language Inference dataset comprising of 10,776 pairs

of statements extracted from Wikipedia and independently annotated as entailments, contradictions, or neutral (Nie et al., 2020). See Supplementary Materials for more details.

Agents

We isolated the analysis only to animals because we are interested in comparing both threatening and non-threatening targets and because prior research on large samples ($N > 2{,}000$) exists on fear ratings of animals to validate the approach used. We used the list of all 735 animals listed on animals.net (Animals Network Editors, 2017). To minimize artificially inflating statistical power from redundancy, we only used the terminal word in an animal's name (e.g., "American crow" is simplified to "crow") resulting in 204 unique animal names. A power analysis that assumed the between-agent effect previously observed is driven entirely by threatening agents (average $\beta=.23$ for threatening agents, and $\beta=0$ non-threatening agents, determined that this study has 97% power to detect the effect.

Natural Language Inference Model

We used the Adversarial XLNet Natural Language Inference model (Nie et al., 2020) as the primary model. XLNet is one of the most performant neural network architectures for drawing inferences from language (Yang et al., 2020). It is pretrained using the Masked Language Modeling task, which extracts sentences from a variety of natural language sources (e.g., Wikipedia, books), and "masks" one of the words. The model then tries to use the surrounding context from the masked word to predict what the missing word is. By performing this task and updating its internal weights based on how it can predict the missing word, the model learns the interchangeability of words and phrases. This model is then finetuned on natural language specific tasks to leverage its generalized semantic understanding (e.g., snow is associated with cold) to a more focused understanding of the logical relationships between

concepts (e.g., snowing implies cold weather; cold weather does not necessarily imply snowing).

Calculating Attribution Score. We calculated the implied association between an attribute and an agent by specifying the agent in the premise, "I am a [e.g., tiger]", and then specifying the attribute as the hypothesis term in the model, "I am [e.g., frightening]." (Figure 4). The probability that the hypothesis contradicts the premise was subtracted from the probability that the hypothesis is entailed with the premise. This difference serves as the attribute score of the natural language model for that attribution for that agent. We repeated this process for each attribute, for each agent, just as a human would rate all targets on cognition, emotion, and fear items.

Cognition and Emotion Scores. We calculated each animal's score on the 18 cognition and emotion items used by Gray et al. (2007). We created composite cognition and emotion scores by applying the same Principal Components Analysis procedure used by Gray et al. to the 204 animal x 18 cognition/emotion item matrix.

Fear Scores. We calculated each animal's score on four different fear attributes: "scary," "frightening," "terrifying," and "unnerving." The ratings on all of the attributes showed high internal consistency ($\alpha = .96$)

Threat Scores. Our proposed moderator of "threat" was also calculated using the composite of ratings for threat-related words relevant to animal behavior (i.e., "threatening," "aggressive", "attacking," "not playful", and "not friendly"; $\alpha = .79$). These five terms are the exact ones used in the "Aggression" subscale of the DPQ, which measures animal personality (Jones, 2008).

Natural Language Inference Validation. We validated the model's inferences of each of the variables of interest using archival data collected from large international samples. The

model's inferences correlated strongly (i.e., near or above conventional reliability thresholds) with targets' cognition (r=0.989, p < .0001, 95% CI =[.687, .969]), emotion (r = 0.881, p < .0001, 95% CI=[0.641, 0.964]), and attributed fear (r=0.669, p < .0001, 95% CI=[0.364,0.844]) that were derived from large scale surveys. In short, this convergence suggests that its attributions align with human perceptions of these attributes (See S1 for a full description of the validation).

Results

We calculated imbalance using the absolute difference between an agent's composite cognition and emotion scores. We examined a multiple regression (all measures are at the between-agent level), which included the standardized measures of threat, imbalance, and the interaction between them. We found a statistically significant interaction between threat and imbalance (b =-0.084, β =-0.084, t_{200} =-2.065, p = 0.040, 95% CI = [-0.163, -0.004]), which qualifies all lower-order effects for threat (b =0.692, β = 0.692, t_{200} = 11.189, p < .001, 95% CI =[0.570, 0.814]) and imbalance (b = 0.034, β = 0.034, t_{200} =0.604, p = .547, 95% CI = [-0.076, 0.143]). We followed the interaction with a simple slopes test. As expected, the interaction shows that the effect of imbalanced minds on fear is driven by threatening agents. For agents the natural language inference model deemed threatening (threat score > 0, n =107), more perceived mind imbalance of an agent was associated with greater fear towards that animal (r = 0.207, p =.032, 95% CI = [0.018, 0.382]). However, for agents that the natural language model deemed less-threatening (threat score < 0 , n =97), there was no association between an agent's imbalanced mind and its scariness (r = -0.072, p = .482, 95% CI = [-0.268, 0.129]).

Summary

Study 3 shows that associations between imbalanced minds and fear are also found in the semantic associations of natural language models trained on billions of examples of human discourse. Important, this relationship is only found for threatening targets. This study highlights the embeddedness of this phenomenon, finding the effect extends to usage within natural language.

STUDY 4: FEAR OF IMBALANCED MINDS HUMANS

So far we focus exclusively on non-human targets (animals or supernatural humans) to find imbalance in capacities for cognition and emotion elicits greater fear. We focus on non-humans as these are stereotypically scarier than humans. But non-humans are also generally seen to have less mind (Morewedge et al., 2007) and capacity for complex emotional experience (Haslam, 2006), which could impact the meaning of a "balanced" mind in these targets. Study 4 addressed this issue by studying imbalanced mind in human targets. To create the perceived imbalances, participants read about a hypothetical brain virus that either increased or decreased cognition and increased/ decreased emotional capacity in infected persons. We predicted that people would rate the infected humans as scarier in scenarios where the virus results in an imbalance between cognition and emotion (high-high; or low-low), rather than equal levels of cognition and emotion (high-low; low-high). We also included additional items to assess perceived control over the target as an expected mediator between imbalance and fear.

