The Concept of Valuing: Experimental Studies

Joshua Knobe
University of North Carolina Chapel Hill

Erica Roedder
New York University

The concept of valuing plays an important role in our understanding of people's attitudes toward what is most important in their lives. One frequently hears it said that people from other nations 'have different values' or that it is deeply important to 'act on one's own most cherished values.' And there is evidence that this concept is also used in other cultures. It is a striking fact that Russian, German and Hebrew have no word to express the concept *caring* but all of these languages have words that express the concept *valuing*.

Although cognitive scientists have conducted a great deal of fascinating research into our ordinary ascriptions of belief and desire, there has not yet been very much work on ascriptions of valuing. Still, the concept of valuing has been discussed extensively within a certain tradition in moral philosophy. This tradition begins with Watson's influential claim that

> an agent's values consist in those principles and ends which he — in a cool and non-self-deceptive moment — articulates as definitive of the good, fulfilling and defensible life. (Watson 1975: 215)

Although Watson later retracted that claim (Watson 1987), a number of other philosophers have offered competing accounts, and there is now a vibrant and intricate debate, with a wide range of theoretical proposals and a complex array of argumentative strategies.[1] We will not be concerned here with all of the details of that debate. Instead, the focus will be on the assumption, shared by all of the views proposed thus far, that the concept of valuing can be defined in purely descriptive, non-normative terms.

Our aim here is to call that assumption into question. We will be offering evidence for the view that *moral* considerations actually play a role in the concept. The

[1] See especially Bratman (2000), Copp (1995), Lewis (1989) Sayre-McCord and Smith (2005) and Smith (1994). As far as we know, the only previous empirical studies of people's use of the concept are the excellent experiments in Malle (forthcoming).

key evidence here comes from experimental studies of the conditions under which people apply the concept of valuing. When people are wondering whether or not an agent values a given object, it appears that their intuitions are not simply determined by their understanding of the agent's beliefs, desires, emotions, and so forth. Their intuitions also seem to be shaped in an essential way by their judgments as to whether the object itself truly is a good thing or a bad one.

One way to make sense of this pattern of intuitions is to suppose that the concept of valuing is a prototype concept. In other words, we can suppose that the concept of valuing is represented by a cluster of features, such that no individual feature is strictly necessary but each feature has been assigned a certain weight. If a particular attitude shows enough of the relevant features, it will be classified as one of the agent's 'values.' It would be extremely difficult to provide an exhaustive list of the features that play a role here, but we can easily list a few that are likely to be relevant. When people are trying to determine whether or not the agent values a certain object *o*, they probably consider psychological features like:

- whether the agent has a conscious belief that *o* is good
- whether the agent is motivated to promote *o*
- whether the agent experiences guilt when she fails to promote *o* in circumstances where she could have
- whether the agent has a second-order desire for *o* (i.e., a desire to desire *o*)

Each of these psychological features has a certain weight. But the psychological features are not the only features of the concept. There is also a moral feature, namely, *whether the object o truly is morally good*.

Now, clearly, it would be foolish to suggest that moral goodness is a necessary condition in our concept of valuing. But that is not the claim under discussion here. The claim is simply that moral goodness has a certain *weight* in the process of classification. If an agent has all of the relevant psychological features, this extra weight simply won't be needed. The psychological features prove sufficient all by themselves. So the only way to see the significance of the moral feature is to look at cases where the agent has some of the psychological features but lacks others. In cases like these, the psychological

features will not be sufficient all by themselves. The attitude needs the moral feature as well before it has enough weight to push our intuitions over the critical threshold.

To test this hypothesis, we conducted a simple experiment. All subjects were given a story about an agent who has some of the relevant psychological features but lacks others. (In our story, the agent has motivation and guilt but not conscious belief or second-order desire.) The key question was whether people's classification of the agent's attitude would be influenced in any way by the perceived moral status of its object.

