Reasoning about sentence and animacy:
Children’s and adults’ inferences about the properties of unseen entities

Kara Weisman
(kweisman@stanford.edu)
Department of Psychology
Stanford University

Ellen M. Markman
(markman@stanford.edu)
Department of Psychology
Stanford University

Carol S. Dweck
(dweck@stanford.edu)
Department of Psychology
Stanford University

Abstract

One striking finding in developmental and cognitive psychology is that people make rich inferences about the intentions and experiences of objects that look nothing like humans or animals. What makes these scenarios appear social, not just mechanical? Three studies explore this foundational level of social cognition: the detection of sentience. We probe inferences among what we posit to be core components of the concept of sentience—affect, autonomy, and perception—as well as physical markers of inanimacy. We find that children and adults share the belief that a fact about one of these three “sentient properties” implies the presence of others, to a moderate degree. Meanwhile, information about sentience blocks inferences of inanimacy. This link between sentience and animacy is particularly strong for US adults and White children, while people from other cultural backgrounds demonstrate a more flexible construal of what kinds of objects might be sentient.

Keywords: inference; sentence; animacy; social cognitive development.

Introduction

Our world is full of sentient beings. Many of these creatures are other humans, but other entities, even unfamiliar ones, can also evoke social responses. Certain behaviors—such as responding contingently, or pursuing a goal—trigger quick, perhaps irresistible, attributions of sentience, whether performed by an animal, a piece of technology, or something altogether unfamiliar. What are the conceptual underpinnings of this most basic level of social cognition?

Even infants are exquisitely sensitive to the presence of sentient creatures in their environment. In the burgeoning field of social cognitive development, seemingly sparse experimental displays elicit rich, spontaneous inferences about such complex phenomena as contingent responding and shared attention (Johnson, Slaughter, & Carey, 1998), goal pursuit and rational action (Csibra, et al., 1999; Luo, 2011), helping/hindering (Hamlin, Wynn, & Bloom, 2007), emotions (Skerry & Spelke, 2014), attachment (Johnson, Dweck, & Chen, 2007), and dominance (Mascaro & Csibra, 2012; Thomsen, et al., 2011) One striking finding from this work is that many of the protagonists that elicit social reasoning look nothing like humans or animals. Instead, they are polygons with minimal eyes, featureless ellipses, wooden boxes, even “blobs” of fiberfill. What renders these scenarios social, rather than merely mechanical? What do such events imply about their protagonists that facilitates reasoning about their experiences and relationships?

A close analysis of these studies suggests that the seemingly simple behaviors used in investigations of social reasoning can be deconstructed into multiple lower-level capacities. An entity that responds contingently, such as a blob that beeps in response to a baby’s vocalizations (Johnson et al., 1998), must perceive its partner’s actions, evaluate this partner as someone it wants to interact with, and generate actions of its own. Likewise, in a typical goal pursuit display (Csibra et al., 1999), an agent perceives a goal, evaluates it as desirable, propels itself toward the goal, adjusts its path to avoid obstacles, and stops upon arrival, at which point it might express happiness about its success. On this analysis, each of these ostensibly “sparse” displays actually presents a rich set of cues to more basic properties of sentience, offering ample evidence to the observer that the target is some sort of creature, and not an inert object.

What are these basic properties? The present studies investigate three low-level capacities implied by many of the events depicted in studies of early social cognition: affect, the experience of affective states, such as positive or negative valence; autonomy, the generation of spontaneous behaviors, such as movements or noises; and perception, the detection of information about the environment, such as sights or sounds. We posit that these capacities are core components of the lay concept of sentience.

