Rethinking people’s conceptions of mental life

Kara Weisman, Carol S. Dweck, and Ellen M. Markman

Contributed by Psychology, Stanford University, Stanford, CA 94305

How do people make sense of the emotions, sensations, and cognitive abilities that make up mental life? Pioneering work on the dimensions of mind perception has been interpreted as evidence that people consider mental life to have two core components—experience (e.g., hunger, joy) and agency (e.g., planning, self-control) [Gray HM, et al. (2007) Science 315:619]. We argue that this conclusion is premature: The experience–agency framework may capture people’s understanding of the differences among different beings (e.g., dogs, humans, robots, God) but not how people parse mental life itself. Inspired by Gray et al.’s bottom-up approach, we conducted four large-scale studies designed to assess people’s conceptions of mental life more directly. This led to the discovery of an organization that differs strikingly from the experience–agency framework: Instead of a broad distinction between experience and agency, our studies consistently revealed three fundamental components of mental life—suites of capacities related to the body, the heart, and the mind—with each component encompassing related aspects of both experience and agency. This body–heart–mind framework distinguishes itself from Gray et al.’s experience–agency framework by its clear and importantly different implications for dehumanization, moral reasoning, and other important social phenomena.

Significance

How do ordinary people make sense of mental life? Pioneering work on the dimensions of mind perception has been widely interpreted as evidence that lay people perceive two fundamental components of mental life: experience and agency. However, using a method better suited to addressing this question, we discovered a very different conceptual structure. Our four studies consistently revealed three components of mental life—suites of capacities related to the body, the heart, and the mind—with each component encompassing related aspects of both experience and agency. This body–heart–mind framework distinguishes itself from the experience–agency framework by its clear and importantly different implications for dehumanization, moral reasoning, and other important social phenomena.

Author contributions: K.W., C.S.D., and E.M.M. designed research; K.W. performed research; K.W. analyzed data; and K.W., C.S.D., and E.M.M. wrote the paper.

Significance

How do ordinary people make sense of mental life? Pioneering work on the dimensions of mind perception has been widely interpreted as evidence that lay people perceive two fundamental components of mental life: experience and agency. However, using a method better suited to addressing this question, we discovered a very different conceptual structure. Our four studies consistently revealed three components of mental life—suites of capacities related to the body, the heart, and the mind—with each component encompassing related aspects of both experience and agency. This body–heart–mind framework distinguishes itself from the experience–agency framework by its clear and importantly different implications for dehumanization, moral reasoning, and other important social phenomena.

Author contributions: K.W., C.S.D., and E.M.M. designed research; K.W. performed research; K.W. analyzed data; and K.W., C.S.D., and E.M.M. wrote the paper.

Reviewers: P.B., Yale University; and S.A.G., University of Michigan.
The authors declare no conflict of interest.

Data deposition: The data for studies 1–4 have been deposited in an Open Science Framework (OSF) project, https://osf.io/m3gwu/.

1To whom correspondence may be addressed. Email: kweisman@stanford.edu or markman@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1704347114/-/DCSupplemental.
Which capacities “go together,” and which are more distinct? Instead, participants in Gray et al.’s original study were led to focus on the similarities and differences among characters: Each participant engaged in a series of pairwise comparisons of characters (e.g., comparing a robot vs. an infant, a woman vs. God) while considering only one mental capacity throughout the session (e.g., one participant would compare characters’ relative capacities for joy, while another participant would compare their capacities for pain). This approach is well suited to reveal how participants think about the similarities and differences among social beings, i.e., how different characters rank in their capacities for joy, pain, and so forth. However the observation that, in the aggregate, participants who considered characters’ capacities for joy ranked them similarly to those who considered characters’ capacities for pain does not necessarily mean that people consider these two mental capacities to be particularly strongly related. Thus, while concepts of experience and agency seem to capture an important aspect of how people compare different beings, it is unclear whether they correspond to folk conceptions of the fundamental components of mental life.