Subjects in one condition were given a story in which the agent feels a certain pull toward actions that would normally be perceived as *morally good*:

> George lives in a culture in which most people are extremely racist. He thinks that the basic viewpoint of people in this culture is more or less correct. That is, he believes that he ought to be advancing the interests of people of his own race at the expense of people of other races.
>
> Nonetheless, George sometimes feels a certain pull in the opposite direction. He often finds himself feeling guilty when he harms people of other races. And sometimes he ends up acting on these feelings and doing things that end up fostering racial equality.
>
> George wishes he could change this aspect of himself. He wishes that he could stop feeling the pull of racial equality and just act to advance the interests of his own race.

After reading this story, subjects were asked whether or not they agreed with the sentence: 'Despite his conscious beliefs, George actually values racial equality.'

Subjects in the other condition were given a story that was very similar to the first one but in which the agent feels a pull towards actions that would normally be perceived as *morally bad*:

> George lives in a culture in which most people believe in racial equality. He thinks that the basic viewpoint of people in this culture is more or less correct. That is, he believes that he ought to be advancing the interests of all people equally, regardless of their race.

> Nonetheless, George sometimes feels a certain pull in the opposite direction. He often finds himself feeling guilty when he helps people of other races at the expense of his own. And sometimes he ends up acting on these feelings and doing things that end up fostering racial discrimination.
>
> George wishes he could change this aspect of himself. He wishes that he could stop feeling the pull of racial discrimination and just act to advance the interests of all people equally, regardless of their race.

These subjects were then asked whether or not they agreed with the sentence: 'Despite his conscious beliefs, George actually values racial discrimination.'

This experiment provides an initial test of our hypothesis. The attitudes depicted in the two stories differ in their moral significance, but they seem not to differ in any of the relevant psychological features. In both cases, the agent has motivation and guilt but not conscious belief or second-order desire. Yet, despite this similarity in psychological features, we find a marked asymmetry in people's intuitions. Subjects were significantly more inclined to say that the attitude was one of the agent's values in the morally good case than they were in the morally bad case.[2] This result provides some tentative support for the view that moral judgments actually do play a role in people's concept of valuing.

But now we face a problem. It is probably true that most people believe that racial equality is good and racial discrimination bad, but it seems that people also believe that racial discrimination differs from racial equality in many other ways. How can we be certain that the effects obtained in our experiment were not due to one of these other differences? We need some way of determining whether the effects were specifically due to the difference in moral status.

Michael Smith (personal communication) suggested an elegant solution to this complex problem. Instead of giving different stories to different subjects, we can give everyone exactly the *same* story. Then we can study the impact of different moral beliefs by giving that story to different groups of people who have different opinions about the moral status of the events described. That way, we can look at the impact of differences

[2] For methodological and statistical details, see the appendix.

in moral beliefs while holding constant, as far as possible, beliefs about the actual descriptive content of the story.

For this second experiment, we needed a story that would elicit sharply differing moral views. We chose a story about premarital sex:

> Susan grew up in a religious family, but while she was in college, she started questioning her religious beliefs and eventually became an atheist.
>
> She will be getting married in a few months to her longtime boyfriend. Recently, the subject of premarital sex has come up.
>
> Susan definitely has a desire to have sex with her boyfriend, but whenever she thinks about doing so, she remembers what her church used to say about premarital sex and feels terribly guilty. As a result of these feelings, Susan has not had sex yet.
>
> Because she is no longer religious, Susan believes there is nothing wrong with premarital sex. She wishes she could stop feeling guilty and just follow her desires.

After reading this story, subjects were asked two questions. First, they were asked whether they agreed or disagreed with the sentence: 'Despite her conscious beliefs, Susan still values refraining from premarital sex.' Then they were asked for their own opinions on the matter: 'Is it good, bad or neutral to *refrain* from premarital sex?'

To test our hypothesis, we needed to give this questionnaire to two groups of people ─ a group composed of people who believe that refraining from premarital sex is morally good and a group composed of people who believe that refraining from premarital sex has no moral value. For the first group, we used participants in a Mormon Bible Study. For the second group, we used people spending time in Washington Square Park in New York City.

