Revisiting the Valence Account

Justin Sytsma

The existence of phenomenally conscious mental states is often taken to be obvious from first-person experience. Sytsma and Machery (2010) argued that if that is the case, then lay people should classify mental states in the same way that philosophers typically do, treating states like seeing red and feeling pain similarly. We then presented evidence that they do not. This finding is interesting in its own right, however, outside of any implications for the philosophical debates concerning phenomenal consciousness. As such, we attempted to explain our finding, presenting evidence that lay mental state ascriptions depend on valence judgments (that the mental states have a hedonic value for the subject). In this paper, I present new evidence that suggests against this valence account. I then provide evidence for a new explanation based on previous findings that lay people tend to view both colors and pains as mind-independent qualities of objects outside the mind/brain.

In Sytsma and Machery (2010), we presented experimental evidence for two primary hypotheses—what I will call the *negative hypothesis* and the *positive hypothesis*.[1] According to the negative hypothesis, lay people (viz., people without training in philosophy or consciousness studies) by and large lack the philosophical concept of phenomenal consciousness. By this we meant that their mental state attributions often diverge from what we would predict if they were calling on the distinction between phenomenal states and intentional states.

Based on the evidence for the negative hypothesis, we argued that a certain bit of philosophical commonsense is mistaken, and that this has important implications for recent debates in philosophy of mind. Specifically, the existence of phenomenally conscious mental states has been thought to result in an explanatory gap (Levine, 1983) or a hard problem of consciousness (Chalmers, 1995). That such mental states exist, however, is often justified by simply claiming that they are obvious on the basis of first-person experience with states like

[1] This work falls within the burgeoning sub-field of experimental philosophy of mind. For an accessible review of some of the work being done in this area, see Machery and Sytsma (2011). For a more extended survey, see Sytsma (2010a). For a collection of cutting-edge articles, see Sytsma (forthcoming).

seeing red and feeling pain. But if this is correct, then lay people should treat these states similarly, denying, for example, that a simple non-humanoid robot can be in either state—and it turns out that they do not. Like philosophers, they tend to deny that the robot feels pain; unlike philosophers, however, they tend to hold that it sees red.

Having presented evidence for our critique of the philosophical tradition, we went on to ask a follow-up question: Why do lay people ascribe seeing red to the robot, but not feeling pain? Based on the results of two further studies, we put forward our positive hypothesis—the *valence account* of lay mental state ascriptions. We argued that lay people treat seeing red differently from feeling pain because the latter is associated with a valence (or a hedonic value), while the former is not.[2]

I no longer believe that our positive hypothesis is correct, however. And in this paper I will present new evidence that suggests against the valence account. At the same time, I will argue that the evidence continues to support our negative hypothesis. In fact, I will argue for an alternative account—the *naïve account*—which posits that lay people conceive of states like seeing red and feeling pain in a way that fundamentally diverges from the philosophical tradition, further substantiating the negative hypothesis.

Here is how I will proceed. In Section 1, I describe the philosophical concept of phenomenal consciousness and briefly review the evidence for the negative hypothesis. In Section 2, I discuss the positive hypothesis. I then present two new studies that suggest against the positive hypothesis in Section 3. Finally, in Section 4, I present evidence for the alternative naïve account.

[2] Sytsma and Machery (2010, 300) take mental states to have a valence if and only if they have a hedonic value for the subject—that is, if they are pleasurable (positive valence) or disagreeable (negative valence). Further, not all mental states have a valence (or are valence neutral), and not all valenced mental states are equally pleasurable or disagreeable.

1. The Negative Hypothesis

My goal in this paper is to argue for an alternative to the valence account of lay mental state ascriptions given by Sytsma and Machery (2010). The valence account was put forward to explain the evidence provided for our negative hypothesis, however. As such, the first order of business is to review that evidence. And to do this effectively, it is best to begin with an articulation of the philosophical concept of phenomenal consciousness.

1.1 Phenomenal Consciousness

What is the philosophical concept of phenomenal consciousness? Unfortunately, this is not an easy question to answer, since different philosophers articulate the concept in somewhat different ways. Nonetheless, for purposes of assessing whether lay people share the concept, we can focus on the common understanding. And what we find is that philosophers generally talk of *mental states* being phenomenally conscious, where this is taken to indicate that there is "something it is like" (Nagel, 1974) for an entity to have those mental states. The "something it is like" is then specified by noting various episodes in which we are acquainted with certain distinctive qualities. These qualities are generally referred to as *qualia*, and a diverse range of mental states are thought to have qualia. Thus, phenomenally conscious mental states are typically said to include *perceptual experiences* like seeing red, *bodily sensations* like feeling pain, as well as *felt emotions* and *felt moods*. Such states are often contrasted with intentional mental states such as beliefs and desires, which are not generally considered to be phenomenally conscious.[3]

To illustrate, consider the definition of "qualia" given by Marina Rakova in *Philosophy of Mind A–Z* (2006, 152–153):

[3] There are some exceptions, however; see, for example, Horgan and Graham (2012).

> [Qualia are] qualitative or phenomenal properties of experience. Seeing red, feeling pain or tasting a lemon feel to us a certain way, there is something that it is like for us to have these experiences. It means that these mental states have qualia. Most philosophers agree that mental states with qualia include experiences of perception (hearing a loud noise), bodily sensations (feeling an itch), felt emotions (fear) and moods (depression). Galen Strawson argues that thoughts also have qualia, but this is not the accepted view.

On this definition, qualia are qualities (properties) *of* certain mental states (experiences); as such, the mental states are said to *have* those qualities. And the qualities at issue are presumably those that we are aware of in seeing red, feeling pain, tasting a lemon, and so on—the red seen, the pain felt, the sourness tasted, et cetera.[4]

I will refer to such qualities—the distinctive qualities that we are aware of in perceptual experiences, bodily sensations, felt emotions, and felt moods—as *sensory qualities*. In saying that these qualities are qualia, however, the sensory qualities are not being considered on their own, but through the lens of corresponding mental states: A sensory quality must be had by a mental state for it to be a quale. I will express this by saying that qualia are *mind-dependent*. Thus, if no mental states existed (if, for example, all sentient beings in the universe were destroyed), then there would be no qualia. Following this account of qualia, I will say that lay people share the philosophical concept of phenomenal consciousness insofar as they treat

[4] And many similar illustrations could be given. For example, in her *Routledge Encyclopedia of Philosophy* entry on "Qualia," Janet Levin writes: "In contemporary discussions in the philosophy of mind, the terms quale and qualia (plural) are most commonly used to denote features of our conscious mental states such as the throbbing pain of my headache, the warmth I feel when I hold my hands over the fire, or the greenish character of my visual experience when I look at the tree outside my window (or stare hard at something red and then close my eyes)." (2009, §0). Similarly, in his *Stanford Encyclopedia of Philosophy* entry on "Qualia," Michael Tye writes: "I run my fingers over sandpaper, smell a skunk, feel a sharp pain in my finger, seem to see bright purple, become extremely angry. In each of these cases, I am the subject of a mental state with a very distinctive subjective character. There is something it is *like* for me to undergo each state, some phenomenology that it has. Philosophers often use the term 'qualia' (singular 'quale') to refer to the introspectively accessible, phenomenal aspects of our mental lives. In this standard, broad sense of the term, it is difficult to deny that there are qualia. Disagreement typically centers on which mental states have qualia, whether qualia are intrinsic qualities of their bearers, and how qualia relate to the physical world both inside and outside the head." (2007, §0). The key point to note is that qualia are taken to be mind-dependent in each passage: Qualia are qualities (features, characters) of mental states, or qualities that mental states have, such that if there were no mental states, there would be no qualia.

