True Happiness: The Role of Morality in the Folk Concept of Happiness

Jonathan Phillips[1*], Julian De Freitas[1], Christian Mott[2],

June Gruber[3] & Joshua Knobe[4,5]

[1] Harvard University, Department of Psychology

[2] Yale University, Law School

[3] University of Colorado Boulder, Department of Psychology and Neuroscience

[4] Yale University, Program in Cognitive Science

[5] Yale University, Department of Philosophy

* Corresponding author

Corresponding author address:

Jonathan Phillips

Harvard University

1482 William James Hall

33 Kirkland St.

Cambridge, MA 02138

phillips01@g.harvard.edu

Word count: 10,625
Figures: 7

Author note

Abstract

Recent scientific research has settled on a purely descriptive definition of happiness that is focused solely on agents' psychological states (high positive affect, low negative affect, high life satisfaction). In contrast to this understanding, recent research has suggested that the ordinary concept of happiness is also sensitive to the *moral* value of agents' lives. Five studies systematically investigate and explain the impact of morality on ordinary assessments of happiness. Study 1 demonstrates that moral judgments influence assessments of happiness not only for untrained participants, but also for academic researchers and even in those who study happiness specifically. Studies 2 and 3 then respectively ask whether this effect may be explained by general motivational biases or beliefs in a just world. In both cases, we find evidence against these explanations. Study 4 shows that the impact of moral judgments cannot be explained by changes in the perception of descriptive psychological states. Finally, Study 5 compares the impact of moral and non-moral value, and provides evidence that unlike non-moral value, moral value is part of the criteria that govern the ordinary concept of happiness. Taken together, these studies provide a specific explanation of how and why the ordinary concept of happiness deviates from the definition used by researchers studying happiness.

Keywords: happiness; morality; emotion; moral cognition; well-being

True Happiness: The Role of Morality in the Folk Concept of Happiness

One important aspect of people's ordinary understanding of the mind is their concept of *happiness*. This paper aims to investigate that ordinary concept. What criteria do people use when trying to determine whether or not someone counts as happy? In considering this question, one obvious place to start is with the accounts of happiness developed by researchers in psychology and philosophy. Within psychology, research on happiness has converged on a conception of happiness that is straightforwardly descriptive. That is, it suggests that whether or not someone is happy is simply a matter of the psychological states a person has (Diener, 2000; Diener, Scollon, & Lucas, 2004; Gruber, Mauss & Tamir, 2011; Lucas, Diener & Lawson, 2003; Zou, Schimmack & Gere, 2013). In contrast to this purely descriptive conception of happiness, philosophers have long argued that whether or not one is happy is at least partially an evaluative question. Put simply, the basic suggestion is that happiness is not only a matter of feeling good, but also a matter of being good (Aristotle, 340 BCE/2002; Foot, 2001; Kraut, 1979).

The studies reported here draw on these theoretical frameworks to explore questions about people's ordinary folk concept of happiness. In particular, the key question is whether (a) people's ordinary concept is purely descriptive or (b) people's ordinary concept actually involves an evaluative component.

Theories of Happiness

The past years have witnessed impressive advances in empirical research on happiness (e.g., Fredrickson, 2001; Gruber, Mauss & Tamir, 2011; Seligman & Csikszentmihalyi, 2000). One key advance within this field has been the convergence on a conception of happiness that involves three elements: *high positive affect*, *low negative affect*, and *high life satisfaction*

(Diener, 2000; Diener, Scollon, & Lucas, 2004; Lucas, Diener & Lawson, 2003; Zou, Schimmack & Gere, 2013). Focusing on this conception, researchers developed a number of methods for measuring positive and negative affective states, as well as general life satisfaction, which has yielded important insights about the descriptive nature of happiness (Csikszentmihalyi & Larson, 1987; Hektner, Schmidt & Csikszentmihalyi, 2007; Diener, Emmons, Larsen & Griffin, 1985; Pavot & Diener, 2008; Watson, Clark & Tellegen, 1988; Watson & Clark, 1994). Critically, this way of measuring and understanding happiness seeks to provide accurate descriptions of the psychological states that people experience, but does not focus or specifically assess moral evaluations (for reviews, see Diener, Scollon, & Lucas, 2004; Lyubomirsky, King & Diener, 2005). Thus, the conception of happiness employed by psychologists is thoroughly descriptive rather than evaluative.

By contrast, a number of philosophers have argued that evaluative considerations are directly relevant to happiness (Aristotle, 340 BCE/2002; Foot, 2001; Kraut, 1979). On this way of understanding happiness, one cannot decide whether or not someone is happy with only descriptive information about the psychological states that person experiences. Instead, whether or not she is actually happy depends in part on whether she is living a life that is *evaluatively good*. On this sort of view, for example, a Nazi commandant who thoroughly enjoys and is satisfied with fulfilling his work in a death camp, cannot actually be happy, no matter which descriptive psychological states he might experience (Foot, 2001, 90-96). Happiness requires one's life actually be good in addition to having the psychological states that make one feel good.

The Ordinary Concept of Happiness

Within the existing literature, it has often been suggested that people's ordinary concept

of happiness is purely descriptive (Cooper, 1987; Crisp, 2008; Feldman, 2010; Gilbert, 2006; Haybron, 2008; 2011; Sidgwick, 1907/1966). As Daniel Gilbert puts it: "Happiness is a word that we generally use to indicate an experience and not the actions that give rise to it. Does it make sense to say, 'After a day spent killing his parents, Frank was happy'? Indeed it does…" (Gilbert, 2006: 36-7).

A series of empirical studies suggest, however, that ordinary assessments of happiness are influenced by the moral value of the agent's life (e.g., Newman et al., 2015; Phillips, et al., 2011). In one study, for example, participants were told about a person who experiences a lot of positive affect, very little negative affect, and is highly satisfied with her life when she reflects on it. Participants were randomly assigned to learn either that this person was living a morally good life or that this person was living a morally bad life. Those who read about the person living a morally good life tended to agree that she was happy, while those who read about a person living a morally bad life tended to disagree (Phillips et al., 2011). Moreover, a number of additional studies suggested that individual differences in participants' moral judgments predicted their assessments of happiness (Phillips et al., 2014).

A question now arises as to what this effect is showing about people's ordinary concept of happiness. We will be focusing here on three basic families of hypotheses.

(1) *Bias or distortion.* One possible view would be that people's concept of happiness is itself purely descriptive and the impact of moral considerations arises solely from some kind of performance error. For example, it might be thought that the entire effect is explained by people getting confused and failing to correctly apply their own concept. Additionally, as this effect involves negative moral assessments, participants' judgments may also be explained by motivational biases such as those that have been well-documented in social psychology (e.g.,

people's intuitions might be distorted by a desire to believe in a 'just world'; Lerner, 1980).

Given the widespread presence of these kinds of motivational biases, it is highly likely that these

forms of motivated reasoning play some role in participants' assessments of happiness. The

essential question is whether these performance errors entirely account for the impact of morality

on assessments of happiness. If this turns out to be the case, people's underlying concept of

happiness could still be purely descriptive (see **Figure 1a**).

(2) *Mediation through descriptive psychological states.* A second possibility would be

that people's concept of happiness is purely descriptive and that the effect of moral

considerations arises solely because these considerations impact people's judgments about

whether or not the agent has the relevant descriptive properties (see **Figure 1b**). Thus, suppose

that people see happiness as basically just a matter of having certain specific psychological states

(high positive affect, low negative affect, judgments of life satisfaction). It might then turn out

that when people believe that an agent's life is morally bad, they tend to think of the agent as not

experiencing these states. This possibility is supported by previous research that demonstrated a

connection between happiness and morality by showing that participants tend to infer that people

are morally good when they report being happy and finding meaning in their jobs (King & Napa,

1998). As with the first hypothesis, previous research gives us reason to expect that such an

effect exists, and the critical question we take up is whether or not changes in the perception of

descriptive psychological states fully accounts for the impact of morality. If this turns out to be

the case, then people's moral judgments would have an impact on happiness attributions, but this

impact would be mediated by people's attribution of purely descriptive psychological states, and

thus the folk concept of happiness would be purely descriptive.

(3) *An evaluative concept of happiness*. Finally, it might be that people's ordinary

concept of happiness is itself fundamentally evaluative. That is, independent of anything about how people attribute descriptive psychological states, people might see moral considerations as directly relevant to the question as to whether an agent is truly happy (see **Figure 1c**). On such a view, people's moral judgments might impact their assessment of the agent's psychological states, but even if we completely hold fixed the assessment of psychological states, there should still be an impact of moral judgment on happiness attributions.

At the core of this last hypothesis is the idea that the impact of moral considerations is not to be explained in terms of a bias, or in terms of an impact on people's attributions of psychological states, but rather in terms of something about people's concept of happiness, i.e., their understanding of what it means for a person to truly be happy. If this hypothesis turns out to be correct, the next step would be to spell it out in more detail by drawing on work in the psychology of concepts. The present paper does not attempt to pursue that next step; we simply ask whether this effect is to be understood in terms of the concept of happiness at all. (For some initial thoughts about how to explain this effect in terms of the psychology of concepts, see the General Discussion.)

The evaluative concept hypothesis might at first seem highly counterintuitive or implausible. However, it should be noted that existing research has provided evidence for analogous hypotheses regarding numerous other concepts. For example, it may seem initially that moral considerations play no role in the concepts of intentional action, causation, freedom, or knowledge, but existing research provides evidence the moral considerations actually do play a role in all of those other concepts (Beebe & Buckwalter, 2010; Knobe, 2003; Kominsky, Phillips, Gerstenberg, Lagnado & Knobe, 2015; Leslie, Knobe & Cohen, 2006; Samland, Josephs, Waldmann & Rakoczy, 2016; Phillips, Luguri & Knobe, 2015; Young & Phillips,

2011). It is therefore at least worth considering the possibility that moral considerations might play a role in the concept of happiness as well.

The Present Investigation

The present investigation aims to decide between these three hypotheses, determining whether the impact of moral considerations is best explained in terms of bias or distortion, mediation through changes in assessments of descriptive psychological states, or the evaluative character of the concept itself.

Five studies consider these hypotheses in turn. Studies 1, 2 and 3 provide evidence that the effect of evaluative judgments on ordinary assessments of happiness is unlikely to be the result of a performance error or motivational bias. Studies 4a and 4b demonstrate that evaluative judgments selectively impact assessments of happiness, and do not influence similar assessments of emotion, suggesting that this effect is not mediated by changes in the perception of purely description psychological states. Study 5 then asks whether the concept of happiness is itself evaluative, and finds direct evidence that it is.