Method

Participants

One hundred participants (39 men, 61 women; mean age = 36 years) were recruited online from the United States through MTurk. The sample size was determined by funding

availability and power analysis using the average between-target effect found in the prior studies suggests that this sample obtains a 74% chance of rejecting the null hypothesis for the expected interaction

Procedure and Target Ratings

Participants responded to a hypothetical scenario describing a virus where infected humans undergo changes to mind (corresponding to four levels of balanced and imbalance in mind) and are subsequently driven to harm others. Participants were first provided with brief definitions of the concepts of cognition and emotion, and were asked to describe them in their own words. Participants were then presented with a hypothetical scenario describing a viral outbreak, where, after a human is infected, they go into a coma, experience behavioral transformations, and are subsequently driven to harm others. Using a 2 x 2 repeated measures design, participants then read four different scenarios where the brain virus affected the mind in different ways: possessing more/less capacity for emotion and sensation, and more/less capacity for reason and intelligence (see Appendix). We randomized the order of the descriptions across participants.

Participants responded to six questions for each infected person. As a manipulation check, the first two questions asked participants to rate how much the infected person exhibits capacities for cognition and emotion, respectively, on the same 7-point scales used in Studies 2 & 3. The next question served as our proposed mediator, and asked, "How much control could you have over this infected person?" (0 = *no control*, 6 = *complete control*). The next two questions asked "How afraid of this infected person would you be, if you encountered them, when by yourself?" and "How dangerous is this infected person?" (0 = *not at all*, 6 = *extremely*). Finally, participants indicated on a sliding scale the absolute minimum distance they would be

willing to tolerate between themselves and the infected person (0-100 feet). As in previous studies, the measures of fear and danger were highly correlated (mean within-target correlation = .87, *SD* = .24), so we computed their mean to create a single composite measure of fear.

Results

Manipulation Check

As a test of our manipulation, we fit two mixed-effects models (treating participants as the only random factor) with ratings of cognition and emotion, respectively, regressed onto variables representing cognition (-.5 = low, .5 = high), emotion (-.5 = low, .5 = high), and their interaction. Cognition ratings revealed a large main effect of cognition (b = 4.170, β = .812, t_{396} = 27.695, $p < .0001$, 95% CI =[0.755, 0.869]) and neither the main effect of emotion (p = .595) nor the cognition x emotion interaction (p = .507) was statistically significant (Low cognition/Low emotion M = -2.11, SD = 1.53; Low cognition/High emotion M = -1.93, SD = 1.71; High cognition/Low emotion M = 2.16, SD = 1.40; High cognition/High emotion M = 2.14, SD = 1.36). Emotion ratings similarly revealed a large main effect of emotion (b = 4.215, β = .824, t_{396} = 29.107, $p < .0001$, 95% CI =[0.768, 0.879]), and no cognition x emotion interaction (p = .863). The main effect of cognition was also statistically significant with a small effect size (b = 0.355, β = 0.069, t_{396} = 2.452, p = .0147, 95% CI =[0.014, 0.125]) (Low cognition/Low emotion M = -2.24, SD = 1.33; Low cognition/High emotion M = 2.00, SD = 1.60; High cognition/Low emotion M = -1.86, SD = 1.79; High cognition/High emotion M = 2.33, SD = 0.93).

Main Analyses

We predicted that hypothetical humans manipulated to have an imbalanced mind (i.e., High cognition/Low emotion; Low cognition/High emotion) would be perceived as scarier than their balanced counterparts (i.e., Low cognition/Low emotion and High cognition/High emotion).

To test this prediction, we analyzed fear ratings using the same 2 (cognition) x 2 (emotion) model as in the manipulation check (The analyses is repeated-measures, and therefore all *N*s = 100). This analysis revealed a main effect of cognition (b = -0.523, β = -0.142, t_{297} = -3.073, p = .002, 95% CI =[-0.232, -0.051]) and a marginal main effect of emotion (b = 0.278, β = .075, t_{297} = 1.632, p = .103, 95% CI =[-0.015, 0.165]). More important, the predicted interaction between cognition and emotion was significant (b = -1.675, β = -.227, t_{297} = -4.926, p < .0001, 95% CI =[-0.317, -0.137), which is equivalent to a planned contrast showing that the two imbalanced targets were perceived as more scary than the two balanced targets (mean difference = -1.68 (-.25 + -.25) = 0.84; Low cognition/Low emotion M = 3.50, SD = 1.86; Low cognition/High emotion M = 4.61, SD = 1.30; High cognition/Low emotion M = 3.81, SD = 1.90; High cognition/High emotion M = 3.25, SD = 1.98).

A similar pattern emerged when we analyzed participants' reports of the absolute minimum distance they would be willing to tolerate between themselves and each kind of infected person. There was a main effect of cognition (b = -0.055, β = -0.115, t_{297} = -3.109, p = .002, 95% CI =[-0.188, -0.043]) and no main effect of emotion (b =, β = -0.001, t_{297} = -0.020, p = .984, 95% CI =[-0.073, 0.072]). But this analysis revealed the predicted interaction between cognition and emotion (b = -22.970, β = -.155, t_{297} = -4.169, p < .0001, 95% CI =[-0.227, -0.082]), demonstrating that participants preferred more distance between themselves and the imbalanced targets than the balanced targets (mean difference = -22.97(-.25 + -.25) = 11.49 feet; Figure 5).

Mediation

So far, we have replicated previous findings that imbalanced minds are scarier than balanced minds. We next investigate the role of perceived controllability on this effect. We expected the imbalanced targets to be perceived as relatively less controllable than the balanced

targets, and in turn, less control should be associated with increased fear towards that agent. We found support for this hypothesis. First, analysis of the control ratings revealed a statistically significant main effect of emotion (b = -0.400, β = -0.109, t_{297} = -2.349, p = .019, 95% CI =[-0.200, -0.018]) and no main effect of cognition (b = -0.160, β = -0.044, t_{297} = -2.349, p = .348, 95% CI = [-0.135, 0.047]). The predicted cognition x emotion interaction was statistically significant (b = 1.040, β = 0.142, t_{297} = 3.053, p = .003, 95% CI =[0.051, 0.233]). In other words, the imbalanced targets were perceived as being less controllable than the balanced targets overall (d = 0.350, 95% CI=[0.120, 0.580]; Low cognition/Low emotion M = 2.71, SD = 2.04; Low Cognition/High Emotion M = 1.79, SD = 1.77; High cognition/Low emotion M = 2.03, SD = 1.68; High cognition/High emotion M = 2.15, SD = 1.71).