By using audiovisual displays of entities responding to social partners, pursuing goals, navigating environments, or engaging in other observable behaviors, previous studies have presented cues to affect, autonomy, and perception in combination. Thus, little is known about the conceptual links between these capacities, or the individual roles they might play in reasoning about sentient creatures. In principle, affect, autonomy, and perception are causally independent: In order for an entity to have one of these abilities it need not have the others. Nonetheless, children or adults might consider some of these capacities to “go together”: e.g., if something is moving around on its own, this might imply that it can also perceive the environment, or experience emotions. The present studies probe the conceptual connections between affect, autonomy, and perception by examining whether learning a fact about one of these capacities licenses inferences about the others. We compare inferences among these three sentient capacities to inferences involving physical cues to inanimacy. We first examine these inferences in US adults.
Study 1

Methods
Participants. 89 adults participated via Mechanical Turk. Participants had gained approval for ≥95% of previous work (≥50 assignments); had US IP addresses; and indicated that they were ≥18 years old; and were paid $0.36 for about 6-7 minutes of their time. Three participants were excluded for failing to complete the survey, and six participants were randomly excluded to preserve counterbalancing, leaving a final sample of 80 participants.

Materials and Procedure. Our method was designed with the primary goal of presenting facts about unknown entities in isolation, rather than in combination. We conveyed information about sentience and animacy via verbal descriptions of unseen targets, rather than using audiovisual depictions of these targets (as in previous studies). Participants were presented with an illustrated story in which a character talked about a series of target entities. On each trial, the character looked into an opaque box, provided one fact about the sentience or animacy of the target inside, and then asked a question about that target’s other properties: e.g., Wow, this one can hear me talking! Hm, does that mean it can be in a bad mood sometimes?

Facts and questions were drawn from the three categories hypothesized to be core components of the concept of sentience—affect (e.g., feels happy right now), autonomy (e.g., is moving around on its own), and perception (e.g., can see what the box looks like)—as well as a fourth category of material cues to inanimacy (e.g., is made out of plastic), yielding 16 possible pairings of fact and question categories. Each participant was presented with 8 of the 16 fact-question pairings; categories appeared equally often in the fact and question positions and in the first and second halves of the testing session, within and across participants. Participants responded on a 4-point scale from Really no to Really yes. Sessions began with 3 warm-up trials intended to familiarize participants with the paradigm and to provide practice using both ends of the response scale.

Analysis Plan and Predictions
Scoring. We scored responses of Really no as -1.5, Maybe no as -0.5, Maybe yes as 0.5, and Really yes as 1.5; this created a (hypothetical) neutral midpoint of 0.

Comparison to neutral. To examine responses to each of the possible pairings of fact and question categories, we conducted a mixed effects linear regression with a random intercept for subject, comparing each of the 16 fact-question pairings to the neutral midpoint of our response scale. We predicted that participants would respond negatively on “inanimate” trials, which probed inferences between sentient properties and inanimate material composition. “Sentient-only” trials, which probed inferences from one sentient property to another (e.g., affect to perception), served to test whether participants considered these capacities to be conceptually linked (in which case they should response positively), or independent (in which case mean responses should not differ from the midpoint).

Planned contrasts. To compare responses to different pairings of fact and question categories, we conducted a separate regression analysis with 10 orthogonal contrasts. Our strongest theoretical prediction was that participants would respond more positively on sentient-only trials than on inanimate trials. We also explored whether, within sentient-only trials, participants would respond more positively on “within-category” trials, in which both the fact and the question were about the same core component, compared to “between-category” trials, in which they were about different core components. See Table 1 for the full list of contrasts, including comparisons of responses to facts and questions about specific sentient capacities (affect vs. other categories, autonomy vs. perception).

Results and Discussion
Comparison to neutral. In general, adults considered capacities for affect, autonomy, and perception to be conceptually linked: Mean responses were moderately positive on sentient-only trials (1.17<r<8.91), suggesting that the fact that an unseen entity had one of these sentient properties implied that it might also have other sentient properties. In addition, learning about the target’s capacity for affect, perception, or autonomy was generally sufficient for adults to rule out the possibility that it was inanimate, as reflected by their strong negative responses to inferences from sentient properties to inanimate materials (r<0.03, t<6.33). Likewise, inanimate material composition was sufficient to rule out the possibility that the target entity had sentient capacities (r<7.50). See Figure 1 (left), for mean responses by fact and question category.