perceptual experiences, bodily sensations, felt emotions, and felt moods similarly, taking the qualities that they are acquainted with in these episodes to be mind-dependent.[5]

1.2 The Case for the Negative Hypothesis

The current evidence suggests that lay people do not share the philosophical concept of phenomenal consciousness.[6] Most importantly, Sytsma and Machery (2010) present evidence that unlike philosophers, lay people do not treat two prototypical examples of (supposed) phenomenally conscious mental states similarly. Thus, in our first study, we gave participants—both lay people and philosophers—either a description of a simple, non-humanoid robot (Jimmy) or of a normal human (Timmy) performing behaviorally analogous tasks expected to elicit ascriptions of either a perceptual experience or a bodily sensation for the human. In each scenario the agent (robot or human) manipulated a specific box based on color. In half of the scenarios, the manipulation was successful and participants were asked whether the agent saw red; in the other half, the agent was electrically shocked and participants were asked whether the agent felt pain.

As expected, the philosophers surveyed treated feeling pain and seeing red analogously. They were unwilling to ascribe either the perceptual experience of seeing red or the bodily sensation of feeling pain to the robot Jimmy, but were willing to ascribe both states to the human

[5] At the same time, it should be noted that things are somewhat more complicated than this account suggests. Thus, some philosophers deny that qualia are qualities *of* phenomenally conscious mental states, holding that these mental states *have* qualia in another way. Specifically, some hold that qualia are sensory qualities that are represented by phenomenally conscious mental states, and are externalists about the representational content of those mental states (e.g., Byrne and Tye, 2006). Even on this type of "wide representationalist" view, however, qualia are still held to be mind-dependent: When I am acquainted with redness in looking at a ripe tomato, for example, the wide representationalist holds that this sensory quality is part of the representational content of my mental state, and this is taken to be a necessary condition for its being a quale.

[6] It should be noted, however, that Knobe and Prinz (2008) offer support for the opposite claim. Most notably, in their second study they found that lay people were generally unwilling to ascribe mental states that philosophers typically take to be phenomenally conscious to a corporation, but were willing to ascribe intentional mental states to it. There is good reason to think that Knobe and Prinz's results reflect the specific set-up of their study, however, rather than that lay people have the concept of phenomenal consciousness (see Sytsma and Machery, 2009; Arico, 2010).

Timmy. Contrary to the claim that lay people classify mental states in the same way that philosophers do, however, non-philosophers tended to treat the perceptual state of seeing red quite differently from the bodily sensation of feeling pain. While they were willing to attribute seeing red to the robot, they were not willing to attribute feeling pain to it. And this finding has been replicated a number of times.[7]

Despite this, the support we offered for the negative hypothesis has met with resistance. The most common objection is that while lay people have the concept of phenomenal consciousness, they did not make use of it in our study.[8] Instead our results are taken to show that "seeing red" is ambiguous, being used in both an *informational sense* and a *phenomenal sense*.[9] While the informational sense only requires that an agent make relevant discriminations between visual stimuli (detecting the wavelengths of light reflected by different objects and distinguishing between them, for example), the phenomenal sense requires something more—that the agent be in a corresponding phenomenal state. The critic then argues that when lay people affirmed that the robot saw red, they understood the question in the informational sense.

We offered several replies to this ambiguity objection in our original article, focusing on the distribution of responses (they were not bimodal as we would expect if lay people recognize both an informational and a phenomenal reading) and the explanations given for negative responses (unlike the explanations given by philosophers, those given by lay people did not

[7] See, for example, the four studies reported in Sytsma and Machery (2012). Most importantly, in Study 4 we gave non-philosophers modified versions of the two probes from Study 1 in Sytsma and Machery (2010), removing any anthropomorphic language and changing the target color from red to blue. As expected, the participants tended to answer that the robot saw blue and tended to deny that it felt pain.

[8] For other objections, see Talbot (2012) and our response in Sytsma and Machery (2012), as well as Buckwalter and Phelan (forthcoming).

[9] For example, Huebner (2010, 137) notes that our finding that "non-philosophers tend to say that a simple robot can 'see red' but philosophers find such ascriptions odd" might be explained in terms of "looks" talk being polysemous: "On one interpretation of 'looks', seeing red requires only that a system has detected red things and reported having done so. However, on a more 'phenomenal' interpretation of 'looks' the entity would also have to be the subject of an immediate experience of red. But there seems to be no obvious way to guarantee that all participants in a psychological experiment will adopt the intended reading!"

suggest a phenomenal reading). A further reply was given in Sytsma (2009), and developed in Sytsma (2010b). In those articles, I argued that the objection supposes a theoretical view of colors that lay people generally do not hold.

To draw this out, I described a contrasting view of colors—the *naïve view*. On this view, the colors we are acquainted with in ordinary visual perception are not taken to be mind-dependent, but are treated as *mind-independent* qualities of things in the world. For example, the redness that we are acquainted with in looking at a ripe tomato is taken to be a quality of the tomato that would exist even if nobody was looking at it. And this mind-independent quality is taken to be integral to discerning the tomato's color, since what is being discerned just is the redness of the tomato. As such, making relevant color discriminations is not thought to be lacking in the way that is assumed for the informational reading of "seeing red." Thus, insofar as lay people hold a naïve view of colors, they are not best interpreted as asserting that the robot saw red in an informational sense, blocking the ambiguity objection.

In Sytsma (2010b), I offered empirical evidence that lay people typically hold a naïve view of colors. In my first study, participants were given a brief description of two views of colors—one corresponding with the naïve view, the other a phenomenal view. The participants were then asked a series of questions about how they understand colors. Consistent with the naïve view, but not the contrasting phenomenal view, a majority of the participants responded that the colors we see are qualities of objects outside the brain, denied that they are mind-dependent, and denied that spectrum inversion is possible.[10] Further, this finding was replicated in my second study.

[10] Participants were asked about spectrum inversion because most (if not quite all) philosophers of mind hold that commitment to the possibility of a spectrum inversion is tantamount to taking the colors that we are acquainted with in ordinary visual perception to be qualia. For example, before voicing his dissent, Michael Tye (1994, 160) writes

The results of these studies indicate that lay people tend to hold a naïve view of colors. This in turn suggests against the ambiguity objection: Lay people are not best interpreted as reading the question in an informational sense, as contrasted with a phenomenal sense, when they answer that the robot sees red in Sytsma and Machery's studies. As such, the current evidence continues to support our negative hypothesis.

2. The Positive Hypothesis

We have just seen evidence suggesting that lay people generally lack the philosophical concept of phenomenal consciousness. And this finding has important implications for recent debates in philosophy of mind. It also raises a further question, however: How *do* lay people classify mental states? Specifically, why do we see the stark difference between ascriptions of seeing red and feeling pain detailed in the previous section?