Study 1: The Ordinary Concept of Happiness, from Experts to the Folk

Previous research provided some evidence that assessments of how happy a person is are affected by whether that person is living a morally good life (Phillips et al. 2014). However, this research employed a between-subjects design and asked completely untrained participants to make assessments of whether or not an agent was happy based on only a single brief vignette. Accordingly, one may worry that these results do not actually reflect the ordinary concept of happiness but arose from a misunderstanding of the task, or reflected a pattern of responses that participants would not actually endorse upon reflection (Tiberius, 2013). To get a better sense for

the robustness of the effect of moral value on assessments of happiness, the current study differed from previous research in two key ways. First, we implemented a mixed within-subjects / between-subjects design, allowing us to ask whether participants would continue to say that those living morally good lives were happier than those living morally bad lives, even when it was made explicit that the only difference between them was a moral difference. Second, we recruited researchers in psychology and philosophy to participate alongside untrained participants. This allowed us to ask whether the effect of moral value on assessments of happiness differed across three different populations: (1) non-researchers, (2) scientific or philosophical researchers, and (3) experts who have studied happiness specifically.

Study 1 Methods

Participants

Previous research on the effect of morality on happiness found a Cohen's d of 1.603 (Phillips et al., 2011). In the present studies, all samples were large enough to have greater than 95% power to replicate this key effect with an alpha of .05. Sample sizes were always determined a priori and the data were never analyzed prior to completing data collection. We aimed to recruit participants in three categories: (1) Non-researchers, who had never conducted philosophical or scientific research, (2) Researchers who had conducted research but were not experts on happiness research in particular, and (3) Experts who studied and conducted research on happiness. 294 Non-researchers (127 females; 2 unreported, M_{age}= 31.21, SD_{age} =11.16) were recruited from the United States using Amazon's Mechanical Turk (Buhrmester, Kwang & Gosling 2011; Gosling, Vazire, Srivastava & John 2004). 299 Researchers and Experts (184 females; 13 unreported, M_{age} = 34.95, SD_{age} = 10.31) were recruited through academic psychology listervs (e.g., International Society for Research in Emotion, Judgment and Decision

Making, Society for Affective Neuroscience, Society for Philosophy and Psychology, Society for

Affective Science, Society for Psychophysiological Research, Society for Research in Child

Development, and Society for Research in Psychopathology).

Measures

      Happiness attribution. Participants rated their agreement with the following statement

"[Agent] is happy" on a scale from 1 ("completely disagree") to 7 ("completely agree").

      Comprehension check. Three items measured whether participants understood that the

agent satisfied all three criteria of the scientific definition of happiness. Specifically, participants

provided a true/false rating for three statements of the following form: (1) "[Agent's name] feels

good and generally experiences a lot of pleasant emotions"; (2) "[Agent's name] rarely

experiences negative emotions", and (3) "[Agent's name] feels highly satisfied with the way

[she/he] lives." Because we wanted to ensure that any difference in responses was due to a moral

difference but not a difference in whether the agent was viewed as having the scientifically

agreed upon traits of happiness, participants who answered "False" on any one of these three

questions were excluded.

      Justification of happiness ratings. Participants in the within-subjects condition were

asked to respond to the following prompt: "Please explain why the difference between the two

passages either was or was not relevant to the question of happiness."

      Demographic items. Demographic items measured participants' age, gender, ethnicity,

education, socio-economic status (SES), and whether or not they conduct philosophical or

scientific research.

      Happiness expertise scale. Participants responded to the three following items on a scale

from 1 ("not at all") to 7 ("completely"): (1) "To what degree do you identify yourself as a

happiness / positive psychology researcher?" (2) "To what degree do you consider yourself an emotion researcher / affective scientist?" (3) "How familiar are you with the research on positive psychology and happiness?"

Study 1 Procedure

Participants were randomly assigned to either a between-subjects version of the study or a within-subjects version of the study. In the between-subjects version, participants were randomly assigned to read one brief vignette about one of three different agents (a janitor, a nurse, or an uncle). For each agent, participants were then randomly assigned to either a "morally good" or "morally bad" condition. In the morally good condition, participants read a paragraph which described the agent as living a morally good life (helping students, caring for sick children, etc.); in the morally bad condition, participants read a paragraph that described the agent as living a morally bad life (cheating on a spouse, killing children, etc.). The full vignettes and all other study materials are available at https://github.com/phillipsjs/trueHappiness. In the within-subjects version, participants were randomly assigned to read both the morally good and morally bad vignettes for one of the three different agents. The two vignettes were presented in counterbalanced order.

Across both conditions, in each of the six vignettes in the 2 (Moral Value) x 3 (Agent) design, the agents' mental states were always described as satisfying the criteria for the scientific definition of happiness, having high levels of positive affect, low levels of negative affect, and high levels of life-satisfaction. Furthermore, the description of the agents' mental states were matched across all vignettes, such that only the specific agent's name and personal pronouns were changed. After reading the vignette(s), all participants completed the happiness attribution item and the comprehension check items. Participants in the within-subjects condition

additionally completed the justification of happiness ratings item. Finally, all participants completed the demographic items and expertise scale.

Data Analysis Approach

When appropriate, our analyses are conducted with linear mixed-effects models that allow for participants and items (e.g., different scenarios) to be modeled as random factors. This approach to analysis allows for greater confidence that significant results will generalize beyond the specific materials used and participants recruited (Baayen, Davidson & Bates, 2008; Gelman & Hill, 2007; Jaeger, 2008). This statistical approach was implemented using the lme4 package in R (Bates, Maechler, Bolker & Walker 2015). To determine the significance of each effect, we compare a model that includes the relevant term in question (as well as all of the other factors that are not currently being investigated) to a model that does not include that term (but does include all of the other factors). The effect is taken to be significant if the fit of the model including the term differs significantly from the fit of the model that does not include that term (Barr, Levy, Scheepers, & Tily, 2013). For simplicity and clarity, follow-up comparisons and effect sizes are included as a way of illustrating the patterns observed in the overall analysis. For each of the analyses we report, similar results were also obtained using the more traditional approach of relying on analyses of variance. The statistical code for running both sets of analyses is provided (along with the data and stimuli) at: https://github.com/phillipsjs/trueHappiness.

Study 1 Results

Preliminary Analyses. Participants were divided into three groups based on their responses to the happiness expertise items: (1) Non-researchers ($n = 324$), (2) Researchers ($n = 140$), and (3) Experts ($n = 129$). On a scale from 1 (no knowledge of happiness research) to 20

(complete expertise on happiness) those in the Experts group had a mean well above the midpoint ($M = 13.56$; $SD = 2.19$); those in the Researchers group had a mean below the midpoint ($M = 7.26$; $SD = 2.09$); and those in the Non-researchers group had almost no knowledge of happiness research ($M = 3.71$; $SD = 1.20$). There were no consistent effects of age, gender, education or SES the critical analyses in these studies. Accordingly, these variables are not discussed further, though the full analyses are available at https://github.com/phillipsjs/trueHappiness

Primary Analyses. We first analyzed participants' responses to the happiness attribution item in the between-subjects version of the study. 45 participants were excluded for failing to answer one of the three comprehension questions correctly. The remaining 269 participants' responses were analyzed by comparing linear mixed effects models that treated Moral Value and Happiness Expertise as fixed factors and Agent as a random factor. The analysis revealed a significant main effect of Moral Value, $X^2 (1) = 28.71$, $p < .001$. Planned comparisons revealed that participants rated the agents as happier in the morally good condition ($M = 6.34$, $SD = 1.09$) than in the morally bad condition ($M = 5.44$, $SD = 1.62$), $t(212.50) = -5.27$, $p < .001$, $d = 0.66$. We did not observe any effects of Expertise or an Expertise x Moral Value interaction, p's $\geq$ .591.[1]

To investigate whether the effect of Moral Value persisted when participants were made aware that the only difference in the scenarios was the moral value of the agent's life, we next analyzed participant's happiness attributions in the within-subjects version. 95 participants were excluded for failing to answer one of the six comprehension questions (three for each of the two vignettes) correctly. The remaining 184 participants' responses were analyzed and again revealed

---

[1] Treating the research scale as a continuous variable rather than factoring participants into expertise groups does not change the pattern of results.

a main effect of Moral Value, X^2 (1) = 48.65, $p < .001$, such that participants again rated the agents as happier in the morally good condition ($M = 6.48$, $SD = 0.84$) than in the morally bad condition ($M = 5.63$, $SD =1.74$), $t(183) = 7.44$, $p < .001$, $d = 0.549$. We did not observe any other main effects or interactions, $p's \geq .645$ (Fig.2).[2]

Study 1 Discussion

Taken together, the results of this study suggest that the influence of moral value on assessments of happiness is highly robust. Moral value influenced participants' judgments whether the study was administered between-subjects or within-subjects, suggesting that at least some participants were willing to more reflectively endorse the relevance of evaluative considerations for happiness. More strikingly, this pattern persisted across all three groups of participants, whether they were untrained in scientific or philosophical study, highly trained in scientific or philosophical study, or even highly trained in the study of happiness itself. In sum, the impact of moral value on assessments of happiness is unlikely to be explained as a misunderstanding of the task or a simple error on the part of participants. Rather, the effect of moral value seems to reflect a robust aspect of how people ordinarily make assessments of happiness.

Study 2: Bias or Distortion - Test 1

While the previous results suggest that the influence of moral value is not due to a simple misunderstanding or lack of competence with the task, another plausible explanation is that the

---

[2] Including all participants, we continue to observe the key main effect of Moral Value, both in the between-subjects condition, X^2 (1) = 51.24, $p < .001$, and in the within-subjects condition, X^2 (1) = 133.98, $p < .001$. Participants' justifications for differences in their happiness assessments were coded as falling within one of four different categories: (1) the agents experienced different descriptive states (~48%), (2) one of the agents was disliked more by the participant (~14%), (3) morality was directly relevant to the agents' happiness (~22%), or (4) other (~16%). We did not find a significant difference in the kinds of responses provided by Non-researchers, Researchers, and Experts, X^2 (6) = 6.96, $p = .143$. All justifications, coding, and analysis are available at https://github.com/phillipsjs/trueHappiness

effect arises primarily from a motivational bias. In particular, it may be that participants found themselves reluctant to attribute positive qualities like happiness to people who they regarded as bad (Ditto, Pizarro & Tannenbaum, 2009), or their assessments of the agent's happiness is distorted by a desire to maintain a belief in a just world (Lerner, 1980). Thus, even if moral evaluations do play a robust role in assessments of happiness, they still may not reflect genuine features of the ordinary *concept* of happiness.

To pursue this possibility, we attempted to directly manipulate the degree to which participants viewed happiness as a genuinely good thing. To the extent that we are able to successfully change participants' views on the desirability or value of happiness, we can then ask whether these changes are reflected in the degree to which moral value influences assessments of happiness. More specifically, if the effect in the previous study arose from a motivational bias, then we would expect this effect to decrease when participants are given evidence that happiness is not actually such a desirable or good thing. Alternatively, if the effect of morality on assessments of happiness reflects an aspect of the ordinary concept of happiness, then we would expect moral value to influence assessments of happiness regardless of the extent to which participants are thinking of happiness as a good or desirable property.