To investigate whether perceptions of controllability mediated the association between mind perception and fear, we calculated the average indirect effect across participants (Bauer et al., 2006). The average total effect was estimated to be 0.842 (p < .0001; 95% CI = [0.515, 1.160] with 10,000 resamples; Figure 6). Imbalanced mind perception was associated with decreased control (b = -0.52, β = -0.142, t_{299} = -3.03 p = .003, 95% CI =[-0.234, -0.050]). Control, in turn, was associated with decreased fear (b = -0.564, β = -0.559, $t_{391.80}$ = -13.630, p < .0001, 95% CI =[-0.640, -0.13]). The average indirect effect was estimated to be 0.292 (p < .0001; 95% CI = [0.0984, 0.50], 0.97 with 10,000 resamples), leaving a direct effect of 0.544 (β = .147, $t_{299.09}$ = 3.92, p = .0001, 95% CI = [0.074, 0.810]). These results suggest that the increased fear felt towards the imbalanced targets is driven, at least in part, by their relative uncontrollability.

Summary

Participants rated four descriptions of humans who were infected with a brain virus which resulted in having either balanced or imbalanced minds. As predicted, targets presented as

having imbalanced capacities for cognition and emotion (low-high/high-low) were perceived as scarier and more dangerous than those presented as having balanced minds (low-low/high-high). The same pattern emerged with the minimum distance people wanted between themselves and each kind of infected person: participants preferred a greater distance between themselves and the imbalanced minds versus the balanced minds. Important, the perceived controllability of the target partially mediated both effects: targets with imbalanced minds were perceived as being relatively less controllable, which in turn was associated with increased fear.

General Discussion

The minds of scary agents –monsters, creatures, and supervillains—capture our attention and our collective imaginations. We suggest a common feature of mind in the scariest agents is an "imbalanced mind"—that is, asymmetric capacities for cognition (e.g., self-control and reasoning) and emotion (e.g., sensations and emotions). Four studies found support for the fear of imbalanced minds. Targets perceived as imbalanced in capacities for cognition and emotion elicited stronger ratings of fear and danger than more balanced targets. This effect was demonstrated among samples of fictional monsters (Studies 1 and 2) and animals (Studies 2 and 3) using a variety of measures of mind and fear. In Study 4, human targets experimentally manipulated to have an imbalanced mind also elicited higher ratings of fear and danger. Study 4 further demonstrated that people perceived imbalanced minds as less controllable than balanced minds, and these controllability perceptions mediate the effect of mind perception on fear. Importantly, scary minds were not consistently predicted by extremely low or high levels of cognition or emotion, but by an imbalance between the two. If extreme levels of cognition and emotion signaled a "scary mind," we should expect the greatest fear from targets either low or

high on both dimensions. But this is not what we observed. Rather, it was the imbalance between these dimensions that consistently predicted fear.

Can Imbalanced Minds be Good?

In this work we were interested in aspects of mind that signal scary agents. These threatening agents demand our attention for the simple reason they have potential to harm, and successfully identifying (and avoiding) such agents is essential to survival. But notably, an agent must at least be perceived as a *potential* threat to elicit fear, and we would not expect imbalance in mind to make an agent scarier if there is no potential threat posed. We operationalize threat as an agent capacity to harm, which could originate from a variety of sources, including mental (e.g., malevolence, aggressiveness) or non-mental (physical attributes, reflexive behaviors).

But this does raise the issue whether imbalance between cognitive and emotional capacity for mind is inherently scary. For other agents, imbalanced minds can be a reason for admiration, if it is accompanied by a benevolent intent. For example, God and Gandhi are also perceived as imbalanced (cognition > emotion; Gray et al., 2007), but these agents are also viewed as benevolent. Rather, an asymmetric attribution of cognition and emotion among benevolent targets "typecasts" them as moral exemplars (Gray & Wegner, 2009), and elicit feelings of inspiration and respect more than fear (Gray & Wegner, 2011). Important here is the perceived intention of the agent. For example, people who conceptualize God as a malevolent and punishing deity are more obedient to rules (Shariff & Norenzayan, 2011). Similarly, other positive targets that are perceived to possess relatively more capacities for emotion than cognition (e.g., children) would also not be expected to elicit fear. Their innocence (i.e., good intentions) combined with a relative inability to do anything harmful makes them completely non-threatening. Instead, they are perceived as agents worthy of moral care and consideration

(Gray & Wegner, 2009) and should elicit emotions such as sympathy (Gray & Wegner, 2011). In some instances, greater "balance" might increase feelings of unease to other agents- in the case of describing super-computers as having the ability to feel emotions (Gray & Wegner, 2012).

Conditions for Scary Minds

For agents who pose a potential threat, the scariest agents are expected to be those with the most "imbalanced mind" that is, asymmetric capacities for cognition and emotion. We focus on an imbalance between perceived cognition and emotion as the principal characteristic of mind that enhances fear, but we do not make any claims whether one type of imbalance (cognition > emotion vs. emotion > cognition) is scarier. In our regression analyses on fear effects of cognition and emotion scores were usually symmetric, but some studies found that more cognitive vs emotional agents evoked more fear, while other times the more emotion skewed agents were scarier (S2). Future research could follow up on contexts that explain-these differences for when one type would be stronger.

Agents with a greater imbalance in mind may also lead to people perceiving an agent as more threatening. If agents with a greater mental imbalance evoke more fear in others, then those agents may develop greater malevolence from those reactions. This process is consistent with self-fulfilling stereotype research where people assigned negative assumptions behave negatively towards another person, and that target then responds with negative behaviors (Synder et al., 1977). Thus, the downstream effects of subsequent interactions with agent's that have imbalanced minds, but are only mildly threatening is that they may become more threatening, making them more susceptible to eliciting fear.

Cultural Limitations

Finally, we note our samples drew exclusively from western cultures, so limits the generalizability of the findings. Cultural traditions have their own prototypes of scary agents-that could differ from those we target here. These traditions may be reflected within language. Therefore, the proposed natural language processing method may be especially suited for examining cross-cultural comparisons in future research to demonstrate whether these associations hold in other languages.