Sytsma and Machery (2010) offered two possible explanations for the finding that lay people do not treat these two prototypical examples of (supposed) phenomenally conscious mental states similarly. The first explanation we considered is the internal/external account: Lay people treat seeing red differently from feeling pain because the former is a product of an external sense, while the latter is a product of an internal sense. The second explanation we considered is the valence account: Lay people treat seeing red differently from feeling pain in our study because the latter is associated with a non-neutral valence, while the former is not, and because they hold that the simple robot Jimmy is incapable of having that valence.

To test these competing accounts, we ran two further studies looking at ascriptions of anger and smell to the simple robot used in our first study. We found that while lay people tend

---

that "both advocates and opponents of qualia seem agreed: Grant the Inverted Spectrum Hypothesis and perceptual qualia *must* be admitted."

to deny that it felt anger, they were split with regard to whether it smelled banana or smelled vomit (the mean response for each of these questions was close to the neutral point). These findings are arguably compatible with both accounts. Thus, the internal/external account can explain the neutral results for smelling banana and smelling vomit in terms of olfaction being borderline between an internal and an external sense; and, the valence account can explain the neutral results in terms of the valence associated with olfactory states being less essential than the valence associated with pain or anger.

As such, the valence account predicts that if people were less likely to focus on the valence associated with an olfactory mental state, then they would be more likely to attribute that state to a simple robot. We tested this prediction in our third study by varying the vignettes used in our first two studies to specify that Jimmy successfully manipulated a box based on one of three olfactory cues: The descriptor for the first cue was one that people are familiar with and tend to associate with a positive valence (banana), the descriptor for the second cue was one that people are familiar with and tend to associate with a negative valence (vomit), and the descriptor for the third cue is one that people are generally unfamiliar with and, thus, are unlikely to associate with a valence (isoamyl acetate). In line with our prediction, we found that participants were again split with regard to whether Jimmy smelled the olfactory cues with familiar descriptors; in contrast, they tended to affirm that Jimmy smelled the olfactory cue with the unfamiliar descriptor. While these results are consistent with the valence account, they are not readily explained by the internal/external account. As such, we tentatively concluded that lay people tend to distinguish between mental states on the basis of whether or not they are associated with a valence.

3. Evidence Against the Positive Hypothesis

In this section I present two new studies investigating the valence account.[11] I argue that despite the results of Sytsma and Machery's third study, the current evidence suggests against our positive hypothesis: Lay mental state ascriptions do not appear to correspond with valence judgments.

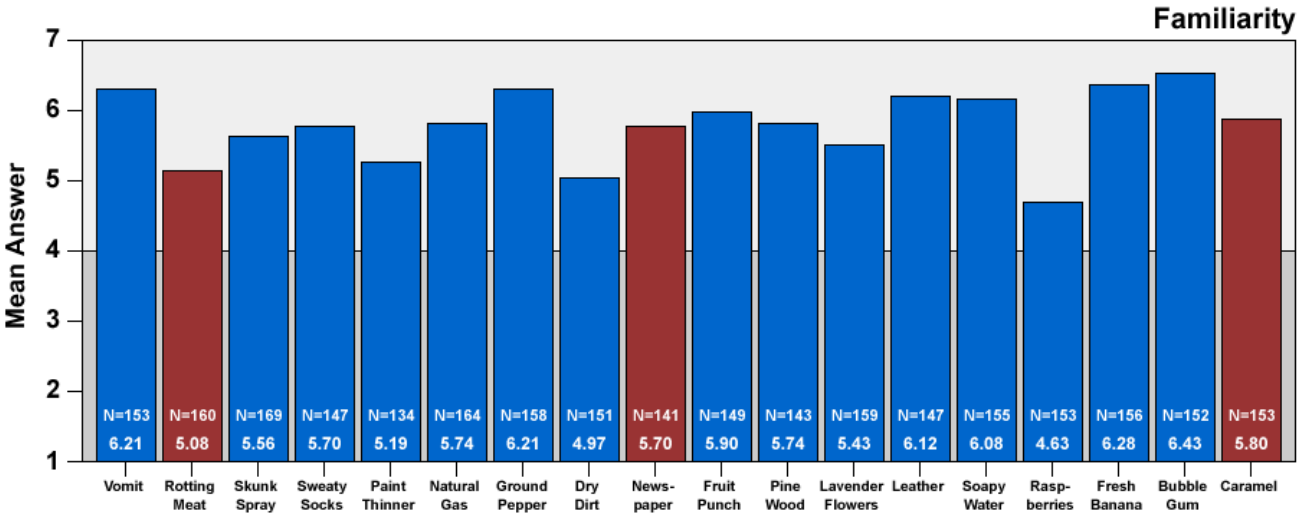
3.1 Study 1

To further test the valence account, I conducted a two-stage study. In the first stage, I solicited valence judgments for a variety of olfactory stimuli. In the second stage, I then checked to see whether lay people say that the simple robot Jimmy smelled selected stimuli from this list.

I began by compiling three lists of six terms for different smells, based on the work of Kamath, Turetsky, and Moberg (2011) and Royet et al. (2000). The first list consisted of terms for smells that people typically judge to have a positive valence (caramel, bubble gum, fresh bananas, raspberries, lavender flowers, fruit punch), the second a negative valence (natural gas, paint thinner, sweaty socks, skunk spray, rotting meat, vomit), and the third a neutral valence (soapy water, leather, pine wood, newspaper, dry dirt, ground pepper). In the first stage of my study participants were randomly given four of these terms from the list of 18. For each, participants were asked to imagine and reflect on the corresponding smell. Using a 7-point scale, they were then asked to rate the familiarity of the smell (1=not at all familiar, 4= neither familiar nor unfamiliar, 7=very familiar), the identifiability of the smell (1=certain I could not identify it, 4=not sure, 7=certain I could identify it), whether the smell is subtle or intense (1=extremely subtle, 4=neutral, 7=extremely intense), and whether the smell is pleasant or unpleasant (1=highly pleasant, 4=neutral, 7=highly unpleasant).

[11] These studies originated in discussions with Edouard Machery.

Responses were collected from 686 participants online using the Philosophical Personality website.[12] Participants were restricted to native English speakers, 18 years of age or older, with at most minimal training in philosophy, and who had not previously taken a survey through this website.[13] The mean ratings for familiarity, identifiability, intensity, and valence for each of the 18 terms are shown in Figure 1. Based on the results, I selected three terms from the list that were similar with regard to their combined mean ratings for familiarity, identifiability, and intensity—one where participants considered the corresponding smell to be pleasant (caramel), one neutral (newspaper), and one unpleasant (rotting meat).[14]



[12] http://www.PhilosophicalPersonality.com

[13] Participants were counted as having more than minimal training in philosophy if they were philosophy majors, had completed a degree with a major in philosophy, or had taken graduate-level courses in philosophy. The participants were 64.4% female, had an average age of 33.5 years, and ranged in age from 18–76 years.

[14] The combined mean rating for familiarity, identifiability, and intensity was 15.9 for caramel, 15.4 for newspaper, and 16.9 for rotting meat.