Study 2 Methods

Participants

Participants were 125 adults (77 females, 3 unreported; $M_{age} = 33.85$, $SD_{age} = 12.12$) from the United States recruited through Amazon's Mechanical Turk. Additional demographic information for this and all following studies can be found at https://github.com/phillipsjs/trueHappiness

Measures

Happiness Attribution. As in Study 1, participants rated their agreement with the statement "[Agent] is happy" on a 1 to 7 scale.

Happiness Check. Three 'happiness check' items were assessed to examine whether we successfully changed people's attitudes toward happiness. These included measuring beliefs about happiness, change in happiness beliefs, and self-reported subjective happiness. For happiness beliefs, participants rated the extent to which they agreed with the statement "Happiness is a good thing". For change in happiness beliefs, participants reported whether or not this research had changed their original view of happiness. For self-reported happiness, participants self-reported their current happiness levels. All items were rated on 1 to 7 scale.

Comprehension Check. To ensure that participants fully understood the information presented during the video lecture, participants were asked to report three new facts they learned about happiness. Participants provided written descriptions of three facts.

Demographic items. Demographic items measured participants' age, gender, ethnicity, education and socio-economic status (SES).

Study 2 Procedure

Participants first watched a brief (less than 3 minute) video-lecture (given by Author JG) based on actual scientific work on happiness (e.g., Fredrickson, 2001; Gruber, et al., 2011; Lyubomirsky et al., 2005). The video lectures were similar in length and amount of content. Participants were randomly assigned to either a 'happiness is good' condition in which they viewed a brief video lecture on studies suggesting that happiness is associated with good characteristics (creativity, prosociality, positive health outcomes) or a 'happiness is bad' condition in which they viewed a lecture on studies suggesting that happiness is associated with bad characteristics (selfishness, discrimination, negative health outcomes). Both conditions

portrayed information from actual scientific studies and were equated for length and

presentation. (Transcripts of each lecture, and the video lecture itself can be found in the

Supplementary Materials available at: https://github.com/phillipsjs/trueHappiness). Participants

were instructed that they would be asked to report three facts they learned from the lecture after

viewing it. They completed this comprehension check and also the happiness check items after

finishing watching the video lecture.

Subsequently, participants read a vignette about an agent that described the agent's life

and psychological states. Participants were randomly assigned to read either about an agent

living a morally good life or about an agent living a morally bad life. As in Study 1, the agent's

psychological states were identical in the two conditions: the agent was described as having high

levels of positive affect, low levels of negative affect, and high levels of life satisfaction. After

reading the vignette, participants completed the happiness attribution and demographic items.

Study 2 Results

Preliminary Analyses

Participants' responses to the comprehension check items were coded according to

whether they were facts reported in the video lecture, and subsequent analyses were conducted

only on participants who correctly provided three facts about happiness from the video lecture

(94%; $n = 117$). We first examined the three manipulation checks to ensure that we successfully

altered the extent to which participants believed that happiness is a good thing. Participants in the

'happiness is bad' condition agreed significantly less with the claim that "Happiness is a good

thing" ($M = 5.23$, $SD = 1.16$) than did participants in the 'happiness is good condition' ($M =$

6.69, $SD = 0.55$), $t(102.82) = -9.10$, $p < .001$, $d = 1.524$. Participants also reported that they

valued happiness significantly less in the happiness is bad condition ($M = 3.80$, $SD = 0.99$) than

in the happiness is good condition ($M = 5.10$, $SD = 1.10$), $t(116) = -6.71$, $p < .001$, $d = 1.254$.

Finally, participants were inclined to describe themselves as less happy in the happiness is bad

condition ($M = 4.52$, $SD = 1.18$) than in the happiness is good condition ($M = 5.04$, $SD = 1.14$),

$t(116) = -2.39$, $p = .019$, $d = 0.446$. Taken together, these results suggest that we were able to

successfully manipulate the extent to which participants perceived happiness as a good thing.

After being exposed to research suggesting happiness may actually be bad, participants were

more likely to disagree that happiness was actually a good thing, reported valuing happiness less,

and were less inclined to describe themselves as happy.

Primary Analyses. The critical question we can now ask is whether these changes in

participants' views about happiness influenced the extent to which moral value affected their

happiness assessments. As in Study 1, there was a main effect of Moral Value $F(1,114) = 78.95$,

$p < .001$, $\eta_p^2 = .409$, such that participants judged agents living morally bad lives to be less happy

($M = 3.88$, $SD = 1.83$) than agents living morally good lives ($M = 6.27$, $SD = 1.04$), $t(85.38) = -$

8.63, $p < .001$, $d = 1.611$. There was no significant main effect of Happiness Research, $F(1,114)$

$= 1.66$, $p = .200$, $\eta_p^2 = .014$, and critically, there was also no interaction between happiness

research and moral value, $F(1,114) = 0.30$, $p = .587$, $\eta_g^2 = .003$ (see **Figure 3**).

Study 2 Discussion

In brief, these results suggest that (1) we successfully manipulated the extent to which

participants perceived happiness to be a good thing, but that (2) even when we decreased

participants' belief that happiness was good, they still showed the same influence of moral value

on their assessments of happiness. This pattern is not predicted by an account on which

morality's effect on assessments of happiness is due to a motivational bias. By contrast, this

pattern would be expected if the effect of moral value reflected an aspect of the ordinary concept of happiness.

Study 3: Bias or Distortion - Test 2

One straightforward prediction of all motivational accounts is that the more people hold a negative attitude toward a given agent, the more they should be reluctant to regard an agent as happy. In other words, even among cases in which people regard the agent's life as morally bad, there should be an effect such that people are *more* reluctant to attribute happiness when they have a strongly negative attitude toward the agent than when they have just a slightly negative attitude.

By contrast, this prediction would not necessarily fall out of theories according to which the concept of happiness is itself evaluative. Instead, the prediction would depend on the details of the theory. The core idea of an evaluative theory of happiness is just that moral considerations play some role in the criteria for determining whether a person is happy. One way to spell out such a theory would be to say that happiness attributions should be in some way proportional to moral goodness, with an agent being regarded as ever less happy as her life becomes ever less morally good. However, it would also be possible to spell out such a theory in other ways, such that moral goodness does play a role but happiness is not simply proportional to goodness. (For further discussion, see General Discussion.)

To explore this issue, we manipulated the degree to which the agent's life was morally bad. Participants in both conditions received a vignette about an agent who had a morally bad life, but participants in one condition were told about an agent whose life was just slightly morally bad, while those in the other were told about an agent who was truly evil. The key question was whether participants would be more reluctant to attribute happiness to the evil

20

agent than to the agent whose life was just slightly morally bad.

Study 3 Methods

Participants

To pretest the difference in the moral value of the agents' lives, we recruited 50 participants (27 females; $M_{age} = 31.80$, $SD_{age} = 9.73$). We additionally recruited 201 adults in the main study (69 females; $M_{age} = 30.75$, $SD_{age} = 8.30$) and asked them to evaluate the happiness of the agents. All participants were from the United States and again recruited through Amazon's Mechanical Turk.

Measures

Moral Value Pretest. Participants rated the moral value of the agent's life on a scale from 1 ('Most immoral life possible') to 7 ('Completely morally fine').

Happiness Assessment. As in Study 1, participants rated their agreement with the statement "[Agent] is happy" on a 1 to 7 scale.

Demographic Items. Demographic items measured participants' age, gender, ethnicity, education and socio-economic status (SES).

Study 3 Procedure

Participants were randomly assigned to read a brief vignette about one of two possible agents. The agent was either described as living a life that was slightly immoral (i.e., a shallow life that involved lying and using other people to get ahead) or as living a life that was fundamentally evil (i.e., a life dedicated to white supremacy and violent terrorism). In both cases, the agent was described as having the three psychological states associated with happiness: high levels of positive affect, low levels of negative affect, and high levels of life-satisfaction. As in the previous studies, the agent's mental states were identical in the two conditions. (All stimuli,

data and analyses are available at: https://github.com/phillipsjs/trueHappiness). After reading the vignette, all participants completed the happiness assessment item and finally the demographic items.

Study 3 Results

We first compared participants' moral judgments, which revealed that they believed the fundamentally evil agent to be living a much worse life ($M = 5.85$; $SD = 1.32$) than the slightly immoral agent ($M = 4.29$; $SD = 1.40$), $t(48) = -4.05$, $p < .001$, $d = 1.15$.

Given this difference, we next compared participants' assessments of happiness for the two agents and found that they actually judged that the fundamentally evil agent was *more* happy ($M = 5.74$; $SD = 1.59$) than the slightly immoral agent ($M = 5.23$; $SD = 1.68$), $t(199) = -2.22$, $p = .028$, $d = 0.313$ (see **Figure 4**).

Study 3 Discussion

To test between general motivational and specific conceptual explanations of the impact of morality, we asked whether increasing the extent to which participants held negative attitudes toward the agent also increased their disinclination to regard the agent as happy. The pattern of responses we observed strongly suggests the opposite. Participants actually regarded the fundamentally evil agent as significantly *happier* than the slightly immoral agent. This result complements the pattern observed in Study 2, and together they help to build a case against an explanation that appeals solely to general motivated cognition or just world beliefs.

The obvious next question is how this puzzling result could be explained in terms of the structure of the concept of happiness. However, before offering an explanation that appeals to some specific structure of the concept of happiness, it is necessary to first determine whether the

concept of happiness should be understood as a descriptive or evaluative one (see **Figure 1b** and **1c**). If the concept of happiness turns out to be purely descriptive, then this result could be explained by suggesting that participants perceived the evil agent as having fewer negative psychological states (e.g., remorse or regret) than the slightly immoral agent. In contrast, if the concept of happiness turns out to be evaluative, then this result could instead be explained by suggesting that there is some value that was relevant to happiness that the evil agent satisfied more than the slightly immoral agent (e.g., having a life that is structured by a deeper meaning). Accordingly, we first turn to addressing the question of whether the folk concept of happiness is descriptive or evaluative. After doing so, we return to this question in the General Discussion and offer one way of understanding these puzzling results.

Study 4a and 4b: Is The Effect Mediated by Changes in The Perception of Descriptive Psychological States?

The first three studies provide evidence that participants' assessments of happiness are affected by their moral evaluations in a way that is unlikely to be the result of a performance error or general motivated cognition. Even setting these explanations to one side, however, there are still two very different ways to explain morality's impact. One possibility is that the ordinary concept, like the scientific understanding, is only concerned with the descriptive psychological states that an agent experiences, and that changes in the moral value of an agent's life change whether the agent is understood as experiencing these states. An alternative possibility is that, unlike the scientific definition, the ordinary concept of the happiness is genuinely evaluative, and the moral value of an agent's life is directly relevant to whether or not an agent is understood to

be happy.

Studies 4a and b provide two tests of the first possibility by asking whether the effect of morality can be explained by changes in the perception of the descriptive psychological states the agent is seen as experiencing.