Planned contrasts. Because adults endorsed inferences among sentient properties and rejected inferences between sentient properties and inanimacy, responses were more positive on sentient-only (M=0.34) vs. inanimate (M=1.03) trials. Responses to within-category inferences were more positive (M=0.57) than responses to between-category inferences (M=0.23), but this was primarily driven by exceptionally positive responses to inferences within the affect category (M=0.95). The strength of these responses could be due in part to our temporal framing: This one feels happy right now. Does that mean it can be in a bad mood sometimes? They are also consistent with the possibility that participants conceived of affective experience as inherently continuous, rather than composed of discrete emotions.2 See Table 1 for the full results of this analysis.

1 A separate analysis of absolute values confirmed that negative responses on inanimate trials were further from the midpoint than positive responses on sentient-only trials, b=0.03, t=16.62.
2 An additional study replicated this finding in a design that included inferences between valence and arousal: e.g., This one feels calm right now. Does that mean it can feel happy sometimes?
Study 2

Study 2 probed the developmental progression of these psychological connections, documenting patterns of inferences among sentient properties and physical markers of animacy in a sample of preschool children.

Methods

Participants. 64 US children ages 4.45–5.68 years (M=4.95 years; 33 boys) participated at a university preschool. School records indicated that 23 (36%) were White and no other race/ethnicity; small numbers of children were Indian (3), Hispanic/Latino (3), African-American (2), East Asian (2), or Middle Eastern (2) and no other race/ethnicity; and 24 (38%) were multiracial. Ethnicity/race was unknown for 5 children (8%). An additional 5 children participated but were excluded for not finishing the study (2), being outside of the target age range (2), or not speaking English (1).

Materials and procedure. Study 2 followed the same procedure as Study 1, with the following exceptions. An experimenter read the story aloud, and children gave verbal responses; typically they spontaneously said Yes or No and were then prompted for a more fine-grained response (e.g., Sort of yes, or really yes?). To accommodate limits on children’s attention, a short break was inserted halfway through test trials, during which the experimenter and child completed an easy puzzle featuring an ice cream cone.

Results and Discussion

On the whole, children’s responses in Study 2 were strikingly similar to those of adults in Study 1. By the age of five years, children seem to have a generally adult-like way of reasoning about capacities for affect, autonomy and perception, considering these three capacities to imply each other to a moderate degree.

Comparison to neutral. Like adults, children considered capacities for affect, autonomy, and perception to be positively related, and generally considered sentient capacities and inanimate material composition to be unlikely to co-occur. Accordingly, on sentient-only trials children’s mean responses were positive (1.30<3.93), and on inanimate trials mean responses were generally somewhat negative (-2.69<-0.30). See Figure 1 (center).

Planned contrasts. Like adults, children gave more positive responses on sentient-only trials (M=0.52) than on inanimate trials (M=0.33). Unlike adults, children gave equally positive responses on within- and between-category sentient-only trials (within: M=0.59, between: M=0.49). Inferences from one kind of affect to another were quite positive, but not exceptional. See Table 1.

Cultural effects. Although we had no a priori hypotheses about cultural differences, the children in our study were quite ethnically diverse, allowing us to conduct exploratory analyses of differences in patterns of responding by racial/ethnic background. For these analyses, we excluded children whose racial/ethnic backgrounds were unknown (N=5) and considered children with at least one non-White parent to be “children of color,” leaving us with a sample of 23 White children and 36 children of color. Among children of color, 81% (N=29) had ≥1 parent with Indian, East Asian, or Middle Eastern heritage. White children and children of color did not differ in mean age (t(57)=0.45, p=0.655) or in gender distribution (χ²(1)=0.24, p=0.625).

For both comparisons to the midpoint and planned contrasts analyses, including interactions with race/ethnicity improved the fit of our models (χ²(16)=28.64, p=0.026; χ²(11)=18.35, p=0.074; respectively). Of particular interest

\[3\] In contrast to Study 1, a separate analysis indicated that negative responses on inanimate trials were closer to the midpoint than positive responses on sentient-only trials, \(b=-0.01, t=-2.68\).
is that while, across the board, responses were more positive for sentient-only trials than for inanimate trials ($b=0.03$, $t=4.54$), this effect was exaggerated for White children, relative to children of color ($b=0.04$, $t=3.15$).