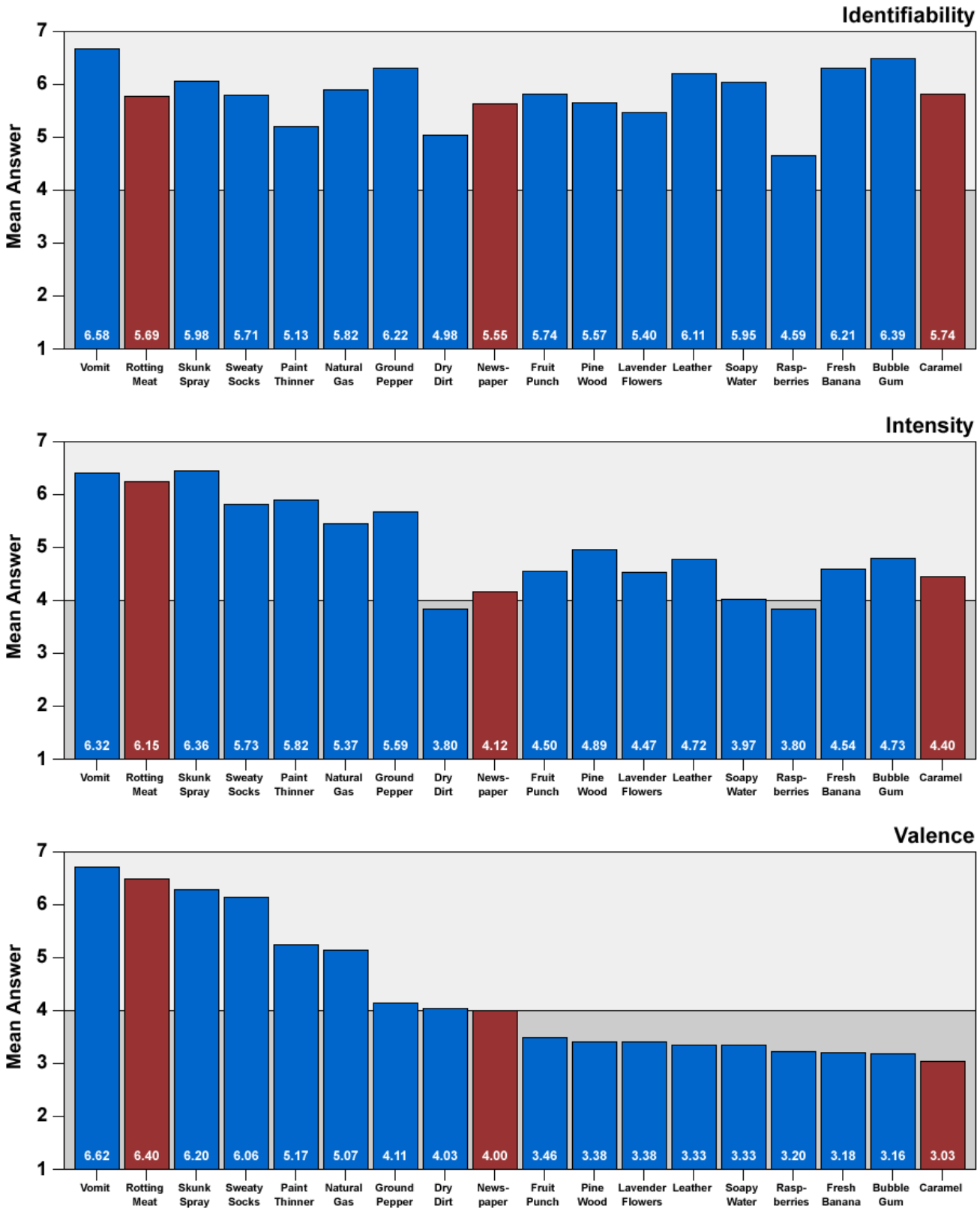


Figure 1: Results of Stage 1 of Study 1 (terms used in Stage 2 shown in red).

In the second stage of the study, 232 new participants were randomly given a probe using one of the three terms selected in Stage 1.[15] The probe for the pleasant stimulus reads as follows:

> Jimmy (shown below) is a relatively simple robot built at a state university. Jimmy is equipped with an odor detector, video camera, wheels for moving about, and two grasping arms for moving objects.
>
> 
>
> As part of an experiment, Jimmy was put into a room that was empty except for three boxes containing fragrance diffusion devices. Each box emitted a different odor. One box emitted the odor of caramel. Each of the other two boxes emitted an odor randomly selected from a long list of common substances. The boxes were identical in all respects except for the odors they emitted.
>
> Jimmy was instructed to put the box emitting the odor of caramel in front of the door. Jimmy performed the task correctly and with no noticeable difficulty. The test was then repeated on three consecutive days, changing the odors used for the other two boxes. For each test the order of the boxes was shuffled. Each time Jimmy performed the task correctly and with no noticeable difficulty.
>
> *Did Jimmy smell caramel?*

Participants answered the question on the same 7-point scale used by Sytsma and Machery (2010)—1 anchored with "clearly no," 4 with "not sure," and 7 with "clearly yes." On a second page, they were then asked the four questions used in Stage 1 for the selected term. The results are shown in Figure 2.

The valence account predicts that the mean response on the Jimmy question for newspaper (valence-neutral) would be higher than for either caramel (positive valence) or rotting meat (negative valence). While I found the predicted pattern of results, the differences between the three means were minimal; further, while the mean rating for newspaper (5.31) was higher than for either caramel (5.15) or rotting meat (4.96), neither difference was significant.[16] I also compared the valence ratings that these participants gave on the second page of the survey to

[15] The same website and restrictions used in Stage 1 were used in Stage 2. The participants for Stage 2 were 65.1% female, had an average age of 33.2 years, and ranged in age from 18–79 years.

[16] Newspaper and caramel: $t(145)=0.4359$, $p=0.3318$, one-tailed; newspaper and rotting meat: $t(163)=0.9906$, $p=0.1617$, one-tailed.

their responses to the question about Jimmy. While the correlation between the ratings on the Jimmy question and the valence question (coded in terms of distance from a neutral rating of 4) across the three probes was in the predicted direction, the correlation was quite small (-0.05164). Although these results perhaps suggest that valence plays some role in lay mental state ascriptions, they are rather underwhelming.