Study 4a did this by asking participants to complete two closely matched measures which asked participants either to assess whether the agent was happy or to assess whether the agent felt good. In both cases, participants read about an agent who was described as living a morally good or a morally bad life, and then were asked to indicate their agreement with a statement that either described the agent as being happy or feeling good.

Study 4b instead did this by developing a novel measure that allowed participants to report, in a very general way, the extent to which they believed the agent experienced positive rather than negative psychological states. To do this, participants selected one of seven faces, which were generated through facial morphing such that the agent's expression ranged from one of high negative affect to one of high positive affect (see **Figure 6** for an example). Participants were again asked to read about an agent who was either living a morally good or bad life, and were then randomly assigned either to assess how the agent felt (on this new measure), or to assess how happy the agent was (using the same measure from previous studies). We additionally manipulated whether the agent's psychological states were described as being positive overall or negative overall, which provided a way of testing whether the newly developed measure succeeded in tracking the agent's descriptive psychological experience.

Together, these studies provide convergent ways to compare the impact of moral evaluations across assessments of happiness and assessments of the descriptive psychological states the agent was perceived as experiencing.

Study 4a Methods

Participants

Participants were 203 adults (121 females, 1 unreported; $M_{age} = 37.55$, $SD_{age} = 13.38$) from the United States again recruited through Amazon's Mechanical Turk.

Measures

Happiness Assessment. Participants were asked to tell us whether they agreed or disagreed with a statement about the agent in the story they read. They then rated their agreement with a statement of the form "[Agent] is happy." on a scale from 1 ('Disagree') to 7 ('Agree').

Affect Assessment. Similarly, participants were asked to tell us whether they agreed or disagreed with a statement about what the agent in the story felt like. They then rated their agreement with a statement of the form "[Agent] feels good." on a scale from 1 ('Disagree') to 7 ('Agree').

Demographic Items. Demographic items measured participants' age, gender, ethnicity, education and socio-economic status (SES).

Study 4a Procedure

Participants were randomly assigned to read a brief vignette about one of four possible agents (mother, janitor, nurse, or uncle). For each agent, participants were randomly assigned to either a "morally good" or "morally bad" condition. This resulted in a total of eight different vignettes. In all cases, the agent was described as experiencing mental states that satisfy the descriptive definition of happiness (high levels of positive affect, low levels of negative affect, and high levels of life-satisfaction). The agent's mental states were phrased in a parallel manner across all vignettes so that only the specific agent name and personal pronouns were changed.

(All stimuli, data and analyses are available at:).After reading the vignette, all participants completed either the affect assessment item or the happiness assessment item (administered between-subjects). Finally, participants completed the demographic items.

Study 4a Results

We observed a main effect of Moral Value, $X^2(1) = 4.89$, $p = .027$, and a marginal effect of Assessment Type $X^2(1) = 3.04$, $p = .081$. Critically, however, these effects were qualified by a significant Moral Value x Assessment Type interaction, $X^2(1) = 5.35$, $p = .021$. For assessments of happiness, participants rated the agents as happier when they were living a morally good life ($M = 6.59$, $SD = 0.88$) than when they were living a morally bad life ($M = 5.87$, $SD = 1.70$), $t(78.47) = -2.73$, $p = .008$, $d = 0.530$. For assessments of affect, however, participants rated the agent as having similar affective states whether that agent was living a morally good life ($M = 6.60$, $SD = 0.65$) or a morally bad life ($M = 6.57$, $SD = 0.81$), $t(97) = -0.19$, $p = .847$, $d = 0.039$ (see **Figure 5**).

Study 4b Methods

Participants

Participants were 403 adults (159 females, 1 unreported; $M_{age} = 29.24$, $SD_{age} = 8.98$) from the United States again recruited through Amazon's Mechanical Turk.

Measures

Happiness Assessment. As in Study 1, participants rated their agreement with the statement "[Agent] is happy" on a 1 to 7 scale.

Emotion Assessment. Participants were told that we were interested in their own

thoughts about how the agent feels. They were then asked to tell us how they thought the agent

felt on a scale of morphed faces from 1 (a photo of a face expressing high negative affect) to 7 (a

photo of a face expressing high positive affect). Four versions of this seven-face scale were

created; one for each of the four agents participants read about. The endpoint photos of the scale

were taken from the Karolinska Directed Emotional Faces (Lundqvist, Flykt & Öhman, 1998)

and the middle five facial morphs were generated using FantaMorph 5 facial morphing software

(See **Figure 5**).

Comprehension Check. To ensure that participants read the vignettes carefully, they

were asked to summarize the life of the person they read about.

Demographic Items. Demographic items measured participants' age, gender, ethnicity,

education and socio-economic status (SES).

Study 4b Procedure

Participants were randomly assigned to read a brief vignette about one of four possible agents

(mother, janitor, nurse, or uncle). For each agent, participants were randomly assigned to either a

"morally good" or "morally bad" condition. For each of these eight possible scenarios,

participants were randomly assigned to read either about an agent experiencing positive mental

states or about an agent experiencing negative mental states. This resulted in a total of sixteen

different vignettes. In the positive mental states condition, the agents' mental states were

described as having the three critical ingredients of happiness: high levels of positive affect, low

levels of negative affect, and high levels of life-satisfaction. In the negative mental states

condition, the agents' mental states were described as lacking these three critical ingredients: the

agent had low levels of positive affect, high levels of negative affect, and low levels of life-

satisfaction. As in the previous studies, the agent's mental states were phrased in a parallel

manner across all vignettes so that only the specific agent name and personal pronouns were changed. (All stimuli, data and analyses are available at:

https://github.com/phillipsjs/trueHappiness). After reading the vignette, all participants completed either the affect assessment item or the happiness assessment item (administered between-subjects). Finally, participants completed the comprehension check question and demographic items.

Study 4b Results

We first tested whether the effect of morality was specific to happiness. Critically, we observed a significant Moral Value x Assessment Type interaction, $X^2 (1) = 8.01$, $p = .005$ (see **Figure 6**). For assessments of happiness, participants rated the agents as happier when they were living a morally good life ($M = 4.11$, $SD = 2.61$) than when they were living a morally bad life ($M = 3.29$, $SD = 2.46$), $t(197) = -2.29$, $p = .023$, $d = 0.325$. For assessments of affect, however, participants rated the agent as having similar affective states whether that agent was living a morally good life ($M = 4.57$, $SD = 2.18$) or a morally bad life ($M = 4.21$, $SD = 2.19$), $p = .240$, $d = 0.165$.

We also observed a Moral Value x Psychological State interaction effect, $X^2 (1) = 5.32$, $p = .021$, such that the agent's life value had a larger effect on assessments when the agent was experiencing positive psychological states, $t(141.05) = -4.872$, $p < .001$, $d = 0.686$, than when the agent was experiencing negative psychological states $t(199) = -1.76$, $p = .080$, $d = 0.249$. In addition, we observed an Assessment Type x Psychological State interaction effect, $X^2 (1) = 12.61$, $p < .001$, such that the agent's psychological states had a somewhat bigger effect on assessments of happiness, $t(154.87) = 26.92$, $p < .001$, $d = 3.846$, than on assessments of general emotion, $t(186.43) = 24.79$, $p < .001$, $d = 3.489$. As expected, we also observed main effects of Moral Value, $X^2 (1) = 24.22$, $p < .001$, Psychological States, $X^2 (1) = 611.83$, $p < .001$, and

Assessment Type, X^2 (2) = 30.57, $p < .001$. There was no three-way Moral Value x

Psychological State x Assessment Type interaction, X^2 (2) = 2.24, $p = .135$.

Study 4a and 4b Discussion

Moral value significantly impacted the extent to which participants understood the agent to be

happy, but did not similarly affect the descriptive psychological states the agent was perceived as

experiencing. This pattern suggests that moral evaluations are not affecting assessments of

happiness solely by changing the descriptive psychological states the agent is perceived as

experiencing. Instead, it may be that the impact of moral value arises because the ordinary

concept of happiness is evaluative and provides a specific role for moral value. We pursue this

possibility in a final study.

Study 5: The Role of Moral Value in The Ordinary Concept of Happiness

If moral judgments actually play a role in the criteria governing people's concept of

happiness, then the impact of morality on happiness attributions should exhibit a specific pattern

that differs sharply from the impact of factors that are not part of the criteria. To see this,

consider for example the impact of being very poor on judgments of whether someone is happy.

While being wealthy is not one of the criteria governing people's concept of happiness, it may

still be the case that believing that someone is very poor will make you less likely to think that

person is happy. Such an effect presumably occurs because believing a person is poor also leads

participants to imagine that this person is unlikely to meet the criteria that actually are part of the

concept of happiness (e.g., experiencing high positive affect and low negative affect). Thus, if

participants found out that the person did experience high positive affect and low negative affect

even though she is poor, then the impact of being poor on happiness attributions should be screened off. By contrast, factors that actually play a role in the criteria governing the concept of happiness should exhibit a different pattern: even if participants are explicitly told about the agent's psychological states, the impact of these factors should remain. We investigate these possibilities with respect to the impact of moral judgments (see Figure 1c).

To ask whether the impact of morality differs in this way from the impact of other factors, we compared a manipulation of non-moral aspects of an agent's life (e.g., having a good vs. bad job) with a manipulation of moral aspects (e.g., caring for vs. harming sick children). Given the previous research on the topic (King & Napa, 1998), as well as the results of the preceding studies, we expect that both moral and non-moral factors are likely to impact assessments of happiness in the absence of information about the agents' descriptive psychological states. The key question is whether these effects will be differentially reduced when explicit information about the agents' psychological states is provided.

Study 5 Methods

Participants

Participants were 400 adults (182 females; $M_{age} = 32.39$, $SD_{age} = 10.25$) from the United States again recruited through Amazon's Mechanical Turk.

Measures

Happiness Assessment. As in the previous studies, participants were asked to tell us whether they agreed or disagreed with a statement about the agent in the story they read. They then rated their agreement with a statement of the form "[Agent] is happy." on a scale from 1 ('Completely disagree') to 7 ('Completely agree').

Demographic Items. Demographic items measured participants' age, gender, ethnicity,

education and socio-economic status (SES).

Study 5 Procedure

Participants were randomly assigned to the moral value condition or the non-moral value condition. Within the moral value condition, participants were randomly assigned either to read about three agents who were described as living morally good lives (helping disabled students, caring for sick children, taking care of a young relative), or to read about three agents who were described as living morally bad lives (stealing from students, harming sick children, molesting a young relative). Within the *non*-moral value condition, participants were randomly assigned either to read about three agents who were living lives that were good in non-moral ways (being very wealthy, having a very comfortable job, living in an exciting area), or to read about three agents who were living lives that were bad in non-moral ways (being very poor, having a terrible job, living in a remote and desolate area). In addition, participants were randomly assigned either to receive or to not receive explicit information about the agents' psychological states. In all cases where the agents' psychological states were described, the agents were described as experiencing the three psychological states associated with happiness: high levels of positive affect, low levels of negative affect, and high levels of life-satisfaction.