Our primary goal in the present studies was to directly investigate the conceptual space of mental capacities themselves. Inspired by Gray et al.’s bottom-up approach, we designed an experimental paradigm that would allow this intuitive ontology of mental life to emerge from participants’ responses organically, without top-down hypotheses about which kinds of mental capacities might go together in people’s reasoning. We led participants to focus on the connections and divisions between different aspects of mental life by asking them to evaluate a wide variety of mental capacities for a single character. For example, one participant would be asked to consider the extent to which a robot is capable of experiencing joy, experiencing pain, seeing things, having intentions, and so forth, while another participant would be asked the same questions about an insect, a human, or some other entity. We used patterns of attributions—i.e., when a participant judged some character to be highly capable of some capacity, which other capacities came along with it—to infer which mental capacities were seen as related and which were considered independent. This approach, in which each individual participant compares and contrasts a range of mental capacities, is a more direct and valid way of assessing participants’ intuitions about the organization of these mental capacities than inferring them from other figures via Gray et al.’s original study design.

We began by exploring attributions of mental capacities to two carefully selected edge cases in social reasoning: a beetle and a robot. This approach offers several advantages. First, beetles are likely considered much more capable of experience and less capable of agency than robots (1). Moreover, because beetles are animals and robots are artifacts, this pair provides a glimpse into the role of biological life in attributions of mental life. Most importantly for our bottom-up approach, the capacities of these edge cases were expected to be controversial, ensuring that not all participants would endorse all capacities (as they might for, say, a human). This allowed us to address the following question: When people disagree about what some entity is capable of, which mental capacities tend to go together?

In study 1, each participant judged either a beetle (n = 200) or a robot (n = 205) on 40 mental capacities, including various affective, perceptual, physiological, cognitive, agentic, social, and other abilities. Study 2 was a direct replication (n = 406). In study 3 (n = 200), each participant judged both a beetle and a robot side by side, allowing us to examine whether the intuitions revealed by studies 1 and 2 are preserved when individual participants are presented with the salient contrast between the two edge cases. In study 4, we broadened the set of characters to include a wide range of entities: Each participant (n = 431) judged one of 21 entities (adult, child, infant, person in a persistent vegetative state, fetus, chimpanzee, elephant, dolphin, bear, dog, goat, mouse, frog, blue jay, fish, beetle, microbe, robot, computer, car, stapler).

This allowed us to assess which aspects of mental life hang together when different entities are considered to have different profiles of mental capacities.

Results
Using a bottom-up approach with the potential to affirm or challenge the experience–agency framework, we discovered a different conceptual organization: a three-way distinction between physiological abilities related to the body, social–emotional abilities related to what we might call the heart, and perceptual–cognitive abilities related to the mind, with each factor encompassing aspects of both experience and agency. In integrating related forms of experience and agency, each of these three factors hints at a coherent and distinct subsystem for making sense of the behavior of other beings. In the case of the body, physiological sensations indicate biological needs, which an animal might address via self-initiated behavior (see ref. 11). In the case of the heart, a person’s emotional life and his or her ability to anticipate others’ emotions bear on the social and moral ramifications of his or her behavior (see refs. 1, 5, and 6). In the case of the mind, perceptual access and cognitive representational abilities combine to influence an agent’s goal-directed actions, as theory of mind researchers have long discussed (9).

Independent exploratory factor analyses for each of our four studies all revealed this same three-factor structure (Table 1; see Figs. S2 and S3 for the mean responses for each mental capacity, by target character.)

The first factor corresponded primarily to physiological sensations related to biological needs, as well as the kinds of self-initiated behavior needed to pursue these needs—in other words, abilities related to the physical, biological body. It was the dominant factor for the following items (from greatest to smallest factor loading, according to study 1): getting hungry, experiencing pain, feeling tired, experiencing fear, experiencing pleasure, having free will (see below), being conscious, feeling safe, having desires, feeling nauseated, feeling calm, and having intentions. Across studies, factor 1 accounted for 36–40% of the total variance in the rotated maximal solution.