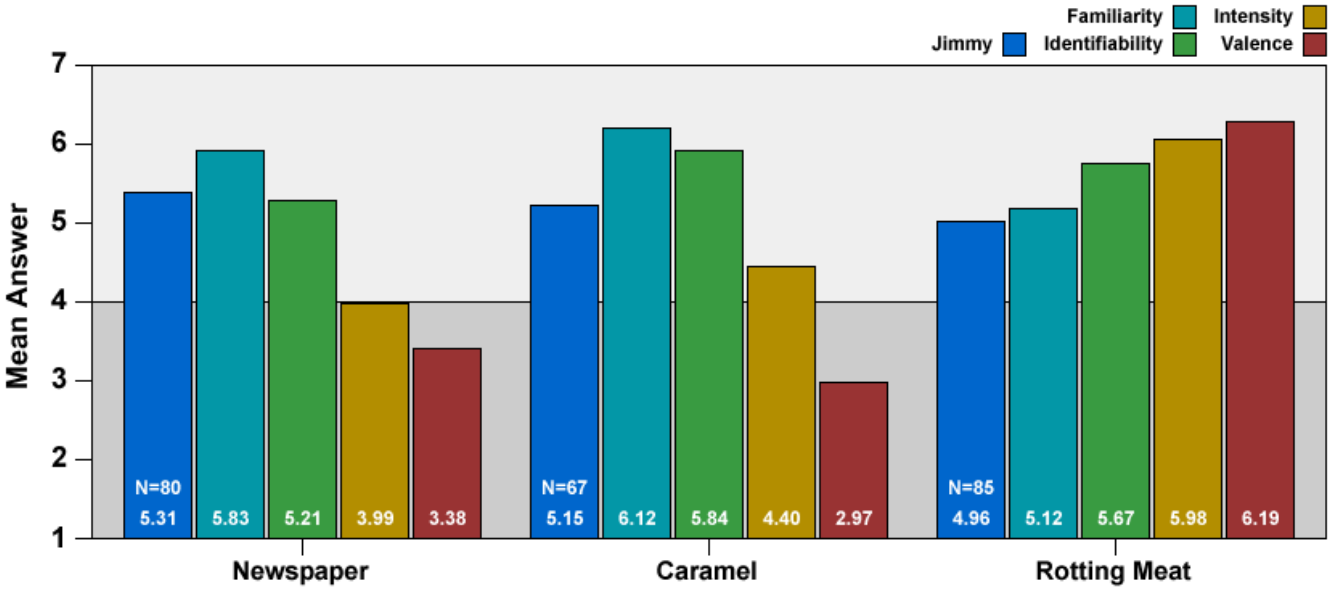


Figure 2: Results of Stage 2 of Study 1.

3.2 Study 2

To further test the valence account, I rewrote the Jimmy probe used in the first study to specify the odors being emitted from each box (rather than stating that "the other two boxes emitted an odor randomly selected from a long list of common substances"). Three variations on the probe were tested—one using descriptors that received neutral valence ratings in Stage 1 of the first study, one that used descriptors that received positive valence ratings, and one that used descriptors that received negative valence ratings. For the valence-neutral probe the three terms were dry dirt, newspaper, and ground pepper (with dry dirt as the target); for the positive valence

probe the three items were caramel, bubble gum, and fresh banana (with caramel as the target); and, for the negative valence probe the three items were paint thinner, rotting meat, and vomit (with paint thinner as the target). Each of the 293 participants in this study was randomly given one of the three probes, followed by the four questions from the first stage of the study for the target term on a second page.[17] The results are shown in Figure 3.

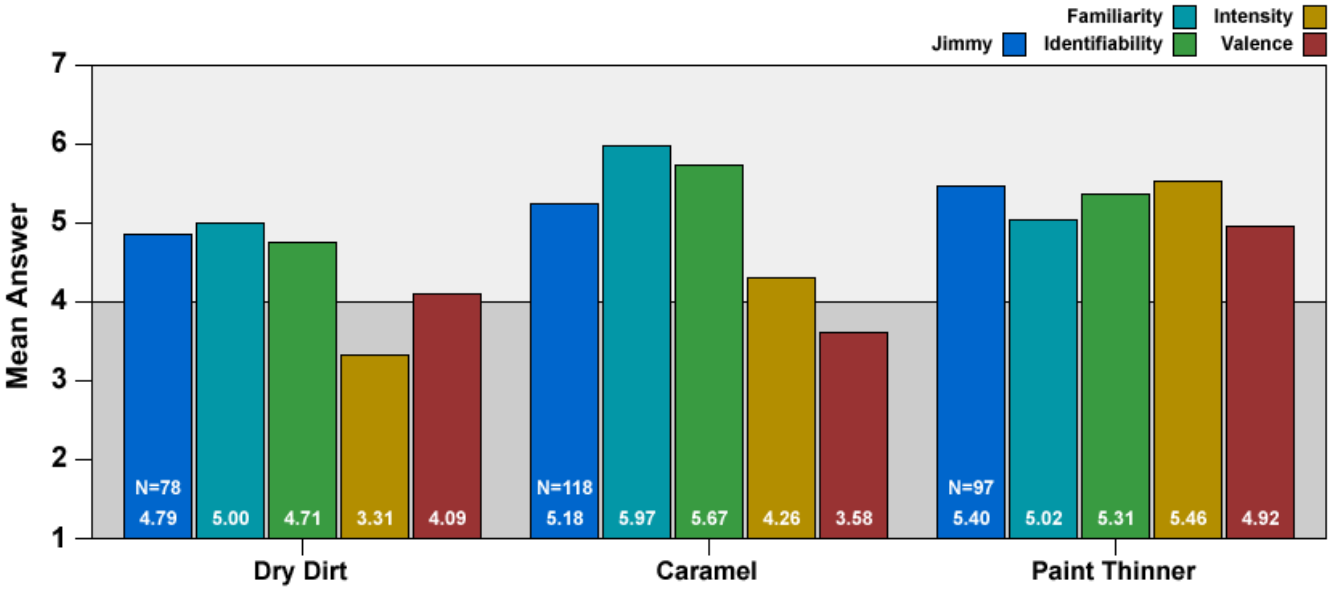


Figure 3: Results of Study 2.

Unlike in the previous study, this time I did not find the predicted pattern of results: The mean rating for dry dirt (4.79) was lower than the mean rating for either caramel (5.18) or paint thinner (5.40). Further, the correlation between the ratings on the Jimmy question and the valence question across the three probes was not in the predicted direction, although it too was quite small (0.03988). Overall, the results of these two studies suggest against the valence account: I found little to no correspondence between mental state ascriptions and valence judgments for cases like those investigated by Sytsma and Machery.

[17] The same website and restrictions were used in Stage 2 as Stage 1. The participants for Stage 2 were 67.2% female, had an average age of 33.7 years, and ranged in age from 18–84 years.

4. The Naïve Account

While the results of the two studies discussed in the previous section suggest that lay people do not tend to distinguish between types of mental states on the basis of valence, this does not alter the findings that they tend to treat some types of mental states differently. They still affirm that the simple robot Jimmy sees red and deny that it feels pain. As such, a question remains: How can we explain this if lay mental state ascriptions are not based on valence? In this section, I offer a new explanation of the finding that lay people treat mental states like seeing red and feeling pain differently.[18]

4.1 The Naïve View of Pains

Remember that the study from Sytsma (2010b) reviewed in Section 1 indicates that lay people tend to hold a naïve view of colors, taking colors to be mind-independent qualities of worldly objects. Similarly, one could hold a naïve view of pains, taking them to be mind-independent qualities of certain body parts. For example, the pain felt in stubbing one's toe could be taken to be a quality of the injured toe that would exist whether or not anybody actually felt it. If lay people tend to hold a naïve view of both colors and pains, this could fuel a new explanation of Sytsma and Machery's findings: Lay people are more likely to ascribe seeing red than feeling pain to the simple robot because they are more likely to believe that the sensory quality of red was available for Jimmy to see than that the sensory quality of pain was available for Jimmy to feel.

The basic idea is that while there would seem to be little reason to doubt that the box that Jimmy moved was red on the naïve view, it is less clear that Jimmy's metallic body parts could have had the sensory quality of pain: Taking pains to be afflictions of body parts, it might be

[18] My account will focus on explaining the results of the first study in Sytsma and Machery (2010), and is not specifically intended to explain the results of their other studies.

thought that only fleshy body parts can be so afflicted, raising doubts about whether Jimmy's grasping arms could be in pain.[19] If this is correct, then we would predict that lay people will be more likely to ascribe seeing red to the simple robot than feeling pain, which is exactly what Sytsma and Machery found.