Overall, this resulted in a mixed design with morality (moral vs. non-moral), value (good vs. bad) and mental state information (present vs. absent) as between-subjects factors and the three agents (a young nurse, a forty-five-year-old man, a school janitor) as a within-subjects factor. (All stimuli, data and analyses are available at: https://github.com/phillipsjs/trueHappiness). After reading each vignette, participants completed the happiness assessment item for that agent, and after assessing the happiness of all three agents, participants completed the demographic items.

Study 5 Results

An analysis of participants' assessments of happiness revealed the key three-way interaction effect between Morality, Life Value and Psychological Information, $X^2(9) = 47.72$, $p < .001$ (see **Figure 7**). We decomposed this interaction by looking separately at the moral and non-moral value conditions.

Considering the Moral Value condition first: when no information was provided about the agents' descriptive psychological states, participants reported that the agent was significantly happier when leading a life that was morally good ($M = 4.65$, $SD = 0.83$) than when leading a life that was morally bad ($M = 2.20$, $SD = 1.20$), $t(85.23) = -11.83$, $p < .001$, $d = 2.38$. Moreover, even when information was provided that the agent had positive descriptive psychological states, participants continued to report that the agent was much happier when leading a morally good life ($M = 6.29$, $SD = 0.62$), than when leading a morally bad life ($M = 4.93$, $SD = 1.94$), $t(59.13) = -4.70$, p , .001, $d = 0.94$.

This pattern in the Non-Moral value condition was quite different. When no information about the agents' descriptive psychological states was provided, participants reported that the agent was much happier when leading a life that was non-morally good ($M = 5.76$, $SD = 0.75$), than when leading a life that was non-morally bad ($M = 2.57$, $SD = 1.18$), $t(84.48) = -16.26$, $p < .001$, $d = 3.22$. In sharp contrast, however, when information was provided that the agent had positive descriptive psychological states, participants no longer reported that the agent was happier when leading a life that was non-morally good ($M = 6.49$, $SD = 0.89$), than when leading a life that was non-morally bad ($M = 6.20$, $SD = 0.93$), $t(98) = -1.61$, $p = .111$, $d = 0.32$.

Study 5 Discussion

The impact of moral value on happiness assessments differed sharply from the impact of *non-moral* value. Non-moral value impacted assessments of happiness in the absence of information about the agent's psychological states, but this effect was eliminated when information about the agents' psychological states was made explicit. This pattern suggests that non-moral value initially impacted assessments of happiness by altering whether or not the agent was perceived as experiencing the descriptive criteria for happiness, and accordingly, providing this information explicitly screened off the impact of non-moral value on assessments of happiness.

The impact of moral value exhibited a different overall pattern. Consistent with previous research (King & Napa, 1998), the impact of moral value was greatest when no information was given about the agents' psychological states, suggesting that people typically exhibit an inferential connection between happiness and morality. In contrast to non-moral value, however, the impact of morality was not completely screened off when information about the agents' psychological states was provided. Thus, these results both provide a conceptual replication of previous findings on the relationship between happiness and morality (King & Napa, 1998), and additionally demonstrate that moral value is special in that it seems to serve as one of the independent criteria that govern the ordinary concept of happiness (see **Figure 1c**).

General Discussion

Five studies explored the role of morality in the ordinary concept of happiness, and collectively shed light on the question of *why* moral judgments affect ordinary assessments of happiness. Study 1 demonstrated the robustness of the effect of moral value, showing that it arises in expert researchers, even in a within-subject design. Studies 2 and 3 then provided evidence that the effect of morality on judgments of happiness is unlikely to be explained by

general motivational biases or beliefs in a just world. Lastly, Studies 4 and 5 tested whether the effect of morality on assessments of happiness occurs *indirectly*, through changes in the psychological states the agent is perceived as experiencing, or *directly*, suggesting that morality is actually one of the criteria of the ordinary concept of happiness. Study 4 revealed that the impact of moral judgments cannot be explained by changes in the perception of descriptive psychological states. Study 5 provided evidence that moral value, unlike non-moral value, has a direct impact on assessments of happiness, suggesting that it is one of the criteria that govern the ordinary concept of happiness.

While previous research demonstrated *that* moral judgments affect assessments of happiness (e.g., Phillips et al., 2011), the present results provide a deeper understanding of why this is the case. Across five studies, we considered three explanations of the effect of morality: (1) biases or performance errors, (2) mediation through descriptive psychological states, and (3) an evaluative folk concept of happiness. While these first two explanations are well-studied and receive support from previous research (King & Napa, 1998; Lerner, 1980), our studies provide evidence for the third explanation. Above and beyond performance errors or changes in descriptive psychological states, we find that morality has a *direct* impact on ordinary assessments of happiness.

Happiness and the Pervasive Impact of Morality

Previous research has found that, in addition to happiness, many others kinds judgments are also impacted by morality, including judgments about knowledge, intentional action, causation, and freedom among others (Beebe & Buckwalter, 2010; Knobe, 2003; Kominsky, et al., 2015; Leslie, et al., 2006; Samland, et al., 2016; Phillips, et al., 2015; Young & Phillips, 2011). Thus, a natural question arises as to whether the effect we have been investigating is

related to these previous findings.

One intriguing possibility is that the theories that have been developed to explain these various effects could also be used to account for the impact of morality on assessments of happiness. In fact, a number of researchers have recently been developing accounts that can be extended to many of the judgments that are impacted by morality (Adams & Steadman, 2004; Alfano, Beebe & Robinson, 2012; Driver, 2008a; 2008b; Phillips et al., 2015; Uttich & Lombrozo, 2010). However, all of these accounts were developed to explain effects that showed precisely the opposite pattern from the one exhibited by assessments of happiness. For example, judgments that an outcome is morally bad lead people to judge that the agent more intended the outcome, was more in favor of it, more decided to bring it about, and so forth (Pettit & Knobe, 2009). In contrast, a judgment that an agent acted immorally leads people to judge that an agent was *less* happy. Thus, it difficult to see how the theories designed to explain those earlier findings could be extended to account for the present effects.

Given the difficulty faced by extant proposals for unifying morality's influence, a plausible alternative possibility is that the effect of morality on judgments of happiness may be best explained by specific features of the concept of happiness. If this is correct, then a better understanding this particular effect we have been investigating will require a better understanding of concepts that involve both evaluative and descriptive criteria.

The Ordinary Concept of Happiness

We have argued that the effects observed in the present studies are to be explained in terms of the structure of the concept of happiness itself. In other words, we suggest that the impact of moral considerations is best explained by suggesting that the concept of happiness is not a purely descriptive one but actually involves a role for moral factors.

This claim immediately opens up a host of new questions. It seems clear that the concept is associated with certain purely descriptive features (concerning the agent's psychological states), but the present studies suggest that it is also associated with certain deeper values (concerning the moral status of the agent's life). What then is the relationship between these aspects of the concept?

While further studies will be required before proposing any definite account, we want to suggest some promising avenues for future research. First, we suggest that a good way to make sense of the puzzling characteristics of the concept of happiness might be to turn to more general theories of concepts. Presumably, the concept of happiness is not completely sui generis. Thus, if we look to a broader class of concepts and try to develop a more general account of how they work, we may be able to gain some valuable insight into the concept of happiness in particular.

It does seem that certain other concepts show a similar duality between descriptive features and deeper values. As an example, consider the concept *scientist*. People seem to associate this concept with various descriptive features (running experiments, doing statistics, developing theories). Yet it also seems that people associate the concept with a deeper value. Consider a physics professor who clearly displays all the descriptive features but who does not seem to be involved in a genuine quest for truth. There seems to be some sense in which such a person would clearly count as a scientist, but all the same, people tend to say that there is a deeper sense in which she is not truly a scientist at all (Knobe, Prasada & Newman, 2013). Conversely, consider a child who has had no formal training in experimentation, statistics, or theory construction, but who is deeply interested in the underlying truths of how things operate and seeks to understand them. There again seems to be some sense in which the child clearly would not count as a scientist, but another deeper sense in which the child truly does seem to be

a scientist. Existing studies show that this sort of pattern arises for a variety of different concepts, including not only *scientist* but also *friend*, *soldier*, *poem* as well as psychological state concepts like *love* (Knobe, Prasada & Newman, 2013).

Within existing research, concepts of this general type have been referred to as 'dual character concepts' (Knobe et al., 2013). What makes such concepts distinctive is the fact that each such concept seems to be associated with two different criteria. On one hand, there are certain purely descriptive features. On the other, there is a deeper value. Both of these criteria seem to play a role in people's application of the concept. (For different approaches to understanding these different criteria and their interrelations, see Leslie, 2015; Del Pinal & Reuter, forthcoming; Buckwalter, Rose & Turri, 2015.)

One possibility is that the pattern of judgments we have observed for the concept of happiness can be understood as just one instance of this far more general phenomenon. People associate the concept of happiness with certain purely descriptive features (positive affect, lack of negative affect, life satisfaction), but they also seem to associate the concept with certain deeper values. When an agent displays all of the descriptive features but fails to embody the deeper values, experimental participants tend to show some reluctance to say that the agent is truly happy.

If this approach is broadly on the right track, a key task for future research will be to say something more about which specific deeper values people associate with the concept of happiness. Given the results of Experiment 3, where people judged an evil agent to be happier than a slightly immoral agent, it cannot be that the only value that is relevant to happiness is simply a matter of living a life with the lowest possible degree of moral wrongness. Thus, the relevant value must extend beyond a simple measure of moral goodness vs. badness, even if it is

related to morality in some way or another. To illustrate with just one example, it could be that the relevant value is a matter of having a life that is truly *meaningful* (for more on meaningfulness, see, Wolf, 2010). In investigating this question, future research could also draw on extant evaluative theories of happiness, which do involve moral considerations but are not simply a matter of minimizing the total number of immoral actions (Aristotle, 340 BCE/2002; Foot, 2001; Kraut, 1979).

Further research could explore this issue on two different levels. On one hand, such research could look to the signature properties of dual character concepts more generally and ask whether the concept of happiness shows these properties. On the other, it could try to work out a more detailed dual character account of the concept of happiness in particular and then put that account to the test directly.

Implications for the Science of Happiness

Lastly, it is important to consider how the present results may bear on the scientific study of happiness. In light of these results, one response would be to conclude that scientists need to revise their definition of happiness, so that it is in line with the ordinary concept. However, we think this would be a mistake. It is not hard to see why when one begins to consider other analogous cases where technical definitions diverge from ordinary concepts. Consider, for example, the technical understanding of *preference* in economics. Now suppose that studies showed that the ordinary folk concept of preference diverges from this technical definition. Clearly, it would be a mistake for economists to revise their definition in light of these results, and doing so would disrupt the progress that has been made in economics. Similarly, we think it would be a mistake for happiness researchers to revise their technical definition of happiness in light of the present results (e.g., by including moral evaluations of the agent's life). In both cases,

the technical definitions play an important theoretical role within these disciplines and need not be aligned with the ordinary concepts.