The second factor corresponded primarily to basic and social emotions, as well as the kinds of social-cognitive and self-regulatory abilities required of a social partner and moral agent. Together, these abilities resonate with the metaphorical sense of heart, most also perhaps with some notions of “spirit” or “soul” (12). Across studies, factor 2 was the dominant factor for the items feeling happy, experiencing pride, experiencing guilt, holding beliefs, feeling disrespected, feeling depressed, understanding how others are feeling, telling right from wrong, and exercising self-restraint and accounted for 26–31% of the total variance.

Finally, the third factor corresponded primarily to perceptual–cognitive abilities to detect and use information about the environment, capacities traditionally associated with the mind. Across studies, it was the dominant factor for the items remembering things, recognizing someone, sensing temperatures, communicating with others, seeing things, perceiving depth, working toward a goal, detecting sounds, and making choices and accounted for 16–20% of the total variance.

See Fig. S1 for a 3D representation of this conceptual space (available online at rpubs.com/kgweisman/bodyheartmind_figureS1) and Table 1 for the complete set of factor loadings.

It is worth noting that additional items not listed above loaded equally strongly on more than one factor, in largely sensible ways. Some forms of perception (e.g., detecting odors) were associated with both body and mind. A few basic emotions (e.g., getting angry, feeling happy) and self-awareness (being self-aware, having thoughts) were associated with both body and heart. The capacity for reasoning about things was related to both mind and heart. The capacity for having intentions loaded relatively equally on all three factors. Such cross-loadings might indicate capacities that are considered
to be combinations of more basic abilities or capacities that feed into more than one conceptual system.

Two items loaded reliably on a single factor but in ways that surprised us: holding beliefs patterned with the social–emotional items related to the heart much more strongly than the perceptual–cognitive items related to the mind, and having free will tracked the physiological phenomena of the body more closely than the social–emotional capacities of the mind. We suspect that these items reflect dissociations between academic terminology (in which “holding beliefs” is equivalent to thinking that some proposition is true, and “free will” might connote the ability to initiate behavior without external causes or constraints; see ref. 14 for an extended discussion of folk concepts of free will).

On the whole, however, the general pattern that emerged from these four studies is clear, highly reliable, and quite different from the experience–agency framework that has been widely assumed to characterize folk beliefs about mental life. Given the range of mental capacities included in each study, a number of additional or alternative factors could have emerged, including experience or agency. However, while we were able to replicate the experience–agency framework using Gray et al.’s (1) character-comparison paradigm [both with Gray et al.’s original 18 mental capacities and with the 40 mental capacities

Table 1. Factor loadings from exploratory factor analyses for all studies (after varimax rotation)

<table>
<thead>
<tr>
<th>A priori category</th>
<th>Item</th>
<th>Factor 1: Body</th>
<th>Factor 2: Heart</th>
<th>Factor 3: Mind</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHY</td>
<td>Getting hungry*</td>
<td>0.92</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>PHY</td>
<td>Experiencing pain*</td>
<td>0.92</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>PHY</td>
<td>Feeling tired</td>
<td>0.83</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>EMO</td>
<td>Experiencing fear*</td>
<td>0.83</td>
<td>0.83</td>
<td>0.79</td>
</tr>
<tr>
<td>EMO</td>
<td>Experiencing pleasure*</td>
<td>0.75</td>
<td>0.73</td>
<td>0.71</td>
</tr>
<tr>
<td>COG</td>
<td>Doing computations</td>
<td>-0.73</td>
<td>-0.73</td>
<td>-0.79</td>
</tr>
<tr>
<td>AGE</td>
<td>Having free will</td>
<td>0.71</td>
<td>0.72</td>
<td>0.70</td>
</tr>
<tr>
<td>AGE</td>
<td>Having intentions</td>
<td>0.54</td>
<td>0.56</td>
<td>0.54</td>
</tr>
<tr>
<td>SOC</td>
<td>Feeling embarrassed*</td>
<td>0.20</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>EMO</td>
<td>Experiencing pride*</td>
<td>0.29</td>
<td>0.27</td>
<td>0.25</td>
</tr>
<tr>
<td>SOC</td>
<td>Feeling love</td>
<td>0.39</td>
<td>0.41</td>
<td>0.44</td>
</tr>
<tr>
<td>SOC</td>
<td>Experiencing guilt</td>
<td>0.27</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>COG</td>
<td>Holding beliefs</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>SOC</td>
<td>Feeling disrespected</td>
<td>0.27</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>EMO</td>
<td>Feeling depressed</td>
<td>0.39</td>
<td>0.41</td>
<td>0.37</td>
</tr>
<tr>
<td>SOC</td>
<td>Understanding how others are feeling†</td>
<td>0.06</td>
<td>0.06</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Loadings greater than or equal to 0.60 or less than or equal to −0.60 are in boldface type. Items are listed according to their dominant factor loading in study 1. Each item is listed with its a priori category membership: AGE, agentic; EMO, emotional; COG, cognitive; PER, perceptual; PHY, physiological; SOC, social; and other/multiple (unmarked).
*Experience dimension from Gray et al. (1).
†Agency dimension from Gray et al. (1).
from our studies 1–4; see Study S1: Replication of Gray et al. (2007) and Study S2: Comparing Characters on 40 Mental Capacities), we saw a very different framework emerge when we asked people to consider the similarities and differences among mental capacities themselves. We observed very similar latent structures across independent analyses, whether participants judged a single edge case in isolation (studies 1 and 2), compared two edge cases that contrasted in biological animacy (study 3), or evaluated a wide range of entities, from inert objects to canonical social partners (study 4). It would be fascinating to explore the cultural and developmental origins of this conceptual system: Are these three ways of reasoning about the behavior of other beings universal, or do they reflect a culturally bounded understanding of the world specific to our US sample? In what ways are intuitions about mental life shaped by culture, religion, education, and personal observation?