Conclusion

There are many real or imagined threatening agents that scare us. Not all scary agents are perceived equally, however, in their cognition and emotional capacities, or in how scary they seem. These studies illustrate that mind perception of scary agents shows a pattern of imbalance between cognition and emotion, in part because the target is judged to be less predictable and controllable. Our perceptions of others minds are crucial in navigating social interactions, and perceptions of an imbalanced mind is one cue that may intensify the fear elicited from those who threaten us.

References

Animals Network Editors. (2017). Animal Index. Retrieved June 14, 2022, from https://animals.net

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*(2), 142–163. https://doi.org/10.1037/1082-989X.11.2.142

Cameron, J. (1991). *Terminator 2: Judgment Day*. Tri-Star Pictures.

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In *Handbook of affective sciences* (pp. 572–595). Oxford University Press.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Friedkin, W. (1973). *The exorcist* (Extended director's cut.). Warner Bros. Entertainment;

Geiger, A., Richardson, K., & Potts, C. (2020). *Neural Natural Language Inference Models Partially Embed Theories of Lexical Entailment and Negation* (arXiv:2004.14623). arXiv. http://arxiv.org/abs/2004.14623

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of Mind Perception. *Science*, *315*(5812), 619–619. https://doi.org/10.1126/science.1134475

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520. https://doi.org/10.1037/a0013748

Gray, K., & Wegner, D. M. (2011). Dimensions of moral emotions. *Emotion Review*, *3*(3), 258–260. https://doi.org/10.1177/1754073911402388

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Griffin, D., Murray, S., & Gonzalez, R. (1999). Difference score correlations in relationship research: A conceptual primer. *Personal Relationships, 6*(4), 505–518. https://doi.org/10.1111/j.1475-6811.1999.tb00206.x

Guthrie, S. E. (1995). *Faces in the Clouds: A New Theory of Religion* (1st edition). Oxford University Press.

Haslam, N. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Jones, A. C. (2008). Development and validation of a Dog Personality Questionnaire [Unpublished Doctoral Dissertation]. *University of Texas, Austin*.

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107–128. https://doi.org/10.1037/h0034225

Lefkeli, D., Ozbay, Y., Gürhan-Canli, Z. et al. (2021) Competing with or Against Cozmo, the Robot: Influence of Interaction Context and Outcome on Mind Perception. *International Journal of Social Robotics, 13*, 715–724. https://doi.org/10.1007/s12369-020-00668-3

Loughnan, S., & Haslam, N. (2007). Animals and Androids: Implicit Associations Between Social Categories and Nonhumans. *Psychological Science*, *18*(2), 116–121. https://doi.org/10.1111/j.1467-9280.2007.01858.x

Morewedge, C. K., Preston, J., & Wegner, D. M. (2007). Timescale bias in the attribution of mind. *Journal of Personality and Social Psychology*, *93*(1), 1–11. https://doi.org/10.1037/0022-3514.93.1.1

Nie, Y., Williams, A., Dinan, E., Bansal, M., Weston, J., & Kiela, D. (2020). Adversarial NLI: A New Benchmark for Natural Language Understanding. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. http://dx.doi.org/10.18653/v1/2020.acl-main.441

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive science*, *30*(3), 531-553. https://doi.org/10.1207/s15516709cog0000_68

Öhman, A. (1986). Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*, *23*(2), 123–145. https://doi.org/10.1111/j.1469-8986.1986.tb00608.x

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Ross, L. (1977). The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process. In *Advances in Experimental Social Psychology* (Vol. 10, pp. 173–220). Elsevier. https://doi.org/10.1016/S0065-2601(08)60357-3

Shariff, A. F., & Norenzayan, A. (2011). Mean gods make good people: Different views of God predict cheating behavior. *International Journal for the Psychology of Religion*, *21*(2), 85–96. https://doi.org/10.1080/10508619.2011.556990

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*(4), 813–838. http://dx.doi.org/10.1037//0022-3514.48.4.813

Thorne, J., Vlachos, A., Christodoulopoulos, C., & Mittal, A. (2018). *FEVER: A large-scale dataset for Fact Extraction and VERification*. arXiv. http://arxiv.org/abs/1803.05355

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2020). *XLNet: Generalized Autoregressive Pretraining for Language Understanding* (arXiv:1906.08237). arXiv. http://arxiv.org/abs/1906.08237

Table 1
Target ratings (Study 1).

Target	Fear		Imbalanced Mind		Cognition		Emotion	
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Demon	6.07	(1.34)	2.50	(2.28)	5.60	(1.90)	3.14	(2.19)
Possessed Human	5.98	(1.41)	1.10	(1.54)	3.39	(2.33)	2.72	(2.08)
Werewolf	5.93	(1.35)	1.07	(1.13)	3.85	(1.78)	3.48	(1.81)
Infected Zombie	5.85	(1.45)	1.41	(1.70)	2.76	(1.82)	1.68	(1.24)
Godzilla	5.64	(1.73)	0.88	(1.09)	3.11	(1.80)	2.73	(1.76)
Vampire	5.39	(1.51)	1.00	(1.38)	6.20	(1.30)	5.37	(1.61)
King Kong	5.28	(1.63)	1.06	(1.12)	3.89	(1.65)	4.07	(1.82)
Predator	5.19	(2.15)	1.35	(1.61)	4.08	(2.32)	2.75	(1.75)
Xenomorph	5.12	(2.22)	1.47	(1.64)	3.69	(2.06)	2.30	(1.59)
Classic Zombie	4.98	(1.96)	1.02	(1.56)	2.47	(1.76)	1.58	(1.10)
Alien	4.77	(1.66)	1.26	(1.51)	6.09	(1.51)	4.87	(1.84)
Troll	4.70	(1.69)	0.88	(1.11)	3.24	(1.52)	2.78	(1.40)
Mummy	4.60	(1.81)	0.95	(1.12)	3.30	(1.73)	2.73	(1.57)
Ghost	4.50	(1.85)	0.74	(1.07)	5.31	(1.82)	4.94	(2.01)
Big Foot	4.28	(1.83)	0.86	(0.97)	4.02	(1.65)	3.69	(1.64)
Frankenstein's Monst.	3.94	(1.63)	0.95	(1.15)	3.95	(1.79)	4.12	(1.99)
Flying Monkey	3.27	(1.60)	0.94	(1.16)	3.60	(1.65)	3.03	(1.61)
The Blob	2.81	(1.83)	0.26	(0.70)	1.80	(1.40)	1.57	(1.21)

Table 2

Target ratings (Study 2a).