At the same time, the present results may bear on the scientific research on happiness in an importantly different way. A number of recent scientific discoveries about happiness have been regarded as highly counterintuitive or even confusing by non-specialists. For example, a number of studies have suggested that having children does not lead people to be happier (Marini, 1980), and may actually decrease their happiness (Evenson & Simon, 2005; Glenn & McLanahan, 1982; Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004; LeMasters, 1957). These findings have been met with some skepticism, provoking media coverage with titles like, "Children don't make you happy, says expert who doesn't have any" ("Children don't make you happy", 2009). To take another example, recent research has also suggested that happiness is positively related to traits like selfishness (Tan & Forgas, 2010), a lack of empathy (Devlin, Zaki, Ong & Gruber, 2014), and risky behavior like alcohol consumption, binge eating and drug use (Cyders & Smith, 2008; Martin et al., 2002). Non-specialists have similarly found these discoveries highly counterintuitive, as is well illustrated by the title one article: "Happiness and selfishness: A paradox" (Vedantam, 2010).

The present studies allow us to understand why non-specialists may have this sort of a reaction. If these findings are interpreted in terms of the folk concept of happiness, then they are in fact highly surprising. After all, the presence of traits like selfishness or a lack of empathy would suggest that one would be unlikely to be experiencing happiness as characterized in the folk concept. However, when one instead understands that these findings are actually about happiness as it has been defined by psychologists (a combination of high positive affect, low negative affect, and high life-satisfaction), these findings cease to be as counterintuitive. In this

39

way, the present studies may help to clarify how findings such as these can be most effectively communicated to non-specialists.

Directions for Future Research

The present studies point to two separate avenues for future research. One way to extend the current findings would be to pursue research that continues to focus on the ordinary concept of happiness and furthers our understanding of how people understand what it means to be happy. At the same time, our findings suggest that another fruitful avenue would be to further investigate what this effect can reveal about evaluative concepts and the influence of morality on non-moral judgments. While these two lines of research can be pursued separately, each should also deepen our understanding of the other.

References

Adams, F., & Steadman, A. (2004). Intentional action in ordinary language: Core concept or pragmatic understanding?. *Analysis*, 64(282), 173–181.

Alfano, M., Beebe, J. R., & Robinson, B. (2012). The centrality of belief and reflection in Knobe-effect cases: a unified account of the data. *The Monist*, 95(2), 264–289.

Aristotle. (2002/340 BCE). *Nicomachean ethics*. (S. Broadie & C. Rowe, Eds.).

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. http://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. doi:10.1016/j.jml.2012.11.001

Beebe, J. R., & Buckwalter, W. (2010). The epistemic side-effect effect. *Mind and Language*, 25(4), 474–498. http://doi.org/10.1111/j.1468-0017.2010.01398.x

Buckwalter, W., Rose, D., & Turri, J. (2015). Belief through Thick and Thin. *Noûs*, 49(4), 748–775. http://doi.org/10.1111/nous.12048

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1), 3–5. http://doi.org/10.1177/1745691610393980

Children don't make you happy, says expert who doesn't have any. (2009, March 20). Retrieved from http://www.telegraph.co.uk/news/health/news/5021035/Children-dont-make-you-

happy-says-expert-who-doesnt-have-any.html

Cooper, J. M. (1987). Contemplation and happiness: A reconsideration. *Synthese*, 72(2), 187–216. http://doi.org/10.1007/BF00413638

Crisp, R. (2008). Well-Being. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Retrieved from http://plato.stanford.edu/archives/win2008/entries/well-being/

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease*, 175(9), 526–36. http://doi.org/10.1097/00005053-198709000-00004

Cyders, M. A., & Smith, G. T. (2008). Emotion-based dispositions to rash action: Positive and negative urgency. *Psychological Bulletin*, 134(6), 807–828.

Del Pinal, G. & Reuter, K. (forthcoming). Dual character concepts in social cognition: Commitments and the normative dimension of conceptual representation. *Cognitive Science*.

Devlin, H. C., Zaki, J., Ong, D. C., & Gruber, J. (2014). Not as good as you think? Trait positive emotion is associated with increased self-reported empathy but decreased empathic performance. *PLOS One*, 9(10), e110470.

Diener, E. (2000). Subjective Well-Being: The Science of Happiness and a Proposal for a National Index. *American Psychologist* 55(1): 34–43. doi:10.1037//0003-066X.55.1.34.

Diener, E., Emmons, R. a, Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49(1), 71–75. http://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Scollon, C. N., & Lucas, R. E. (2004). The evolving concept of subjective well-being: The multifaceted nature of happiness. In P. T. Costa & I. C. Siegler (Eds.), *Advances in*

cell aging and gerontology: Vol. 15 (pp. 187-220). Amsterdam: Elsevier.

Ditto, P.H., Pizarro, D.A., & Tannenbaum, D. (2009). Motivated moral reasoning. In D.M. Bartels, C.W. Bauman, L.J. Skitka, & D.L. Medin (Eds.), *Moral judgment and decision making* (pp. 307- 338). San Diego, CA: Academic Press

Driver, J. (2008a). Attributions of causation and moral responsibility. In W. Sinnott-Amstrong (Ed.) *Moral psychology (Vol. 2):, The cognitive science of morality: Intuition and diversity* (pp. 423-439). Cambridge, MA: MIT Press.

Driver, J. (2008b). Kinds of norms and legal causation: Reply to Knobe and Fraser and Deigh, in W. Sinnott-Amstrong (Ed.) *Moral psychology (Vol. 2):, The cognitive science of morality: Intuition and diversity* (pp. 459-461). Cambridge, MA: MIT Press.

Evenson, R. J., & Simon, R. W. (2005). Clarifying the Relationship Between Parenthood and Depression. *Journal of Health and Social Behavior*, 46(4), 341–358. http://doi.org/10.1177/002214650504600403

Feldman, F. (2010). *What Is This Thing Called Happiness*? OUP Oxford.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226. http://doi.org/10.1037/0003-066X.56.3.218

Foot, P. (2001). *Natural Goodness*. New York, NY: Oxford University Press.

Glenn, N. D., & McLanahan, S.. (1982). Children and Marital Happiness: A Further Specification of the Relationship. *Journal of Marriage and Family*, 44(1), 63–72. http://doi.org/10.2307/351263

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gilbert, D. (2006). *Stumbling on Happiness*. New York, NY: Vintage Books.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies?: A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, 59(2), 93–104. http://doi.org/10.1037/0003-066X.59.2.93

Gruber, J., Mauss, I. B., & Tamir, M. (2011). A Dark Side of Happiness? How, When, and Why Happiness Is Not Always Good. *Perspectives on Psychological Science*, 6(3), 222–233. http://doi.org/10.1177/1745691611406927

Haybron, D. M. (2008). *The Pursuit of Unhappiness: The Elusive Psychology of Well-Being*. OUP Oxford.

Haybron, D. (2011). Happiness. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Retrieved from http://plato.stanford.edu/archives/fall2011/entries/happiness/.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience Sampling Method: Measuring the Quality of Everyday Life*. Thousand Oaks, CA: Sage Publications.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434–446. http://doi.org/10.1016/j.jml.2007.11.007

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: the day reconstruction method. *Science*, 306(5702), 1776–80. http://doi.org/10.1126/science.1103572

King, L. A., & Napa, C. K. (1998). What makes a life good?. *Journal of Personality and Social Psychology*, 75(1), 156–165.

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, 63(279), 190–194.

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect theory of mind and moral judgment. *Psychological Science*, 17(5), 421–427.

Knobe, J., Prasada, S., & Newman, G. E. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition*, 127(2), 242–57. http://doi.org/10.1016/j.cognition.2013.01.005

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition*, 137, 196-209.

Kraut, R. (1979). Two Conceptions of Happiness. *Philosophical Review*, 88(2), 167. http://doi.org/10.2307/2184505

LeMasters, E. E. (1957). Parenthood as Crisis. *Marriage and Family Living*, 19(4), 352–355.

Lerner, M. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.

Leslie, S.-J. (2015). "Hillary Clinton is the Only Man in the Obama Administration": Dual Character Concepts, Generics, and Gender. *Analytic Philosophy*, 56(2), 111–141. http://doi.org/10.1111/phib.12063

Lucas, R. E., Diener, E., & Larsen, R. J. (2003). Measuring positive emotions. In S. J. Lopez & C. R. Snyder (Eds.), *Positive psychological assessment: A handbook of models and measures* (pp. 201-218). Washington, DC: American Psychological Association.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: does happiness lead to success? *Psychological Bulletin*, 131(6), 803–55. http://doi.org/10.1037/0033-2909.131.6.803

Marini, M. M. (1980). Effects of the number and spacing of children on marital and parental satisfaction. *Demography*, 17(3), 225–42.

Martin, L. R., Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Criqui, M. H., & Schwartz, J. E. (2002). A Life Course Perspective on Childhood Cheerfulness and its Relation to Mortality Risk. *Personality and Social Psychology Bulletin*, 28(9), 1155–1165. http://doi.org/10.1177/01461672022812001

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. *Similarity and analogical reasoning*, 179, 195.

Newman, G., De Freitas, J. & Knobe, J. (2015). Beliefs about the True Self explain asymmetries based on moral judgment. *Cognitive Science*, 39, 96-125.

Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology*, 3(2), 137–152. http://doi.org/10.1080/17439760701756946

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language*, 24(5), 586-604.

Phillips, J., Luguri, J. B., & Knobe, J. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities. *Cognition*, 145, 30-42.

Phillips, J., Misenheimer, L., & Knobe, J. (2011). The Ordinary Concept of Happiness (and Others Like It). *Emotion Review*, 3(3), 320–322. http://doi.org/10.1177/1754073911402385

Phillips, J., Nyholm, S., & Liao, S. (2014). The Good in Happiness. In J. Knobe, T. Lombrozo, & S. Nichols (Eds.), *Oxford Studies in Experimental Philosophy* (Vol. 1). New York, NY: Oxford University Press.

Samland, J., Josephs, M., Waldmann, M. R., & Rakoczy, H. (2016). The role of prescriptive

    norms and knowledge in children's and adults' causal selection. *Journal of Experimental*

    Psychology: General, 145(2), 125–130.

Seligman, M. E. P. Csikszentmihalyi, M. (2000). Positive psychology. *American Psychologist*, 1,

    5–14.

Sidgwick, H. (1966/1907). *The methods of ethics*. New York: Dover Publications.

