Love and Happiness

Jonathan Phillips[1], Luke Misenheimer[2] and Joshua Knobe[1,3]

[1]*Program in Cognitive Science, Yale University;* [2]*Department of Philosophy, University of California, Berkeley;* [3]*Department of Philosophy, Yale University*

Consider people's ordinary concept of *belief*. This concept seems to pick out a particular psychological state. Indeed, one natural view would be that the concept of belief works much like the concepts one finds in cognitive science – not quite as rigorous or precise, perhaps, but still the same basic type of notion.

But now suppose we turn to other concepts that people ordinarily use to understand the mind. Suppose we consider the concept *happiness*. Or the concept *love*. How are these concepts to be understood? One obvious hypothesis would be that they are best understood as being more or less like the concept of belief. Maybe these concepts, too, pick out a particular mental state and thereby enable people to predict, explain and understand others' behavior.

We will argue that this hypothesis is mistaken. Instead, we suggest that the different concepts people use to understand the mind are fundamentally different from each other. Some of these concepts do indeed serve simply to pick out a particular mental state, but others allow a role for *evaluative* judgments. So, for example, our claim will be that when people are wondering whether a given agent is truly 'happy' or 'in love,' they are not merely trying to figure out whether this agent has a particular sort of mental state. They are also concerned in a central way with evaluating the agent herself.

In short, our aim is to point to a striking sort of difference between the different concepts that people use to pick out psychological attitudes. We will be trying to show that evaluative judgments play a role in the concepts of love and happiness that they do not play in other concepts that might at first seem quite similar.

To examine these issues, we conducted three simple studies. Here we provide only a brief description of each, but a full description of all the methods and results is available in the supplementary materials (Phillips et al. 2010).

Study 1

Nyholm (2009) has shown that people's intuitions about whether an agent is truly happy can be affected by their views about whether this agent actually is leading a good life. We hypothesized that although this effect arises for the concept *happiness*, it would not arise for the concept *unhappiness*. In other words, we hypothesized that the concept of happiness allows a role for evaluative judgments but that the concept of unhappiness serves simply to pick out a particular mental state.

To test this hypothesis, we used a 2 x 2 design in which we independently varied the concept that participants were asked to apply (happiness vs. unhappiness) and the nature of the life they were asked to evaluate (good life vs. bad life). Each participant was therefore assigned to one of four possible conditions:

Happiness/Good Life. Maria is described as a caring individual with a great family life and a variety of meaningful friendships and projects. She enjoys her day-to-day activities and feels like there isn't anything she would rather be doing with her life. Participants are asked whether Maria is happy.

Happiness/Bad Life. Maria is described as a vapid individual who has no real friendships and no goals beyond going to parties and gaining greater social status. Nonetheless, she enjoys her day-to-day activities and feels like there isn't anything she would rather be doing with her life. Participants are asked whether Maria is happy.

Unhappiness/Good Life. Maria is described as a caring individual with a great family life and a variety of meaningful friendships and projects. Nonetheless, she feels terrible all the time and regards her life as fundamentally a failure. Participants are asked whether Maria is unhappy.

Unhappiness/Bad Life. Maria is described as a vapid individual who has no real friendships and no goals beyond going to parties and gaining greater social status. She feels terrible all the time and regards her life as fundamentally a failure. Participants are asked whether Maria is unhappy.

Participants rated the degree to which they agreed with the statement that Maria is happy [unhappy] on a scale from 1 to 7.

The mean ratings for each condition are displayed in Figure 1. The results showed the predicted difference between judgments of happiness and judgments of

unhappiness. In judgments of happiness, there was a highly significant difference between the good life and the bad life. By contrast, in judgments of unhappiness, there was nothing even approaching a significant effect.
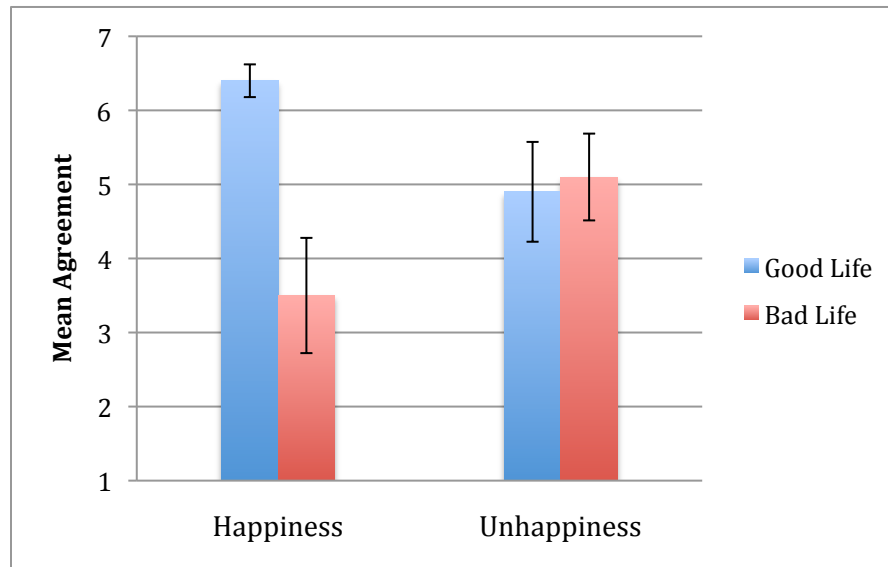

Figure 1. Attributions of happiness and unhappiness by condition. Error bars show SE mean.

In short, people's evaluations of Maria's life appeared to be having a significant impact on attributions of happiness but not on attributions of unhappiness. [1]

Study 2

We then hypothesized that precisely the same effect would arise for the contrast between *love* and *lust*. That is, we hypothesized that attributions of love involve a kind of evaluation, whereas attributions of lust serve simply to pick out a particular mental state.