	Fear		Imbalanced Mind		Cognition		Emotion	
Target	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Supernatural								
The Devil	6.37	(1.50)	1.60	(1.63)	1.27	(1.54)	0.56	(2.13)
Werewolf	6.26	(1.17)	1.28	(1.02)	-0.97	(1.35)	0.21	(1.29)
Possessed Human	6.09	(1.50)	1.48	(1.52)	-1.03	(2.04)	0.06	(1.87)
Infected Zombie	6.00	(1.51)	0.85	(1.18)	-1.85	(1.58)	-1.80	(1.29)
Classic Zombie	5.49	(1.72)	0.31	(0.56)	-2.32	(1.31)	-2.09	(1.37)
Vampire	5.37	(1.93)	0.82	(1.11)	0.31	(1.47)	0.39	(1.67)
Alien	5.34	(1.73)	1.18	(1.43)	1.28	(1.32)	0.21	(1.52)
Mummy	5.14	(1.78)	0.89	(0.92)	-1.70	(1.40)	-1.47	(1.58)
Ghost	4.77	(1.70)	1.13	(1.31)	-0.73	(1.73)	-1.06	(1.66)
Frankenstein's Monst.	4.54	(1.85)	0.84	(0.98)	-1.16	(1.56)	-0.66	(1.90)
The Blob	4.14	(2.12)	0.64	(0.90)	-2.21	(1.16)	-2.26	(1.23)
Natural								
Great White Shark	6.31	(1.28)	0.98	(1.02)	-1.28	(1.40)	-1.16	(1.51)
Lion	6.24	(1.38)	1.02	(0.99)	-0.38	(1.65)	0.11	(1.54)
Cobra	6.11	(1.37)	0.84	(0.85)	-1.26	(1.47)	-1.28	(1.48)
Tiger	6.07	(1.48)	0.80	(0.82)	-0.26	(1.51)	0.13	(1.57)
Crocodile	6.02	(1.36)	0.86	(0.81)	-1.14	(1.52)	-1.22	(1.47)
Bear	5.84	(1.48)	0.98	(0.85)	-0.82	(1.47)	-0.22	(1.43)
Wolf	5.31	(1.73)	0.69	(0.65)	-0.03	(1.66)	0.12	(1.59)
Gorilla	5.02	(1.64)	0.95	(0.96)	-0.24	(1.54)	0.31	(1.44)
Tarantula	4.91	(2.02)	0.78	(1.22)	-1.11	(1.78)	-1.56	(1.63)
Swarm of Wasps	4.58	(1.91)	0.97	(1.21)	-1.51	(1.53)	-1.67	(1.64)
Bat	3.44	(1.90)	0.71	(0.77)	-1.28	(1.43)	-0.92	(1.33)

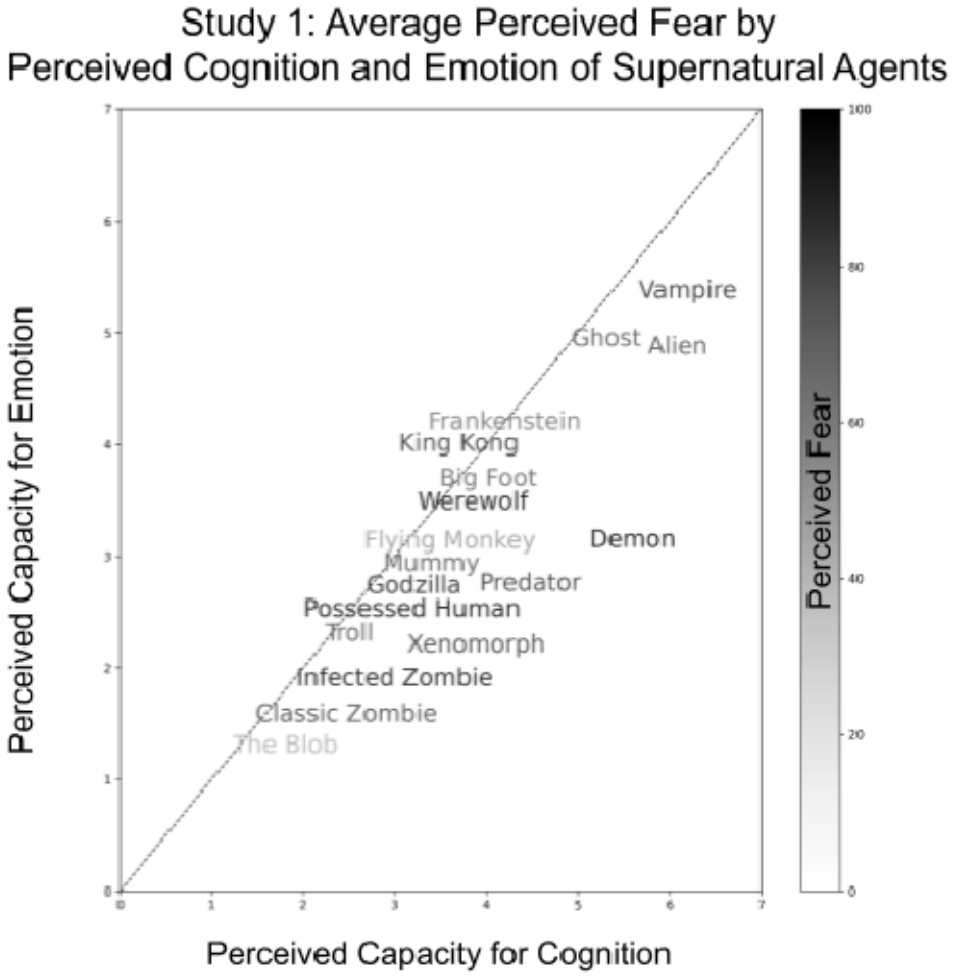
Table 3

Target ratings (Study 2b).