Tan, H. B., & Forgas, J. P. (2010). When happiness makes us selfish, but sadness makes us fair:

    Affective influences on interpersonal strategies in the dictator game. *Journal of*

    Experimental Social Psychology, 46(3), 571–576.

    doi:http://dx.doi.org/10.1016/j.jesp.2010.01.007

Tiberius, V. (2013). Well-being, Wisdom, and Thick Theorizing: on the Division of Labor

    between Moral Philosophy and Positive Psychology. In S. Kirchin (Ed.), *Thick Concepts*

    (pp. 217–256). OUP Oxford.

Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational

    explanation for the side-effect effect. *Cognition*, 116(1), 87–100.

Vendantam, S. (2010, March 27). Happiness and Selfishness: A Paradox [Web log post].

    Retrieved from https://www.psychologytoday.com/blog/the-hidden-

    brain/201003/happiness-and-selfishness-paradox

Watson, D., & Clark, L. A. (1994). The PANAS-X: Manual for the positive and negative affect

    schedule-expanded form. *The British Journal of Clinical Psychology the British*

    Psychological Society, 65, 836–851. http://doi.org/10.1037/0022-3514.65.4.836

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures

    of positive and negative affect: the PANAS scales. *Journal of Personality and Social*

Psychology, 54(6), 1063–1070. http://doi.org/10.1037/0022-3514.54.6.1063

Wolf, S. (2010). *Meaning in life and why it matters*. Princeton University Press.

Young, L. & Phillips, J. (2011). The paradox of moral focus. *Cognition*, 119(2), 166—187.

Zou, C., Schimmack, U., & Gere, J. (2013). The validity of well-being measures: A multiple-indicator–multiple-rater model. *Psychological Assessment*, 25(4), 1247–1254. http://doi.org/10.1037/a0033902

Figure Captions

Figure 1. (a) Schematic depiction in which moral evaluations affect assessments of happiness by giving rise to performance errors. (b) Schematic depiction in which morality indirectly affects assessments of happiness by changing the perception of descriptive psychological states. (c) Schematic depiction in which morality is directly relevant to the ordinary concept of happiness, independent of changes in the perception of descriptive psychological states.

Figure 2. Participants' agreement that an agent is happy as a function of the participants' level of expertise and the moral value of the agent's life, which was manipulated both between-subjects (left) and within-subjects (right). Error bars indicate +/- 1 SE.

Figure 3. Participants' agreement that agents living morally bad (dark bars) and morally good (light bars) lives are happy, for participants who were presented with research indicating that happiness is bad (left) or good (right). Error bars indicate +/- 1 SE.

Figure 4. Participants' assessments of the moral value of the agents' lives (left) and their happiness (right) as a function of the whether the agent was described as being slightly morally bad (dark bars) or being fundamentally evil (light bars). Error bars indicate +/- 1 SE.

Figure 5. Participants' happiness assessments (left) and the emotion assessment (right) as a function of the whether the agent was living a morally good (light bars) or morally bad (dark barks) life. Error bars indicate +/- 1 SE.

Figure 6. Participants' happiness assessments (left) and emotion assessments (right) as a function

of the whether the agent had positively or negatively valenced psychological states and whether the agent was living a morally good (light bars) or morally bad (dark barks) life. Error bars indicate +/- 1 SE.