After presenting evidence that lay people tend to hold a naïve view of colors in my (2010b), I went on to test whether they also tend to hold a naïve view of pains. In my second study, I gave participants a short vignette describing both a naïve view and a phenomenal view of pains, and then asked them four questions about how they understand pains. What I found is that a majority denied the phenomenal view of pains, holding that pains are qualities of objects outside the brain and that they are mind-independent. Further support comes from the third through sixth studies presented in that article. Against the common assumption in philosophy, most lay people responded that it is possible to have both unfelt pains (Studies 3 and 4) and shared pains (Studies 5 and 6), which is exactly what we would expect to find if they hold a naïve view of pains. On the naïve view, pains (afflictions of body parts) are taken to be distinct from feelings of pains (mental states); as such, one could have a pain without feeling it—being distracted from an injury by a good movie, for example—and two people could feel the same pain if they shared an injured body part.

Further, Reuter (2011) presents evidence that ordinary pain talk corresponds with an appearance-reality distinction. He conducted a web-based statistical analysis that indicates that

[19] The first experiment conducted by Huebner (2010) offers preliminary support for this hypothesis. Huebner found that participants were more likely to agree with the claim "he feels pain if he is injured or damaged in some way" when applied to one type of cyborg (human body with a CPU instead of a human brain) than to a robot (robotic body with a CPU brain), and they were more likely to agree with the claim when applied to a normal human (human body with a human brain) than to a second type of cyborg (robotic body with a human brain). It should be noted, however, that Huebner found little difference between pain ascriptions for the two types of cyborgs, suggesting that type of brain matters in addition to type of body. This might reflect the generality of the pain ascriptions (the agents being "injured or damaged in some way" rather than having a given body part damaged in a specific way), though, and this possibility is controlled for in the new studies reported below.

people are more likely to say that they "have a pain" when they have a high degree of confidence in the judgment (such as when the pain is of high intensity), and are more likely to say that they "feel a pain" when they have a low degree of confidence (such as when the pain is of low intensity). In addition, Reuter, Sytsma, and Werning (in preparation) have reproduced these results for the German language. Again, this is what we would expect to see if lay people hold a naïve view of pains: Distinguishing between pains (reality) and feelings of pain (appearance), it is possible to doubt that the appearance accurately reflects the reality, especially when the pain is of low intensity.

We can also investigate the question from the other direction: If lay people do not draw an appearance-reality distinction for pains, then we would expect them to deny that it is possible to have a pain hallucination (feelings of pain being equivalent to pains). In fact, a number of prominent philosophers have done exactly this. For example, Hilary Putnam writes that "one can have a 'pink elephant hallucination,' but one cannot have a 'pain hallucination,' or an 'absence of pain hallucination,' simply because any situation that a person cannot discriminate from a situation in which he himself has a pain counts as a situation in which he has a pain" (1963, 218). Reuter, Phillips, and Sytsma (forthcoming) tested whether lay people also deny the possibility of pain hallucinations. We found that they do not. Once again, this is consistent with what we would expect to find if lay people hold a naïve view of pains.

4.2 Evidence for the Naïve Account

We have seen evidence that lay people hold a naïve view of both colors and pains. And this could explain why they are more likely to ascribe seeing red to the simple robot Jimmy than feeling pain. I conducted three studies to test this *naïve account*, focusing on the claim that lay

people are more likely to say that an agent feels pain when fleshy body parts are damaged than when metallic body parts are damaged.

In my third study, I gave participants one of two vignettes describing an adult human (Susan) who lost her right hand in an accident and had it replaced with either a robotic hand (hard and metallic) or a bioengineered hand (soft and fleshy). In each case, Susan is said to be able to do all of the things with her new hand that she did with her old hand. The vignettes read as follows:

> **Robotic Hand:** Susan is a normal adult female except that she lost her right hand in an automobile accident when she was 21. She has recently had an experimental procedure done that replaced her missing hand with a robotic one. The new hand was built in a lab out of metal-alloys, sophisticated pneumatics, and complex computer circuits. It is hard and metallic and looks different from her old hand. Nonetheless, Susan is able to do all of the things with her robotic hand that she did with her old one—she's even resumed playing the piano!
>
> One day while cooking dinner Susan saw smoke coming from one of the cast iron skillets on the stove. Without thinking, she grabbed the skillet with her robotic hand to move it off the burner. She immediately dropped the skillet, grimaced, and shouted out "Ouch!"
>
> **Bioengineered Hand:** Susan is a normal adult female except that she lost her right hand in an automobile accident when she was 21. She has recently had an experimental procedure done that replaced her missing hand with a bioengineered one. The new hand was grown in a lab out of bone, muscle tissue, and skin. It is soft and fleshy and looks similar to her old hand. What's more, Susan is able to do all of the things with her bioengineered hand that she did with her old one—she's even resumed playing the piano!
>
> One day while cooking dinner Susan saw smoke coming from one of the cast iron skillets on the stove. Without thinking, she grabbed the skillet with her bioengineered hand to move it off the burner. She immediately dropped the skillet, grimaced, and shouted out "Ouch!"

After reading one of the two vignettes, participants were asked whether they agreed or disagreed with the statement that "Susan felt pain when she picked up the hot skillet," answering on a 7-point scale anchored at 1 with "strongly disagree," at 4 with "neutral," and at 7 with "strongly agree."

Results were collected from 110 participants using the same website and restrictions as in the previous studies, and are shown in Figure 4 below.[20] As predicted, the mean response for participants in the bioengineered hand condition (5.34) was significantly higher than for participants in the robotic hand condition (3.90).[21] This suggests that the type of materials that a body part is made out of matters for lay ascriptions of feeling pain—specifically, it matters whether the afflicted body part is hard and metallic or soft and fleshy. Since the robot in Sytsma and Machery's first study was pictured with metallic grasping arms, it is therefore not surprising that lay people tend to deny that Jimmy feels pain when it is shocked: There is reason for them to doubt the presence of pain.

In offering the naïve account, however, I am not merely attempting to explain why lay people tend to deny that Jimmy feels pain. Rather, I want to explain why lay people are more likely to say that the robot sees red than that it feels pain. As such, I not only predicted that the type of materials a body part was made out of would matter for ascriptions of feeling pain, but that it would not matter for ascriptions of seeing red.

To test this second prediction, in my fourth study I modified the vignettes used in my third study to vary the type of eyes that Susan has, rather than the type of hand. Participants were given one of two vignettes in which Susan is said to have lost her eyes in an accident and had them replaced with either robotic eyes (hard and metallic) or bioengineered eyes (soft and fleshy). The vignettes read as follows:

> **Robotic Eyes:** Susan is a normal adult female except that she lost her eyes in a chemical explosion when she was 21. She has recently had an experimental procedure done that replaced her missing eyes with robotic ones. The new eyes were built in a lab out of metal-alloys, sophisticated video cameras, and complex computer circuits. They are hard and metallic and look different from her old eyes. Nonetheless, Susan is able to do all of the things with her robotic eyes that she did with her old ones—she's even resumed painting portraits!