Neither experience nor agency emerged as a unitary factor in any of these studies. Instead, distinctions among three varieties of experience—physiological sensations (body), emotions (heart), and perceptual abilities (mind)—were particularly prominent in the latent structure underlying participants’ responses. Beyond this, we were particularly interested, and initially rather surprised, to see that agentic capacities were also distributed across these three factors in reliable ways. We now argue that each of these three factors encompasses conceptually related pairings of experiential and agentic capacities.

These conceptual relationships are especially clear for heart and mind. Within the social–emotional capacities that characterized the heart factor we find both experiences of emotion (e.g., happiness, guilt) and four of the seven items that constituted Gray et al.’s (1) original agency dimension: understanding how others are feeling, telling right from wrong, exercising self-restraint, and (in two of four studies) having thoughts. These items reflect a particularly social and moral form of agency. Likewise, within the perceptual–cognitive capacities that characterized the mind factor are both perceptual experiences (e.g., vision, hearing) and the remainder of Gray et al.’s original agency items: remembering things, communicating with others, and working toward a goal (along with one of our additional agentic items, making choices). This might reflect the role of perception and cognition in action-planning and goal pursuit.

Moreover, the collection of items related to the body hints at a conceptual relationship worthy of further investigation. Like the other factors, the body factor included a distinct variety of experience: physiological sensations (e.g., hunger, pain). However, it also included two items that represent important aspects of agency: having free will and having intentions. Although this pattern was unexpected—particularly because it contrasts so vividly with academic accounts of free will as rational self-determination—it emerged across all four studies. We speculate that this reflects the importance of self-initiated behaviors and self-propelled motion in lay people’s understanding of biological life and animacy (11).

To recap, we set out to investigate people’s intuitive ontology of mental life: What do people consider to be the fundamental components of the mind? In four studies, we documented a robust and reliable answer to this question (at least for US adults): Participants’ mental capacity judgments were undergirded by a three-part conceptual structure, resonating with notions of body, heart, and mind.

The body–heart–mind framework is a substantially different perspective on people’s beliefs about mental life than Gray et al.’s experience–agency framework. What accounts for these differences, and how do they play out in everyday social reasoning?