Comparison to adults. The general pattern of responses was highly similar for adults and children, although children’s responses were generally more positive than adults’. Whereas adults gave tempered positive responses on sentient-only trials and strong negative responses on inanimate trials, children gave positive responses on sentient-only trials and more tempered (though still mostly negative) responses on inanimate trials. In the aggregate, children were more willing to entertain the idea that an entity that engages in autonomous behavior could be composed of metal, plastic, glass, or clay; children’s responses to questions about the potential sentient properties of inanimate objects were likewise more moderate than adults.’

A supplemental regression analysis confirmed that children’s responses were more positive than adults’ ($b=0.42$, $t=4.75$), and the difference between sentient-only and inanimate trials was attenuated for children ($b=-0.03$, $t=-5.56$). Children also gave more positive responses than adults on sentient-only trials when the base fact was about autonomy ($b=0.16$, $t=1.97$), and on inanimate trials when the question was about affect ($b=0.23$, $t=2.78$).

**Table 1:** Results of mixed effects linear regressions with random intercepts for subjects and planned orthogonal contrasts.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Study 1 (US adults)</th>
<th>Study 2 (US children)</th>
<th>Study 3 (Ind. adults)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>$t$</td>
<td>$p$</td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td>-0.26</td>
<td>-7.40</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>All trials</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Sentient-only vs. inanimate</td>
<td>0.09</td>
<td>26.50</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Sentient-only trials</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Within- vs. between-categories</td>
<td>0.11</td>
<td>4.58</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3. Fact: affect vs. other categories</td>
<td>0.06</td>
<td>2.77</td>
<td>0.006</td>
</tr>
<tr>
<td>4. Fact: autonomy vs. perception</td>
<td>-0.02</td>
<td>-0.41</td>
<td>0.683</td>
</tr>
<tr>
<td>5. Question: affect vs. other categories</td>
<td>0.08</td>
<td>3.11</td>
<td>0.002</td>
</tr>
<tr>
<td>6. Question: autonomy vs. perception</td>
<td>0.04</td>
<td>1.00</td>
<td>0.318</td>
</tr>
<tr>
<td><strong>Inanimate trials</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Fact: affect vs. other categories</td>
<td>-0.06</td>
<td>-1.49</td>
<td>0.136</td>
</tr>
<tr>
<td>8. Fact: autonomy vs. perception</td>
<td>0.21</td>
<td>2.86</td>
<td>0.004</td>
</tr>
<tr>
<td>9. Question: affect vs. other categories</td>
<td>-0.11</td>
<td>-2.68</td>
<td>0.008</td>
</tr>
<tr>
<td>10. Question: autonomy vs. perception</td>
<td>0.09</td>
<td>1.15</td>
<td>0.253</td>
</tr>
</tbody>
</table>

---

* 4 This model also revealed that responses on inanimate trials were less negative when the fact was about autonomy, compared to perception ($b=0.42$, $t=2.33$); this effect was marginally smaller for White children ($b=-0.50$, $t=1.74$).

* 5 A supplemental regression analysis confirmed that children’s responses were more positive than adults’ ($b=0.42$, $t=4.75$), and the difference between sentient-only and inanimate trials was attenuated for children ($b=-0.03$, $t=-5.56$). Children also gave more positive responses than adults on sentient-only trials when the base fact was about autonomy ($b=0.16$, $t=1.97$), and on inanimate trials when the question was about affect ($b=0.23$, $t=2.78$).