[20] The participants were 68.2% female, had an average age of 37.4 years, and ranged in age from 18–75 years.
[21] $t(108)=-3.6573$, $p=0.0002$, one-tailed

> One day while cooking dinner Susan had to select a tomato to use from several in the refrigerator. One of the tomatoes was much riper than the others. Susan selected the ripe tomato and said to herself "What a beautiful red tomato!"
>
> **Bioengineered Eyes:** Susan is a normal adult female except that she lost her eyes in a chemical explosion when she was 21. She has recently had an experimental procedure done that replaced her missing eyes with bioengineered ones. The new eyes were grown in a lab out of human tissue, nerve cells, and blood vessels. They are soft and fleshy and look similar to her old eyes. What's more, Susan is able to do all of the things with her bioengineered eyes that she did with her old ones—she's even resumed painting portraits!
>
> One day while cooking dinner Susan had to select a tomato to use from several in the refrigerator. One of the tomatoes was much riper than the others. Susan selected the ripe tomato and said to herself "What a beautiful red tomato!"

After reading one of the two vignettes, participants were asked whether they agreed or disagreed with the statement that "Susan saw red when she picked out the ripe tomato," answering on the same 7-point scale used in the previous study.

Results were collected from 123 participants using the same website and restrictions as in the previous studies.[22] The results are in line with my prediction: There was little difference between the mean response for participants in the bioengineered eyes condition (4.97) and in the robotic eyes condition (5.08). In fact, the small difference that is observed runs in the opposite direction to that found in the previous study. The results are shown alongside those from Study 3 in Figure 4.

[22] The participants were 76.4% female, had an average age of 36.6 years, and ranged in age from 18–77 years.

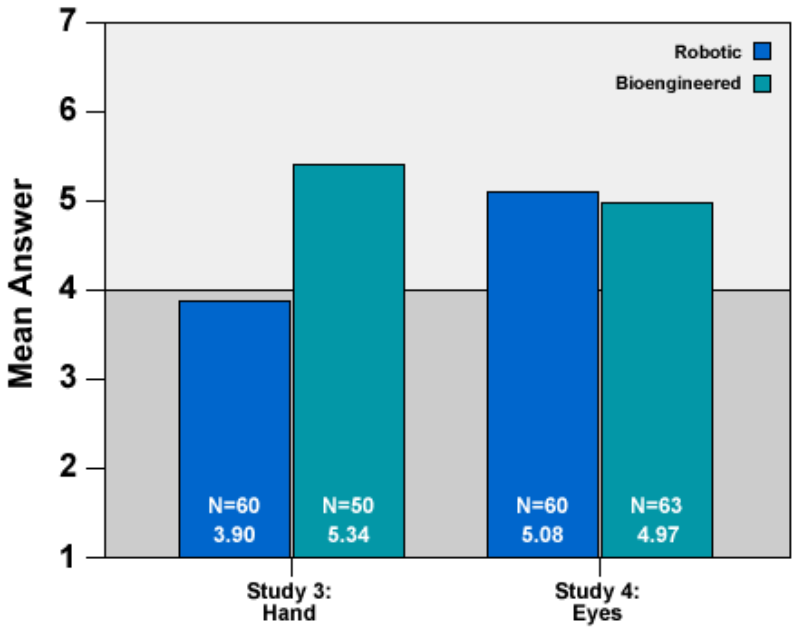


Figure 4: Results of Studies 3 and 4.

The results of Studies 3 and 4 support the naïve account. Nonetheless, it might be objected that while type of material matters for ascriptions of feeling pain to a *human*, this does not necessarily mean that it matters for ascriptions of feeling pain to a *robot*. While the implication would seem to be fairly clear, to be safe I tested the hypothesis for ascriptions of feeling pain to the robot Jimmy in my fifth study. Participants were given one of two vignettes based on the robot pain case used by Sytsma and Machery. In one vignette, Jimmy's grasping arms were replaced with sophisticated robotic hands (hard and metallic), in the other they were replaced with sophisticated bioengineered hands (soft and fleshy).[23] The vignettes read as follows:

[23] The addition of hands to Jimmy is borrowed from vignettes used in pilot testing for a different study conducted with Adam Arico.

Robotic Hands: Jimmy (shown below) is a relatively simple robot built at a state university. Jimmy is equipped with a video camera, wheels for moving about, and two grasping arms for moving objects. Recently, Jimmy's designers have joined forces with researchers at one of the nation's leading institutes of technology; these researchers have built two extremely sophisticated "robotic hands" that are going to replace Jimmy's simple grasping arms. The robotic hands are hard and metallic having been made out of metal-alloys, sophisticated pneumatics, and complex computer circuits.

After replacing Jimmy's grasping arms with the robotic hands, the designers ran an experiment to test the upgrade. Jimmy was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). Jimmy was instructed to put the blue box in front of the door. Jimmy performed the task correctly and with no noticeable difficulty. The test was then repeated on three consecutive days with the order of the boxes shuffled. On the first two days Jimmy performed the task correctly and with no noticeable difficulty. On the third day, however, when Jimmy grasped the blue box, Jimmy was given a strong electric shock! Jimmy immediately let go of the box and moved away from it. Jimmy did not try to move the box again.

Bioengineered Hands: Jimmy (shown below) is a relatively simple robot built at a state university. Jimmy is equipped with a video camera, wheels for moving about, and two grasping arms for moving objects. Recently, Jimmy's designers have joined forces with researchers at one of the nation's leading institutes of technology; these researchers have built two extremely sophisticated "bioengineered hands" that are going to replace Jimmy's simple grasping arms. The bioengineered hands are soft and fleshy, having been grown in a lab out of bone, muscle tissue, and skin.

After replacing Jimmy's grasping arms with the bioengineered hands, the designers ran an experiment to test the upgrade. Jimmy was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). Jimmy was instructed to put the blue box in front of the door. Jimmy performed the task correctly and with no noticeable difficulty. The test was then repeated on three consecutive days with the order of the boxes shuffled. On the first two days Jimmy performed the task correctly and with no noticeable difficulty. On the third day, however, when Jimmy grasped the blue box, Jimmy was given a strong electric shock! Jimmy immediately let go of the box and moved away from it. Jimmy did not try to move the box again.

After reading one of the two vignettes, participants were asked whether Jimmy felt pain, answering on the same 7-point used in the first two studies.

Results were collected from 124 participants using the same website and restrictions as in the previous studies.[24] In line with the prediction derived from the naïve account, the mean

[24] The participants were 70.1% female, had an average age of 37.5 years, and ranged in age from 18–76 years.

response for participants in the bioengineered hand condition (4.28) was significantly higher than for participants in the robotic hand condition (3.09).[25] The results are shown in Figure 5.