Although further research would be required to answer this question in full, we speculate that these two frameworks—body–heart–mind and experience–agency—capture distinct aspects of social reasoning. On the one hand, people are faced with the question of how to organize and make sense of the wide range of mental capacities that one must take into account to represent another mind. Our studies, in which participants were encouraged to consider the similarities and differences among mental capacities, suggest that the body–heart–mind framework serves this function in the population we tested. On the other hand, people are also faced with the question of how to organize and make sense of the wide range of beings in the world, from humans and animals to supernatural beings and social technologies, and the different roles they play in social interactions (5). Gray et al.’s (1) study, in which participants were encouraged to consider the similarities and differences among characters, suggests that the experience–agency framework might serve this function. If these are indeed two different conceptual frameworks serving distinct roles in social reasoning, this raises interesting questions about when one framework takes precedence over the other or how they might operate in tandem.

Consider the domain of morality. One of the most influential applications of the experience–agency framework has been to argue that moral responsibility is critically tied to attributions of agency, while victimhood (moral patiency) is tied to attributions of experience (5, 6; cf. ref. 7). However, if people were to limit themselves to asking, “What did the perpetrator do, and what did the victim feel?” they would have only a skeletal understanding of the moral significance of any given event. Beyond thinking about the contrast between the two characters involved in this interaction, an observer might also think about the different aspects of mental life that guide their actions. From a bodily perspective, the observer might ask, “What physical actions did they take? Did anyone experience pain?” but beyond this, the observer might also adopt a more perceptual–cognitive perspective, wondering, “What did each person see, hear, and know? How was each person understanding the situation?” or a more social–emotional perspective, asking, “Did these people demonstrate self-restraint or consideration of each other’s feelings? What emotions did each person experience?” The observer might even violate the division between agency/responsibility and experience/patiency by taking into account the patient’s agency in assessing his or her moral status (e.g., “Did the patient provoke the agent’s actions?”) or by considering the agent’s experience in judging his or her culpability (e.g., “Does the agent feel remorse?”). However, the observer would also need information about different aspects of agency (and perhaps experience) to assess moral responsibility; likewise, he or she might weigh the importance of different varieties of experience (and perhaps agency) to assess victimhood. We speculate that this process of moving back and forth between thinking about the roles that different characters are playing (agent vs. patient) as well as the more nuanced analysis of the different aspects of mental life that guide their behavior (body, heart, mind) may more closely approximate moral reasoning in everyday life.

Another compelling case study of mind perception is a natural experiment currently playing out in much of the developed world: the introduction of increasingly sophisticated “intelligent” and “social” technologies. How are we to make sense of a self-driving car, a robotic caregiver, or a virtual personal assistant? When do these kinds of technologies feel useful, interesting, and fun, and when do they feel uncanny (15)? To take one example, Gray et al.’s work suggests that people perceive robots to be very low on experience and middling on agency (1) and that increasing people’s perceptions of robots’ experience (but not agency) elicits feelings of uncanniness (3). However in the current studies, participants actually endorsed some kinds of experience quite strongly for robots (e.g., sensing temperatures, seeing things) and rejected quite a few agentic abilities (e.g., having free will, telling right from wrong) (Fig. 1 and Figs. S2 and S3), suggesting that there might be more to beliefs and feelings about robots than their capacities for experience and agency, broadly writ. In particular, people’s beliefs
and attitudes toward robots might be guided by their conceptual understanding of the components of mental life. For example, thinking about robots and other artificial intelligences as having perceptual–cognitive abilities might feel perfectly comfortable to many people, while entertaining the possibility of technological beings having social–emotional or bodily capacities may be more unsettling. Examining human–robot interactions through the lens of body–heart–mind framework could bring some clarity to these complicated and increasingly relevant issues.

Just as notions of bodies, hearts, and minds are likely to shape the ways we humanize technological beings, this three-part conceptual structure might also shape the converse processes of dehumanizing and devaluing fellow human beings. Claims that some people are somehow less than human—in particular, that they lack certain mental capacities—have been used to justify oppression throughout history and through the present day (e.g., ref. 16). If experience and agency are understood to be the two components of mental life, as has been widely assumed, it is logical to suppose that there may be two corresponding forms of dehumanization: the denial of experience, and the denial of agency (17). However, current theories of dehumanization are considerably richer and more nuanced than this (18). The body–heart–mind framework might be a more useful framework for making sense of dehumanization, which, from our perspective, might take the form of augmenting or diminishing mental capacities in any of our three domains (body, heart, or mind). This resonates well with most recent work on dehumanization, which, from our perspective, might take the form of augmenting or diminishing mental capacities in any of our three domains (body, heart, or mind). This resonates well with most recent work on dehumanization, which, from our perspective, might take the form of augmenting or diminishing mental capacities in any of our three domains (body, heart, or mind).