Target	Threat		Imbalanced Mind		Cognition		Emotion	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Supernatural								
Werewolf	6.74	(0.44)	1.50	(0.51)	-1.54	(0.80)	-0.05	(0.75)
Infected Zombie	6.45	(0.96)	1.50	(0.57)	-2.56	(0.77)	-1.06	(0.93)
Vampire	6.04	(0.97)	0.86	(0.69)	-0.59	(0.90)	0.10	(0.74)
Devil	5.88	(1.04)	0.67	(0.58)	0.16	(1.36)	0.02	(0.89)
Mummy	5.83	(1.01)	0.75	(0.65)	-2.06	(1.01)	-1.38	(1.03)
Classic Zombie	5.56	(1.39)	1.01	(0.60)	-2.19	(0.95)	-1.19	(0.78)
Frankenstein's Monst.	5.43	(0.94)	1.59	(0.86)	-1.52	(0.91)	0.06	(0.69)
Possessed Human	5.29	(1.67)	0.74	(0.46)	-1.35	(1.32)	-0.85	(1.05)
The Blob	4.92	(1.67)	0.82	(1.02)	-2.25	(1.40)	-1.90	(1.04)
Alien	4.82	(1.20)	0.57	(0.46)	0.54	(0.79)	-0.03	(0.53)
Ghost	4.29	(1.16)	1.10	(0.93)	-0.04	(1.31)	-0.94	(0.96)
Natural								
Tiger	6.17	(1.10)	1.23	(0.72)	-1.51	(1.00)	-0.28	(0.66)
Lion	6.03	(0.88)	1.37	(0.54)	-1.74	(0.71)	-0.37	(0.79)
Shark	6.03	(1.28)	1.17	(0.59)	-1.80	(1.00)	-0.70	(0.86)
Crocodile	5.86	(0.69)	0.92	(0.50)	-1.59	(1.08)	-0.68	(0.93)
Wasp	5.79	(1.11)	0.81	(0.72)	-1.98	(0.84)	-1.36	(1.01)
Cobra	5.75	(1.52)	0.73	(0.48)	-1.41	(1.00)	-0.68	(0.89)
Gorilla	5.59	(1.30)	1.01	(0.35)	-1.17	(0.55)	-0.15	(0.33)
Wolf	5.57	(1.40)	1.34	(0.76)	-1.44	(0.94)	-0.10	(0.64)
Bear	5.50	(1.06)	1.10	(0.50)	-1.44	(0.91)	-0.38	(0.52)
Spider	5.24	(1.40)	1.11	(0.88)	-1.63	(1.12)	-0.86	(0.82)
Bat	3.47	(1.65)	0.69	(0.56)	-1.67	(0.85)	-1.18	(0.55)

Figure 1.

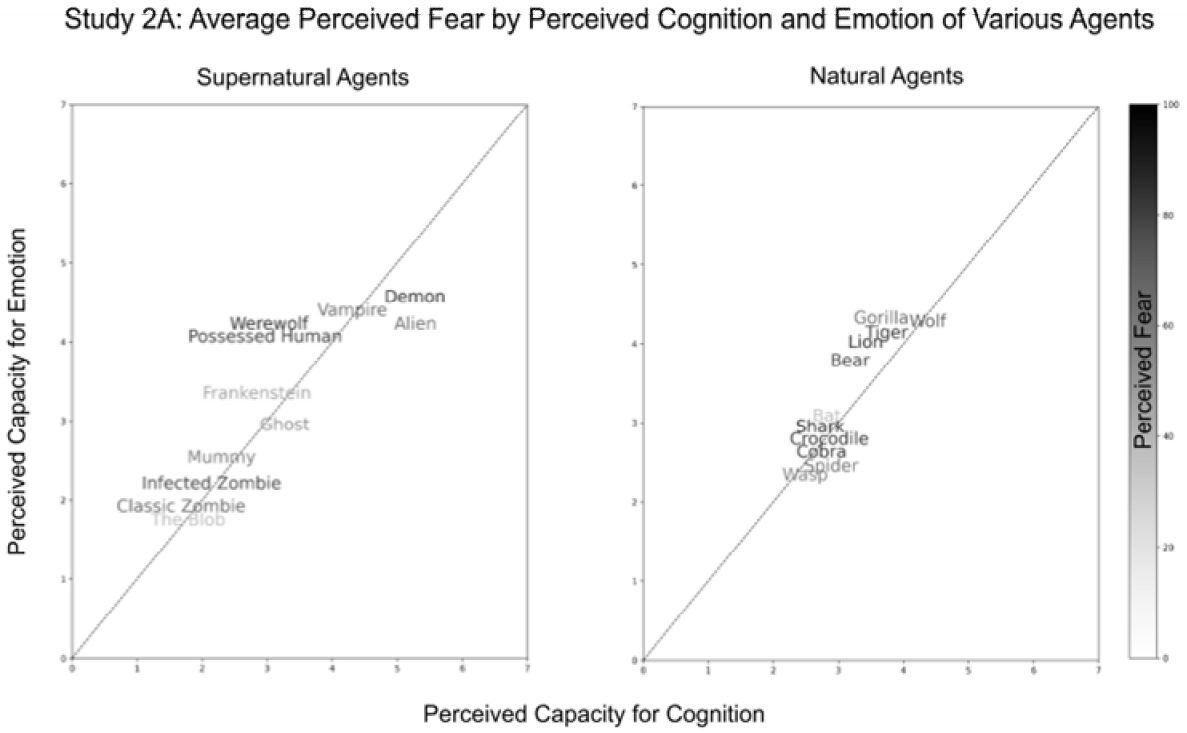
Scatterplot Depicting Perceived Fear, Cognition, and Emotion of Various Supernatural Agents



Note. A 45-degree line indicates where each dimension of mind (Cognition and Emotion) are balanced. Distance from the line indicates greater degrees of imbalance. Perceived fear of each agent is normalized and labels are jittered for visualization purposes.

Figure 2.