---

[1] One worry about this first study is that the asymmetry might have arisen simply because participants in the Happiness/Bad Life condition thought that Maria was merely deceiving herself into a superficial appearance of positive mental states. To address this concern, Liao and Phillips (unpublished data) conducted a follow-up study in which all participants were given a vignette about a person who had negative mental states and the only difference between conditions lay in whether the question was about the degree to which this person was 'happy' vs. the degree to which the person was 'unhappy.' Even in this revised version, participants continued to show the asymmetry – tending to say that the person was more 'happy' in the good life condition than in the bad life condition, while saying that she was equally 'unhappy' in both conditions.

To test this hypothesis, we conducted a study in which each participant was assigned to one of two conditions:

Good Guy. Susan has a boyfriend who seems worthy of love (thoughtful and kind) and also worthy of lust (physically attractive).

Bad Guy. Susan has a boyfriend who seems unworthy of love (cruel and irresponsible) and also unworthy of lust (overweight and missing a few teeth).

In both conditions, participants were told that Susan has never been more attracted to another man and can't imagine spending her life with anyone else. Each participant was then asked (a) whether they agreed that Susan is experiencing love and (b) whether they agreed that Susan is experiencing lust. Answers to each of these questions were marked on a scale from 1 to 7.

The mean ratings for each question are displayed in Figure 2. Once again, we found the predicted difference between the two concepts. Participants were significantly more likely to say that Susan was experiencing love for the good guy than for the bad guy, whereas there was no significant difference between the two in attributions of lust.
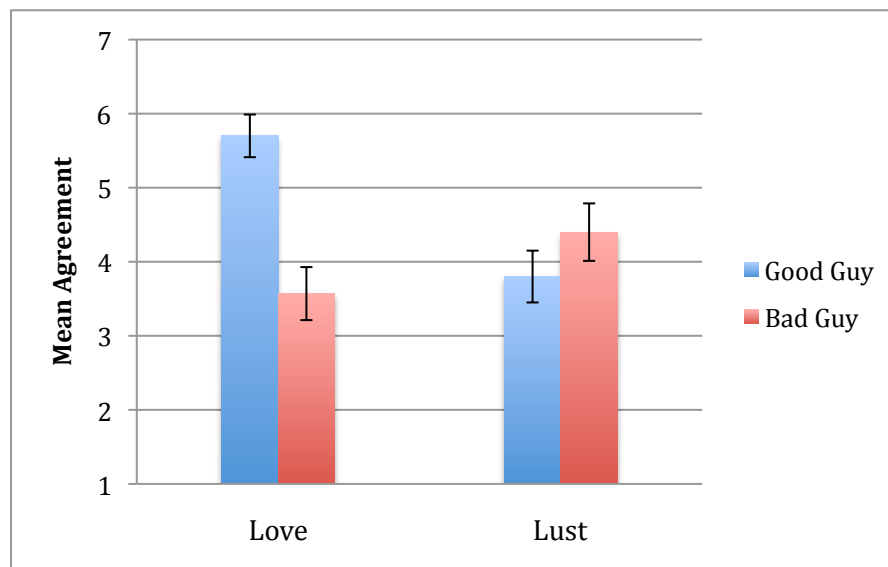

Figure 2. Attributions of love and lust by condition. Error bars show SE mean.

Study 3

Finally, we compared the concept of *valuing* to the concept of *thinking good*. The hypothesis was that intuitions about sentences like 'She values that' would be influenced by evaluative judgments (Knobe & Preston-Roedder 2009), whereas intuitions about sentences like 'She thinks that is good' would serve simply to pick out a particular mental state.

To test this hypothesis, we conducted an archival study using naturally occurring data. We began by using Google to identify passages that contained expressions that were either of the form 'He values *x*' or 'She values *x*' or of the form 'He thinks *x* is good' or 'She thinks *x* is good.' Two raters then coded these passages as to whether the writer of the passage appeared to believe that the object of the attitude actually was good. So, for example, suppose that the passage said:

Like every other reasonable and fair-minded person, he values prayer.

The raters would then conclude that the writer probably thought that prayer actually *was* good. By contrast, suppose that a passage said:

Like so many other victims of religious indoctrination, he thinks prayer is good.

The raters would then conclude that the writer probably thought that prayer *wasn't* good.

As predicted, evaluative judgments appeared to be having an impact on attributions of valuing but not on attributions of thinking good. The writer was coded as regarding the object as good in the vast majority of valuing passages (93%), whereas the percentage of thinking good passages in which the writer regarded the object as good did not differ significantly from chance (40%).

Conclusion

We have now reported three experiments examining three different pairs of concepts. In each case, we found the same basic pattern – with evaluative judgments showing an impact on the use of one concept but not showing an impact on the use of another, apparently similar concept. In particular, people's evaluative judgments

appear to be playing a role in attributions of *happiness*, *love* and *valuing*, but not in attributions of *unhappiness*, *lust* and *thinking good*.

 We are not at all sure why these different concepts differ in this way, but it does seem that there is some broader principle or generalization to be found here. Perhaps the matter can be addressed in future research.

References

Knobe, J. & Preston-Roedder, E. (2009). The ordinary concept of valuing. *Philosophical Issues*. 19: 131-147.

Liao, S. & Phillips, J. (unpublished data). [Further Asymmetries in Attributions of Happiness and Unhappiness], University of Michigan.

Nyholm, S. (2009). Moral judgments and happiness attributions. Unpublished manuscript, University of Michigan.

Phillips, J., Misenheimer, L. & Knobe, J. (2009). Love and Happiness: Supplementary Materials. Available at: http://pantheon.yale.edu/~jk762/LovHapSupp.pdf