In these studies, participants rated target characters on 40 mental capacities. The goal of these studies was to assess lay people’s understanding of the structure of mental life by evaluating which mental capacities tend to hang together in people’s judgments.

### Materials and Methods

The studies reported in this paper were approved by the Stanford Administrative Panel on Human Subjects in Nonmedical Research. Informed consent was obtained from all participants. Stimuli, surveys, data, and analysis code for studies 1–4 are available at https://osf.io/m3gwu.

In these studies, participants rated target characters on 40 mental capacities. The goal of these studies was to assess lay people’s understanding of the structure of mental life by evaluating which mental capacities tend to hang together in people’s judgments.

### Materials

**Surveys.** Studies were administered online. Participants first saw a photograph of the character(s) they were assigned to evaluate and read, “You have been assigned to evaluate the following entity: [target]. On the following page, you will see a list of mental capacities. For each mental capacity, please indicate the extent to which you believe a [target] has this capacity. Please note: We care only about your opinion or best guess—please do not do any external research about these questions.”

Participants then proceeded to a page featuring the same photograph(s), with the following text prominently displayed: “On a scale of 0 (Not at all capable) to 6 (Highly capable), how capable is a [target] of . . . ?” This was followed by the 40 mental capacities and one attention check (“Please select 4”), presented in a random order for each participant. Participants were required to answer every question.

In studies 1, 2, and 4, each participant evaluated one randomly assigned target character. In study 3, every participant evaluated two target characters presented side by side (left–right position was determined randomly). See Characters, below.

Finally, participants were presented with demographics questions (Table S1). **Characters.** For all studies, each target character was accompanied by a brief label (e.g., “a beetle”) and a color photograph of that character. Participants were not provided with any further information about these characters.

For studies 1–3, the list of characters included a beetle and a robot. For study 4, the list of characters again included a beetle and a robot, as well as a stapler, a car, a computer, a microbe, a fish, a blue jay, a frog, a mouse, a goat, a bear, a dolphin, an elephant, a chimpanzee, a fetus, a person in a persistent vegetative state, an infant, a child, and an adult. Note that this set of characters included close variants of all 13 characters used in Gray et al.’s (1) dimensions of mind perception study, with the following exceptions: We included only one typical human adult, instead of three, in favor of including a wider range of entities toward the middle of the spectrum from inert objects to humans, and we excluded the recently deceased adult and God, whose mental capacities we anticipated would be difficult or confusing to evaluate exclusively, particularly for nonreligious participants.

See Table S2 for descriptions of these images, links to image sources, and condition randomization for each study.

### Mental capacities

The 40 mental capacities included in all studies were generated from an a priori conceptual analysis of candidate ontological categories of mental life, with the constraint that each category should include at least five items of varying valence, complexity, and phrasing:

- **Affective experiences:** feeling happy, feeling depressed, experiencing fear, getting angry, feeling calm, experiencing joy
- **Perceptual abilities:** detecting sounds, seeing things, sensing temperatures, detecting odors, perceiving depth
- **Physiological sensations related to biological needs:** getting hungry, feeling tired, experiencing pain, feeling nauseated, feeling safe

![Image](https://example.com/image.png)

*Fig. 1.* Mean ratings of the 40 mental capacities for a subset of the 21 entities in study 4. (See Figs. S2 and S3 for mean ratings for the full set of entities in all studies.) Participants responded on a scale of 0 (not at all capable) to 6 (highly capable). Error bars are nonparametric bootstrapped 95% CIs. Mental capacities are grouped according to their dominant factor in any study (stapler (n = 17); robot (n = 22); beetle (n = 19); goat (n = 20); elephant (n = 21); adult (n = 21)).
iv) Cognitive abilities: doing computations, having thoughts, reasoning about things, remembering things, holding beliefs
v) Agentic capacities: having free will, making choices, exercising self-restraint, having intentions, working toward a goal
vi) Social abilities: feeling love, recognizing someone, communicating with others, experiencing guilt, feeling disrespected, understanding how others are feeling, feeling embarrassed
vii) Additional items that fell into none or more than one of these categories: being conscious, being self-aware, experiencing pleasure, having desires, telling right from wrong, having a personality, experiencing pride

Note that this set of mental capacities included close variants of all 18 mental capacities used in Gray et al.’s (1) original experience-agency study (Table 1).