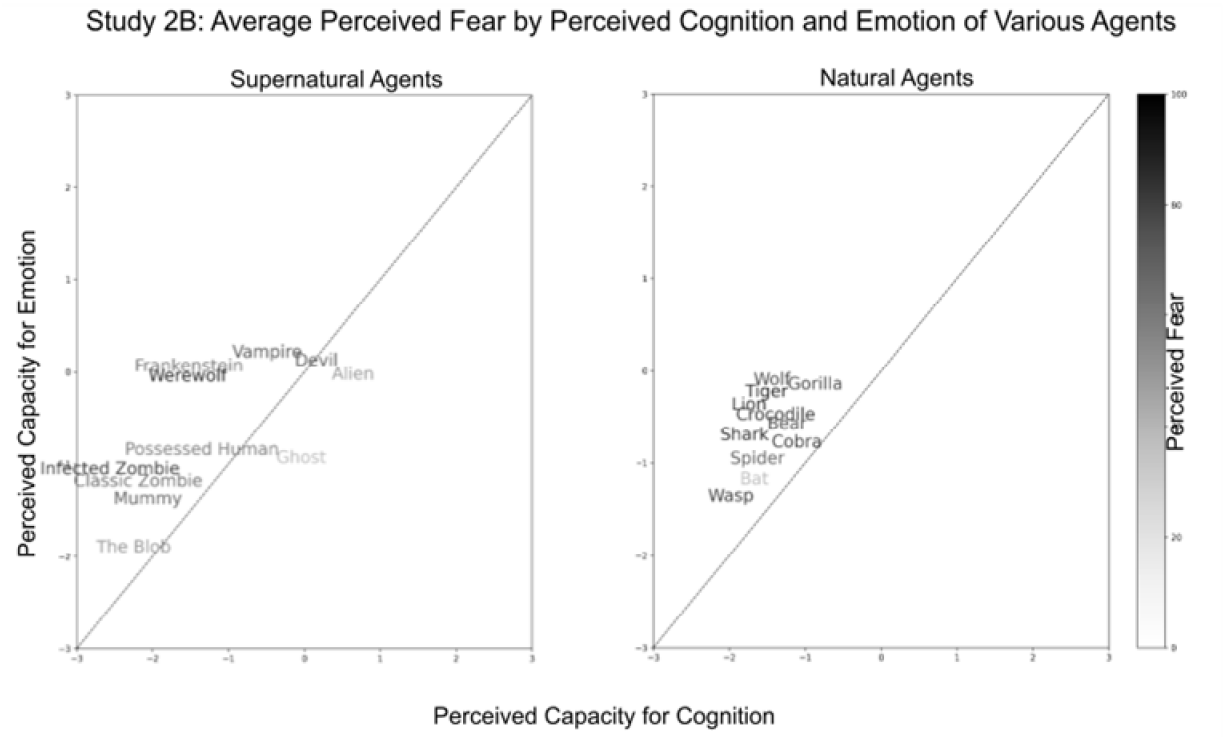
Scatterplot Depicting Perceived Fear, Cognition, and Emotion of Various Supernatural and Natural Agents in Study 2a



Note. A 45-degree line indicates where each dimension of mind (Cognition and Emotion) are balanced. Distance from the line indicates greater degrees of imbalance. Perceived fear of each agent is normalized and labels are jittered for visualization purposes.

Figure 2b.

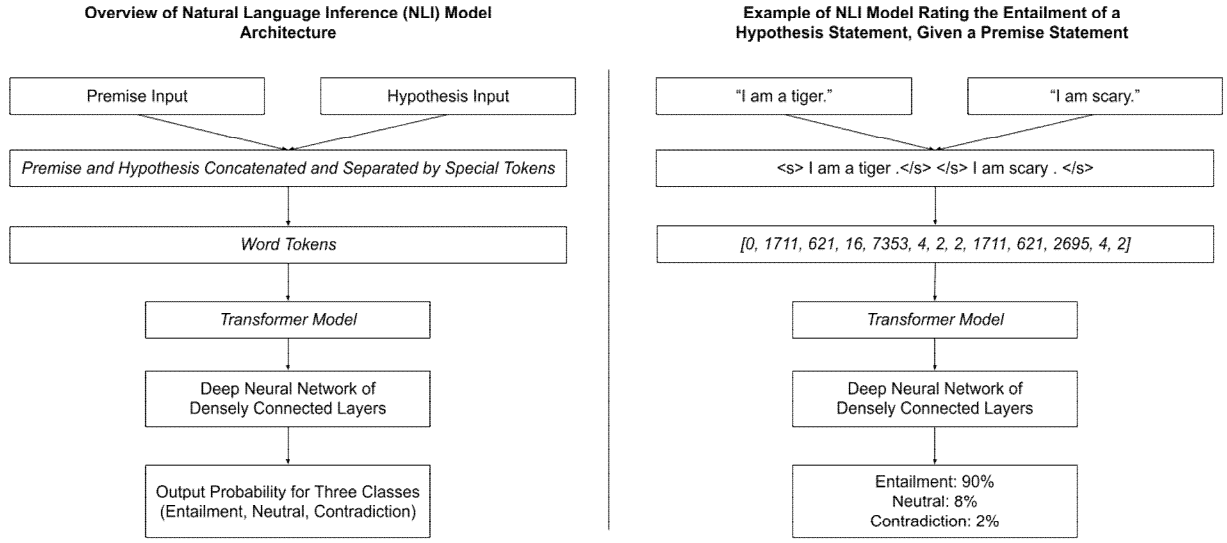
Scatterplot Depicting Perceived Fear, Cognition, and Emotion of Various Supernatural and Natural Agents in Study 2b



Note. A 45-degree line indicates where each dimension of mind (Cognition and Emotion) are balanced. Distance from the line indicates greater degrees of imbalance. Perceived fear of each agent is normalized and labels are jittered for visualization purposes.

Figure 4.