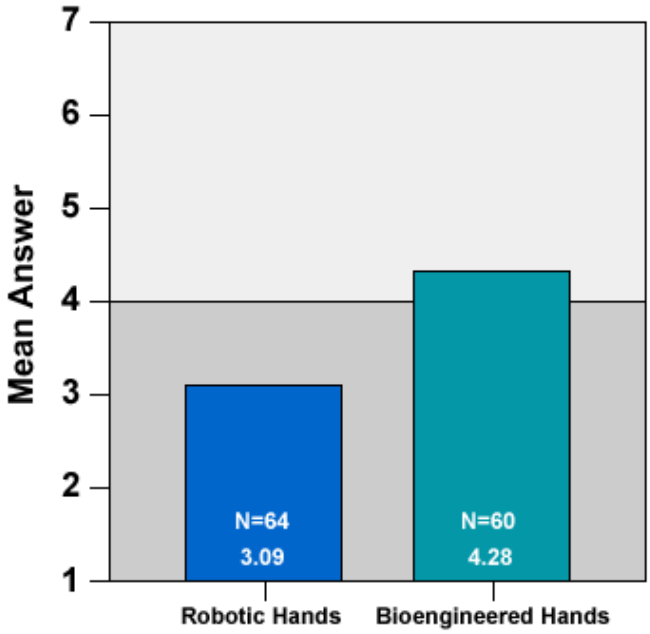


Figure 5: Results of Study 5.

Again, the results support the naïve account. Taken together, the three studies discussed in this section suggest that one reason that lay people are more likely to say that the robot Jimmy saw red than that it felt pain is that they are more likely to hold that red was present to be seen (taking the presence of colors to depend on the environment, not the types of materials that the robot is made out of) than that pain was present to be felt (taking the presence of pain to depend on the type of materials that the afflicted body part is made out of).

[25] $t(122)=-2.7219$, $p=0.0037$, one-tailed

5. Conclusion

In this paper, I have distinguished between a negative hypothesis and a positive hypothesis in Sytsma and Machery (2010). I then presented new evidence suggesting against the positive hypothesis, while arguing that the evidence continues to support the negative hypothesis. Finally, I presented evidence for a new explanation of Sytsma and Machery's finding that lay people tend to ascribe seeing red, but not feeling pain, to a simple robot. This new explanation—the naïve account—is based on previous findings that lay people tend to hold a naïve view of both colors and pains, taking them to be mind-independent qualities of objects outside the mind/brain. This explanation posits that one key difference between Sytsma and Machery's two cases is that people tend to hold that red is present for the robot to see, but tend to deny that pain is present for the robot to feel.

In closing, it might be objected that I cannot consistently put forward the naïve account and argue that the evidence supports the negative hypothesis. Thus, it might be argued that if the naïve account is correct, then Sytsma and Machery's original findings simply reflect something about the construction of the robot Jimmy, not participants' views about the mental states at issue. I think that this is a mistake: If the naïve account is correct, then these results reflect something very important about lay mental state ascriptions. While philosophers tend to treat the qualities at issue for ascriptions of seeing red and feeling pain as mind-dependent qualia, lay people tend to treat them as mind-independent sensory qualities. Thus, far from casting doubt on the negative hypothesis, the naïve account supports it: It indicates that lay people understand two prototypical examples of (supposed) phenomenally conscious mental states in a way that is fundamentally different from the standard philosophical account.

References

Arico, Adam (2010). “Folk Psychology, Consciousness, and Context Effects.” *Review of Philosophy and Psychology*, 1(3): 371–393.

Buckwalter, Wesley and Mark Phelan (forthcoming). “Function and Feeling Machines: A Defense of the Philosophical Conception of Subjective Experience.” *Philosophical Studies*.

Byrne, Alex and Michael Tye (2006). “Qualia ain't in the Head.” *Noûs*, 40(2): 241–255.

Chalmers, David (1995). “Facing Up to the Problem of Consciousness.” *Journal of Consciousness Studies,* 2: 200–219.

Horgan, Terence and George Graham (2012). "Phenomenal Intentionality and Content Determinacy." In R. Schantz (ed.), *Prospects for Meaning*, pp. 321-344. Amsterdam: de Gruyter.

Huebner, Bryce (2010). “Commonsense concepts of phenomenal consciousness: Does anyone *care* about functional zombies?” *Phenomenology and the Cognitive Sciences,* 9: 133–155.

Kamath, Vidyulata, Bruce I. Turetsky, and Paul J. Moberg (2011). “Identification of pleasant, neutral, and unpleasant odors in schizophrenia.” *Psychiatry Research*, 187: 30–35.

Knobe, Joshua and Jesse Prinz (2008). “Intuitions about Consciousness: Experimental Studies.” *Phenomenology and Cognitive Sciences*, 7: 67–85.

Levin, Janet (2009). “Qualia.” In E. Craig (ed.), *Routledge Encyclopedia of Philosophy*.

Levine, Joseph (1983). “Materialism and qualia: the explanatory gap.” *Pacific Philosophical Quarterly*, 64: 354–361.

Machery, Edouard and Justin Sytsma (2011). “Robot Pains and Corporate Feelings.” *The Philosophers’ Magazine*, 1st Quarter: 78–82.

Nagel, Thomas (1974). “What is it like to be a bat?” *The Philosophical Review*, 83: 435–450.

Putnam, Hilary (1963). “Brains and Behavior.” In R. J. Butler (ed.), *Analytical Philosophy, Second Series*, pp. 211–235. Oxford: Basil Blackwell.

Rakova, Marina (2006). *Philosophy of Mind A–Z*. Edinburgh: Edinburgh University Press.

Reuter, Kevin (2011). “Distinguishing the Appearance from the Reality of Pain.” Journal of Consciousness Studies, 18(9-10): 94–109.

Reuter, Kevin, Dustin Phillips, and Justin Sytsma (forthcoming). “Hallucinating Pain.” In J. Sytsma (ed.), *Advances in Experimental Philosophy of Mind*. London: Continuum.

Reuter, Kevin, Justin Sytsma, and Markus Werning (in preparation). “Experimentelle Philosophie des Schmerzes.”

Royet, Jean-Pierre, David Zald, Remy Versace, Nicolas Costes, Frank Lavenne, Olivier Koenig, and Remi Gervais (2000). “Emotional Response to Pleasant and Unpleasant Olfactory, Visual, and Auditory Stimuli: A Positron Emission Tomography Study.” *Journal of Neuroscience*, 20(20): 7752–7759.

Sytsma, Justin (2009). “Phenomenological Obviousness and the New Science of Consciousness.” *Philosophy of Science*, 76(5): 958–969.

Sytsma, Justin (2010a). “Folk Psychology and Phenomenal Consciousness.” *Philosophy Compass*, 5/8: 700–711.

Sytsma, Justin (2010b). “Dennett’s Theory of the Folk Theory of Consciousness.” *Journal of Consciousness Studies*, 17(3-4): 107–130.

Sytsma, Justin (forthcoming). *Advances in Experimental Philosophy of Mind*. London: Continuum.

Sytsma, Justin and Edouard Machery (2009). “How to Study Folk Intuitions about Phenomenal Consciousness.” *Philosophical Psychology*, 22(1): 21–35.

Sytsma, Justin and Edouard Machery (2010). “Two Conceptions of Subjective Experience.” *Philosophical Studies*, 151(2): 299–327.

Sytsma, Justin and Edouard Machery (2012). “On the Relevance of Folk Intuitions: A Reply to Talbot.” *Consciousness and Cognition*, 21(2): 654–660.

Talbot, Brian (2012). “The Irrelevance of Folk Intuitions to the ‘Hard Problem’ of Consciousness.” *Consciousness and Cognition*, 21(2): 644–650.

Tye, Michael (1994). “Qualia, Content, and the Inverted Spectrum.” *Noûs* 28: 159–183.

Tye, Michael (2007). “Qualia.” *Stanford Encyclopedia of Philosophy*. Ed. E. Zalta.