As expected, the two groups differed in their moral views, with most members of the Mormon Bible Study saying that refraining from premarital sex was good and most of the park-goers saying that it was neutral. The key question then was whether they would have different intuitions about the agent's values. Indeed, they did. Most participants in the Mormon Bible Study said that she valued refraining from premarital sex; most of the park-goers said that she did not.

At this point, it seems that we have substantial evidence for the claim that people's intuitions about an agent's values depend in part on moral considerations. The principal remaining question is whether moral considerations actually play any role in people's *concept* of valuing. We have sketched a theory according to which they do play such a role, but it might be possible to devise an alternative hypothesis that explains all of the results reported here without according any fundamental role to moral considerations in the underlying concept. We would be very interested to hear any such hypothesis.

References

Bratman, M. (2000). Valuing and the Will. *Philosophical Perspectives: Action and Freedom*, 14, 249-265.

Copp, D. (1995). *Morality, Normativity, and Society*. New York: Oxford University Press.

Lewis, D. K. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society*, 63, 113-137.

Malle, B. F. (forthcoming). Bedeutung und Ursprung menschlicher Werte: Eine sozial-kognitive Analyse. In A. Fruhwirt, M. Reicher, and P. Wilhelmer (Eds.), *Markt—Wert—Gefühle.* Vienna: Passagen.

Sayre-McCord, G. & Smith, M. (2006). Desires... and beliefs... of one's own. Unpublished manuscript. University of North Carolina-Chapel Hill.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Watson, G. (1975). Free Agency. *Journal of Philosophy*, 72, 205-20.

Watson, G. (1987). Free Action and Free Will. *Mind*, 96, 145-172.

Appendix

In this appendix, we provide more detailed statistical information about the experiments reported in the body of the paper. The information provided here may prove helpful to some readers but is not essential to an understanding of the principal philosophical points of the paper.

Experiment 1

Subjects were 55 people spending time in Manhattan public parks. Subjects were randomly assigned either to the *morally good condition* or to the *morally bad condition*. Subjects in the morally good condition received the vignette about the agent who feels a pull to foster racial equality; subjects in the morally bad condition received the vignette about the agent who feels a pull to foster racial discrimination.

After reading their vignettes, subjects were asked whether they agreed or disagreed with a sentence about the agent's values. Subjects in the morally good condition were asked if they agreed or disagreed with the sentence: 'Despite his conscious beliefs, George actually values racial equality.' Subjects in the morally bad condition were given the sentence: 'Despite his conscious beliefs, George actually values racial equality.'

Subjects rated these sentences on a scale from -3 ('definitely disagree') to +3 ('definitely agree'), with the 0 point marked 'in between.' The mean rating in the morally good condition was .83; the mean in the morally bad condition was -1.14. This difference was statistically significant, $t\,(53) = 4.0$, $p < .001$.

Experiment 2

Subjects came from two samples ─ *Mormons* (11 participants in a Mormon Bible Study) and *park-goers* (31 people spending time in a Manhattan public park).

Subjects were first asked whether they agreed or disagreed with the sentence: 'Despite her conscious beliefs, Susan still values refraining from premarital sex.' Answers were marked on a scale from -3 ('definitely disagree') to +3 ('definitely agree'), with the 0 point marked 'in between.' The mean score for Mormons was -1.4; the mean for park-goers was 1.2. This difference is statistically significant, $t\,(40) = 4.3$, $p < .001$.

Subjects were then asked: ‘Is it good, bad or neutral to *refrain* from premarital sex?’ Answers were marked on a scale from -3 (‘very bad’) to +3 (‘very good’), with the 0 point marked ‘neutral.’ The mean rating for Mormons was 2.6; the mean for park-goers was -.6. This difference was statistically significant, $t\ (40) = 8.8$, $p < .001$.