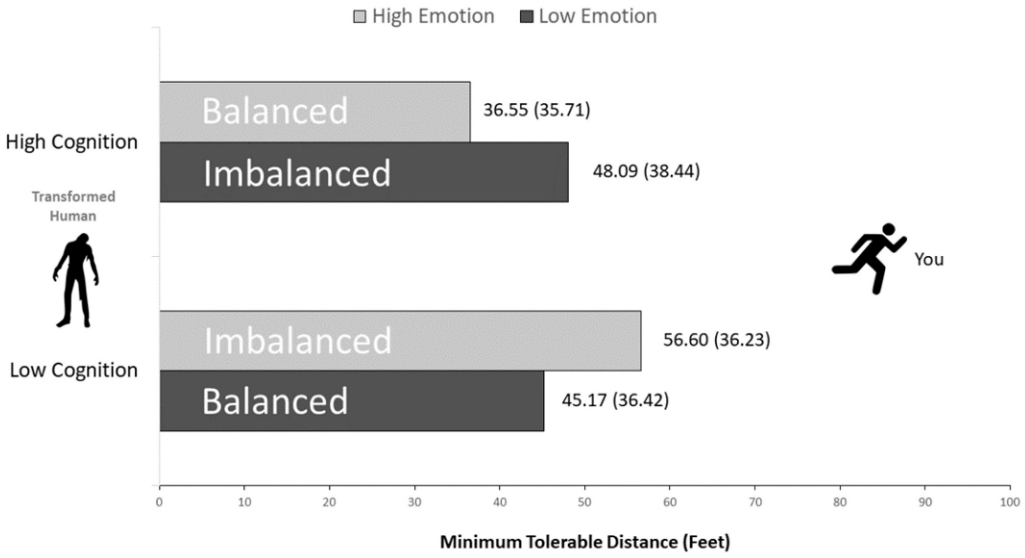
Illustration of Natural Language Inference Model



Note. The above figure shows how a natural language inference model accepts a premise statement and a hypothesis statement. It then scores how well the hypothesis can be inferred, given the premise. The panel on the right provides a concrete example of how the model would score related to animal perception. We subtracted the contradiction score from the entailment score to obtain a single value of that attribute for that target.

Figure 5.

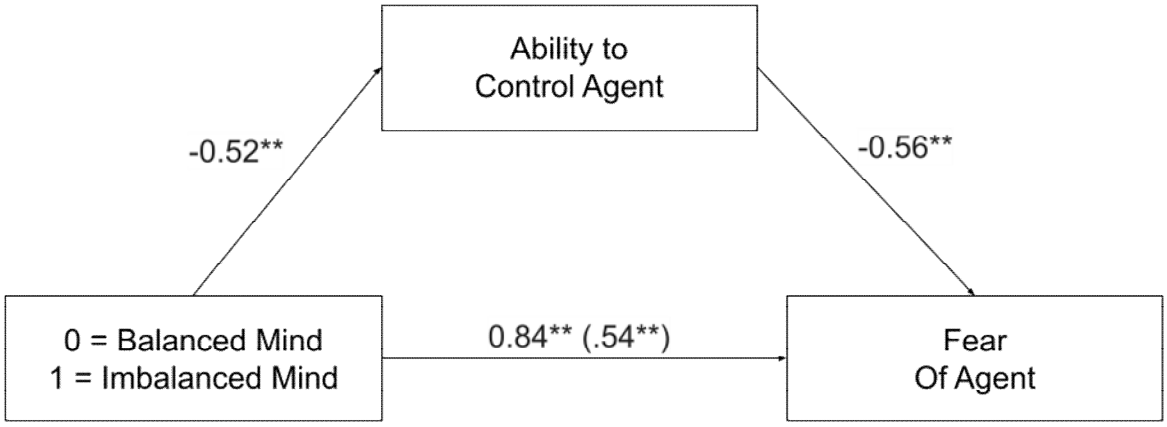
Minimum distance participants were willing to tolerate between themselves and each kind of infected person (Study 4)



Note. Values represent means with standard deviations in parentheses.

Figure 6.

Mediation diagram of imbalanced mind perception on threat through control (Study 4)



Note: **p* < .05, ***p* < .01. Indirect effect = -.29. Proportion of Total Effect mediated by Control =36%

Appendix

Descriptions of hypothetical infected persons (Study 4)

We would now like you to imagine that there is a terrible viral outbreak that poses a very serious threat to the human race as we know it. The 3 phases of this "brain virus" are as follows:

Infection: Symptoms of the brain virus appear quickly. Within one or two hours, the victim will develop a headache, fever, chills and other flu-like symptoms.

Coma: Slow pulse and shallow breathing characterize a brief coma that lasts between four and six hours.

Transformation: Victims awaken from their coma with a transformed mind.

Little else is known about the virus, but it is clear that infected persons become driven (whether consciously or unconsciously) to harm those who are not already infected with the virus.

However, given that their minds are transformed, persons infected with this brain virus may be more or less of a threat depending on the results of this transformation.

Imagine an infected person that has the following traits…

Low Cognition/Low Emotion:

> Imagine an infected person that experiences emotions much less intensely than before the transformation, and is relatively insensitive to feelings of pleasure and pain. They also think, reason and plan much worse than before the transformation, and are extremely unintelligent. In sum, they don't show much emotion, are unreactive to pain (not easily agitated), and unintelligent.

Low Cognition/High Emotion:

> Imagine an infected person that experiences emotions much more intensely than before the transformation, and is relatively more sensitive to feelings of pleasure and pain. They also think, reason and plan much worse than before the transformation, and are extremely unintelligent. Further, they lack self-control which causes poor regulation of their extreme emotions. In sum, they are highly emotional, highly reactive to pain (easily agitated), unintelligent, and lacking in self-control.

High Cognition/Low Emotion:

> Imagine an infected person that experiences emotions much less intensely than before the transformation, and is relatively insensitive to feelings of pleasure and pain. They also think, reason and plan much better than before the transformation, and are extremely intelligent. In sum, they feel no emotions, are unreactive to pain (not easily agitated), and are highly intelligent.

High Cognition/High Emotion:

> Imagine an infected person that experiences emotions much more intensely than before the transformation, and is relatively more sensitive to feelings of pleasure and pain. They think, reason and plan much better than before the transformation, and are extremely intelligent. Further, their high self-control causes good regulation of their extreme emotions. In sum, they are highly emotional, highly reactive to pain (easily agitated), intelligent, and have high self-control.