**Study 3**

Exploratory analyses of children’s response patterns suggested that conceptual connections between sentient properties and physical markers of animacy might vary by children’s cultural exposure, with children of color (most of whom were of Indian, East Asian, or Middle Eastern heritage) demonstrating more tolerance of the idea that an entity might be both sentient and inanimate. Although unpredicted, this finding is in line with known cultural differences in both domain-general cognitive styles (Nisbett, Peng, Choi, & Norenzayan, 2001) and domain-specific beliefs about the ontological and moral status of inanimate objects (see Footnote 8). If children’s construals of sentience and animacy depend on cultural input, we would predict that adults from different cultures—particularly cultural contexts that are known to differ in cognitive styles and religious/philosophical orientations—would also vary in their profile of inferences on our task. To this end, Study 3 probed these inferences in a sample of Indian adults who were likely embedded in a context that encourages more continuous and dialectical reasoning (Nisbett et al., 2001).

**Methods**

Participants. 80 English-speaking adults with Indian IP addresses participated via Mechanical Turk.

**Results and Discussion**

Comparison to neutral. Like US adults, Indian adults considered capacities for affect, autonomy, and perception to be positively related: Mean responses were moderately positive on sentient-only trials (0.33<-$t$<3.74). Indian adults considered sentient capacities to be somewhat unlikely to co-occur with inanimate material composition, as reflected...
by their negative responses on inanimate trials (-5.53 < r < -1.95). See Figure 1 (right).

**Planned contrasts.** Like US adults, Indian adults gave more positive responses on sentient-only trials (M=0.26) than on inanimate trials (M=-0.53). Unlike US adults, Indian adults gave equally positive responses on within- (M=0.38) and between-category (M=0.20) sentient-only trials and did not respond particularly positively to within-category inferences about affect (M=0.25). See Table 1.

**Comparison to US adults.** Overall, Indian adults’ pattern of responses was very similar to that of US adults. However, Indian adults were more moderate in their responses, particularly on inanimate trials: While US adults considered sentient properties sufficient grounds to rule out the possibility of inanimacy (and vice versa), Indian adults indicated that it was unlikely—but not impossible—that inanimate objects could have sentient properties, or that sentient beings could be composed of metal, plastic, glass, or clay. As predicted, this difference echoes the post-hoc finding in our child sample, with similar patterns of responses for White children and US adults on the one hand, and for children of color and Indian adults on the other.7

**General Discussion**

In these studies, we set out to explore the conceptual underpinnings of attributions of sentience. These studies demonstrated that adults and young children consider capacities for affect, autonomy, and perception to mutually imply each other to a reliable, moderate degree. By five years of age, children and adults from multiple cultural backgrounds converge on a similar pattern of inferences among these capacities, suggesting that they share a similar construal of the psychological connections among what we consider central components of the concept of sentience.

Across our studies, inferences between affect, autonomy, and perception were generally positive, but quite modest in strength. At the outset of these studies, it seemed plausible that the psychological connections between at least some pairs of properties might be strong enough to elicit stronger inferences in our task: It could have been the case that learning that an unseen target is happy, or can hear someone talking, or is moving itself around would be enough to trigger a whole suite of sentient properties, leading participants to strongly endorse inferences about a variety of other capacities. Instead, both adults and children tended to give responses that were only slightly above the neutral midpoint of our scale, both in the aggregate and at the subject level. Although isolated information about a capacity for affect, autonomy, or perception was sufficient to suggest the possibility of other sentient properties, it was not enough to fully convince participants that the unseen target possessed the entire suite of sentient properties. The adults and children in our studies seem to have adopted a flexible stance regarding the specific capacities that a sentient creature might possess: While in the absence of other information they might expect such a creature to have some variety of affective, autonomous, and perceptual abilities, they would not be “thrown off” by a creature that did not have some specific ability (e.g., to see, or to make noises). In fact, many individual creatures in the real world, including humans, lack one or more of the specific capacities included in the present studies—if every such instance constituted a major violation of people’s concept of sentience, social reasoning would be much more difficult.