Methods. Participants. Participants in all studies participated via Amazon Mechanical Turk (MTurk). All participants had gained approval for ≥95% of previous work on MTurk (≥50 assignments), had verified accounts based in the United States, and indicated that they were ≥18 y old. Repeat participation (within each study and across studies) was prevented.

Study 1 (n = 405): Participants were paid $0.30 (median duration: 2.67 min). An additional 48 respondents were excluded for not completing the survey (n = 14), failing the attention check (i.e., failing to select “4” for an item that read “Please select 4”); n = 19), or not providing a year of birth (n = 15).

Study 2 (n = 406): Participants were paid $0.30 (median duration: 2.52 min). An additional 13 respondents were excluded for not providing a year of birth.

Study 3 (n = 400): Participants were paid $0.50 (median duration: 5.46 min). An additional 24 respondents were excluded for not completing the survey (n = 13), failing the attention check (n = 7), or not providing a year of birth (n = 4).

Study 4 (n = 431): Participants were paid $0.30 (median duration: 2.88 min). An additional 40 respondents were excluded for not completing the survey (n = 15), failing the attention check (n = 24), or not providing a year of birth (n = 1).

See Table S1 for demographic information.

Recruitment. Participants were recruited using the Human Intelligence Task title “Short psychology study judgments of familiar things” along with a time estimate, a description reading, “Complete a very short psychology study,” and the keywords “psychology, survey, study.” Participants were recruited between December 14, 2015, and January 14, 2016.

Statistical analysis. For each study we conducted exploratory factor analyses using Pearson correlations to find the minimum residual solution. [Note that factor analyses using polychoric correlations and/or oblimin rotation, principal components analyses (PCAs), correspondence analyses, and item response analyses all yielded similar latent structures.]

We first examined unrotated maximal solutions. To determine the maximum number of factors to extract, we used the following rule of thumb: With p observations per participant, we can extract a maximum of k factors, where (p − k) × 2 > (p + k), i.e., k < p/3. Thus, with 40 mental capacity items, we could extract a maximum of 13 factors.

To determine how many factors to retain, we used the following preset retention criteria, considering the unrotated maximal (13-factor) solution: Each factor must (i) have an eigenvalue >1.0, (ii) individually account for >5% of the total variance, and (iii) be the dominant factor (i.e., the factor with the highest factor loading) for at least one mental capacity item after rotation. In all studies, this yielded exactly three factors.

In the primary analysis for each study, we examined and interpreted varimax-rotated solutions, extracting only the number of factors that meet the retention criteria just described. See Table 1 for factor loadings for all mental capacities on these three rotated factors, for all studies.

See Supporting Information for methods, results, and discussion of two additional studies: Study S1 (results: Table S3) and Study S2 (results: Table S4).

ACKNOWLEDGMENTS. Special thanks to Susan Gelman and Paul Bloom, whose challenging and insightful reviews truly pushed our thinking forward. We also thank Sophie Bridges, Eleanor Chestnut, and David McClure for feedback on the manuscript; Heather Gray for providing information and materials for the replication study; and Michael Frank for guidance in conducting the replication. This material is based on work supported by the National Science Foundation Graduate Research Fellowship Program under Grant DGE-1147474 (to K.W.), and by a William R. and Sara Hart Kimball Stanford Graduate Fellowship (to K.W.).

5. Gray K, Wautz A, Young L (2012) The moral dyad: A fundamental template unifying social abilities: feeling love, recognizing someone, communicating with others, experiencing guilt, feeling disrespected, understanding how others are feeling, feeling embarrassed.