In light of our findings, we might reconsider the traditional characterization of experimental displays in the recent social cognitive literature as simple or sparse: The fact that isolated verbal information about a single sentient property does not license stronger inferences about other sentient properties emphasizes the rich amount of information contained in more traditional audiovisual displays. In our studies, information about an entity’s capacity for, e.g., autonomous movement was not redundant with information about its perceptual abilities or its affective experiences. Thus, behaviors such as navigating through an environment, which requires both perceptual access and a capacity for autonomous behavior, should be considered to contain rich information about multiple aspects of the identity or nature of the target—i.e., multiple, modestly correlated cues to sentience. Although scenes portraying social contingency, goal pursuit, helping and hindering, etc., are beautifully controlled with regard to the phenomena they were created to depict, they are nonetheless incredibly rich in content from the perspective of detecting which components of the displays are potential social partners.

How does this understanding of sentience relate to concepts of animacy? In contrast to the consistent pattern of moderately positive inferences between sentient properties, which held true across our samples, participants from different cultural groups diverged in their responses to questions about whether entities could be both sentient and inanimate. US adults considered it highly unlikely that something made of metal, plastic, glass, or clay could have any of the sentient properties included in our studies—or that an entity with a capacity for affect, autonomy, or perception could be composed of these materials. In contrast, Indian adults were more moderate in their responses to these questions, indicating that it was not impossible that inanimate entities could have sentient properties. Among children attending a university preschool in the US, White children’s responses were very similar to those of US adults, while, in the aggregate, children of color—most of whose ethnic backgrounds suggested

---

6 As in Study 1, a separate analysis confirmed that negative responses on inanimate trials were further from the midpoint than positive responses on sentient-only trials (b=0.02, t=8.07).

7 A supplemental analysis revealed that US adults gave more negative responses overall (b=-0.15, t=-2.30), and the difference in responses on sentient-only vs. inanimate trials was exaggerated for US adults (b=0.03, t=5.96). Exceptionally positive responses to inferences within the affect category were particular to US adults (b=0.09, t=2.22).
exposure to East or South Asian cultures—gave responses that were more similar to those of Indian adults.

Following work in cultural psychology (e.g., Markus & Kitayama, 1991), we speculate that these diverging response patterns might emerge as a result of learning both domain-general cognitive styles and domain-specific beliefs. From a cognitive perspective, a liberal understanding of sentience and animacy as non-mutually exclusive is congruent with the “holistic” reasoning style common to many Eastern cultural contexts, in which continuity and relationships between entities are the focus of thought, and dialectical oppositions are tolerated and even valued. In contrast, a sharp distinction between sentient animate beings, on the one hand, and non-sentient inanimate objects, on the other hand, is more consistent with the “analytic” reasoning style dominant in the Western world, which deals in individual objects, belonging to discrete categories and bound by rules of logic (Nisbett et al., 2001). In combination with religious and philosophical beliefs about sentience, souls, and minds, these cognitive styles could plausibly support cultural differences in the degree to which information about sentience licenses inferences about animacy, and vice versa.

Taken together, our studies provide evidence of culturally invariant conceptual connections among affect, autonomy, and perception, while also suggesting possible cultural differences in how people reason about what kinds of entities might have these sentient properties. Attritions of sentience or mind are critical to both psychological and philosophical accounts of sociomoral reasoning (Gray, Young, & Watz, 2012), and the inferences documented in these studies may play a critical role in allowing people to abstract away from direct observations of unfamiliar entities (human or otherwise) toward conceptual representations of these entities as sentient, facilitating decisions about whether to accord them moral status. Our findings suggest that different cultures might build on these connections in different ways, perhaps connecting them to categorical construals of biological animacy, or instead employing a more fluid representation of the physical instantiation of sentient beings. In turn, these construals might influence how people in these communities interact with the wide variety of creatures in their worlds.

Acknowledgments

Special thanks to the staff and families at Bing Nursery School, and to Co Tran. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-114747, and by a William R. & Sara Hart Kimball Stanford Graduate Fellowship.

References

Skerry, A. E., & Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. Cognition, 130(2), 204-16.

---

8 E.g., inanimate objects are thought to be sentient in some Buddhist traditions (Rambelli, 2007), and the Shinto term kami refers to the spirit or essence of individual people, places, natural objects, and some tools (Ono & Woodard, 1962)—whereas souls in the Judeo-Christian tradition reside only in beings that have been given “the breath of life” (Genesis 2:7, King James Bible).