Figure 7. Happiness assessments of agents living lives that were either morally good or bad (left panel) or non-morally good or bad (right panel) and for whom participants were either given information about the agent's positive psychological states (right-hand side of each panel) or not (right-hand side of each panel). Dark bars indicated that agent was described as having a bad life; light bars indicate the agent was described as having a good life. Error bars indicate +/- 1 SE.
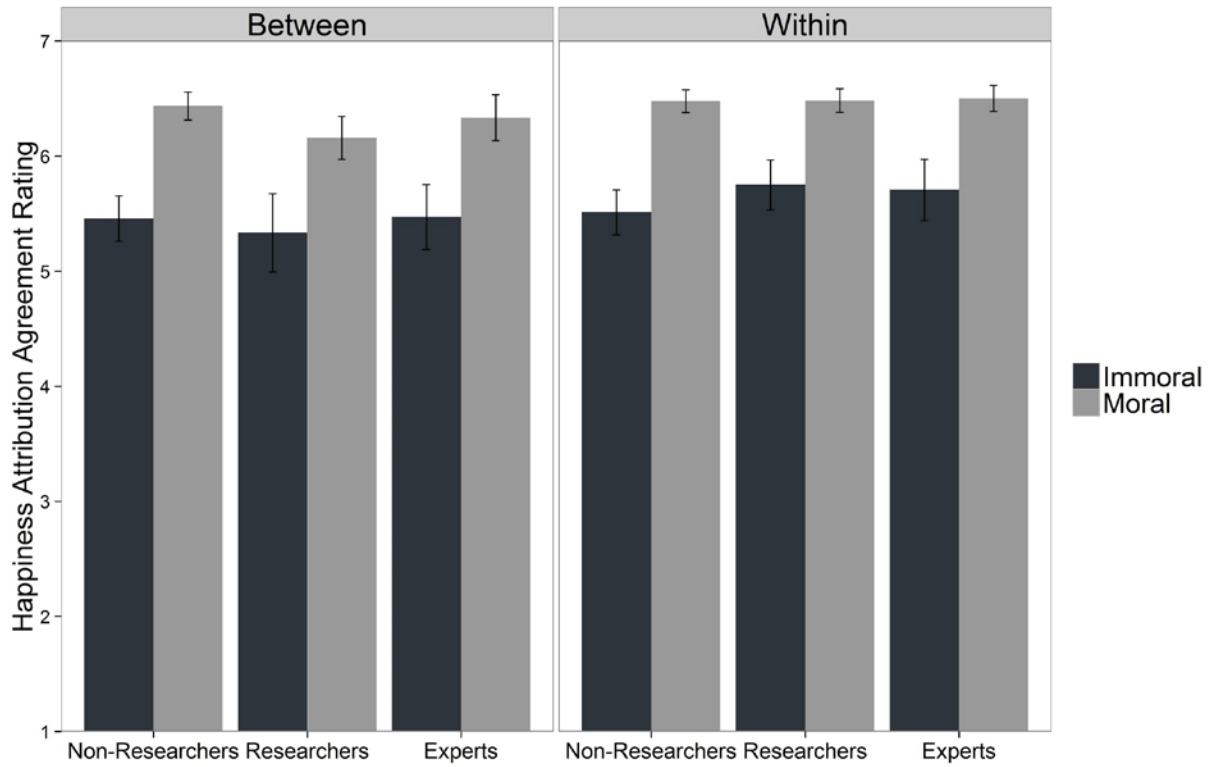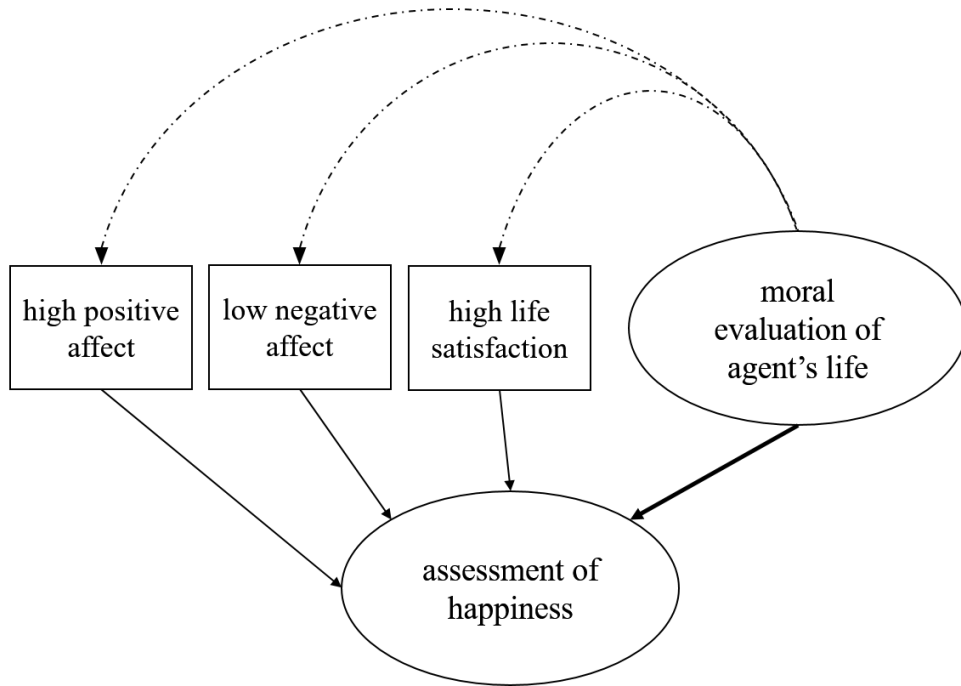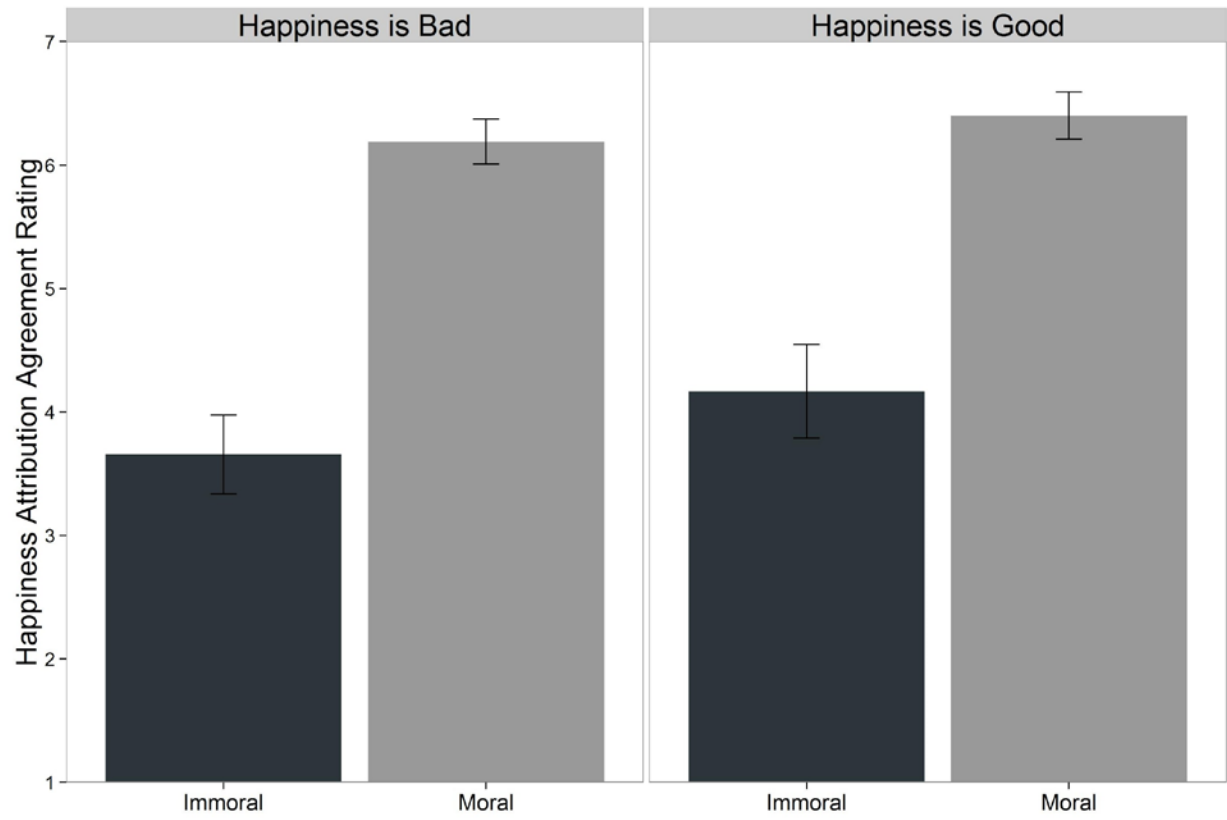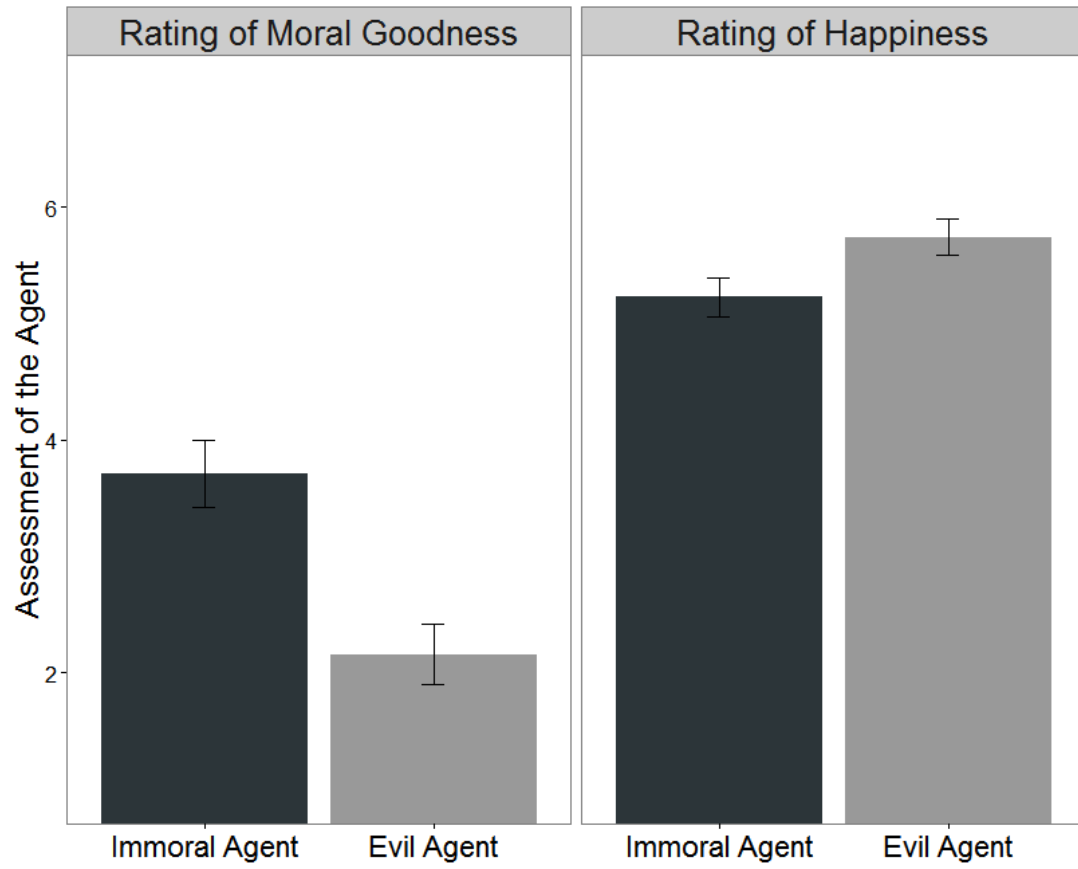
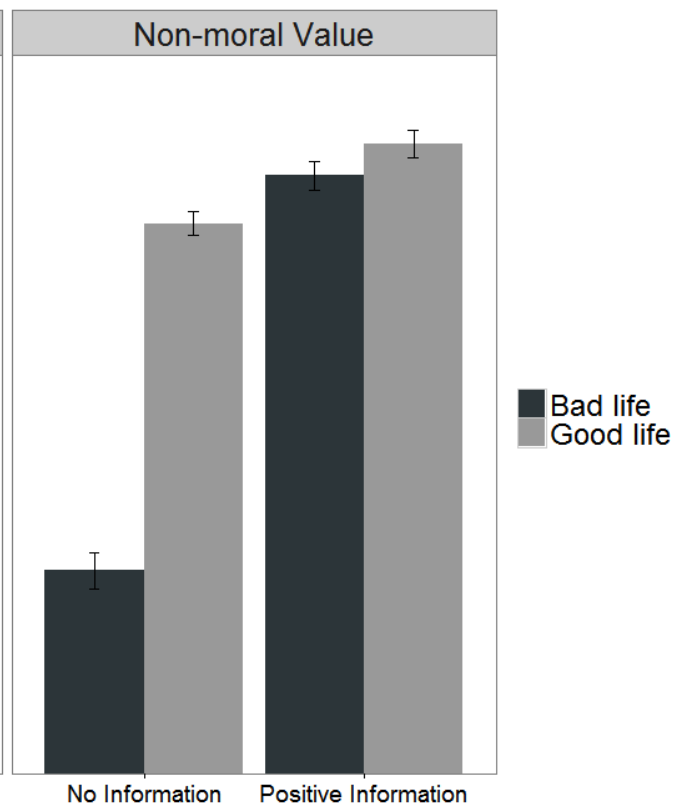a) **Performance error model**
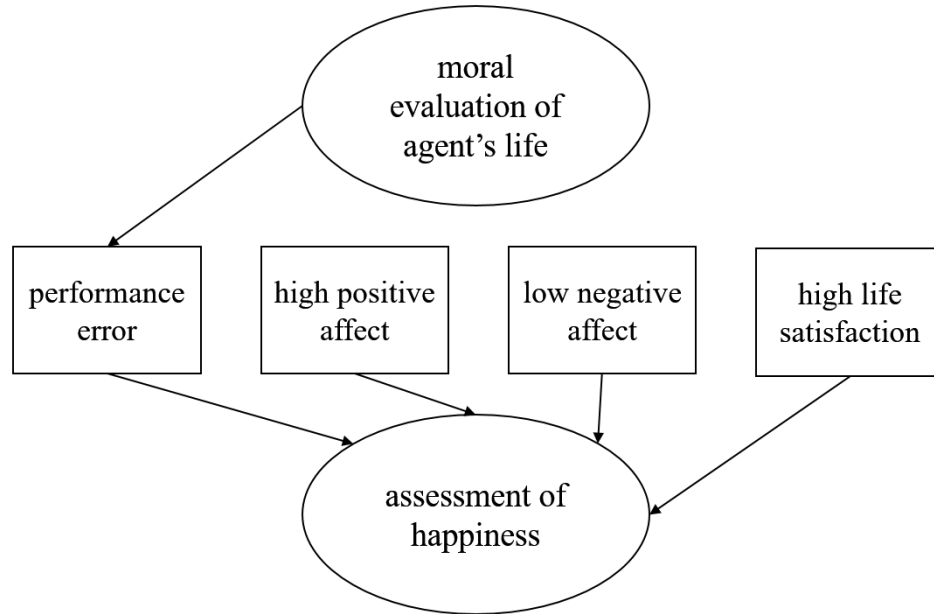
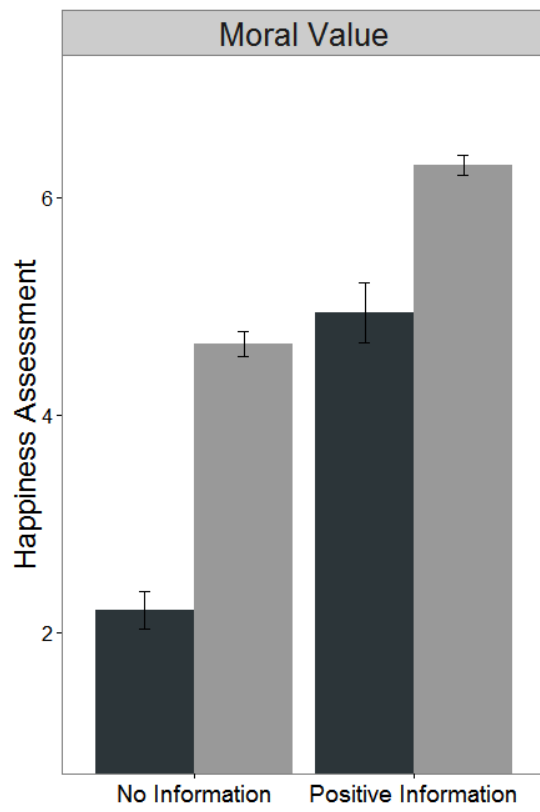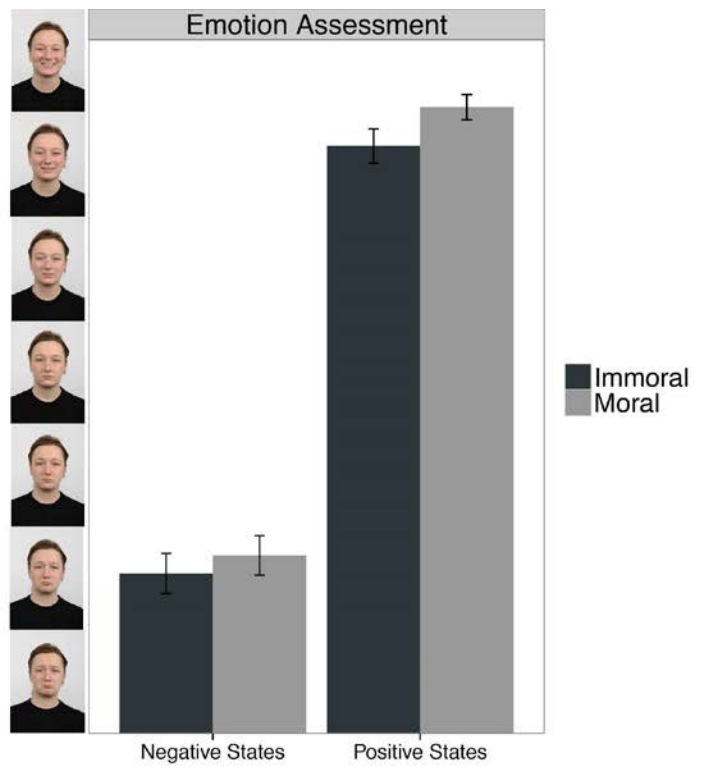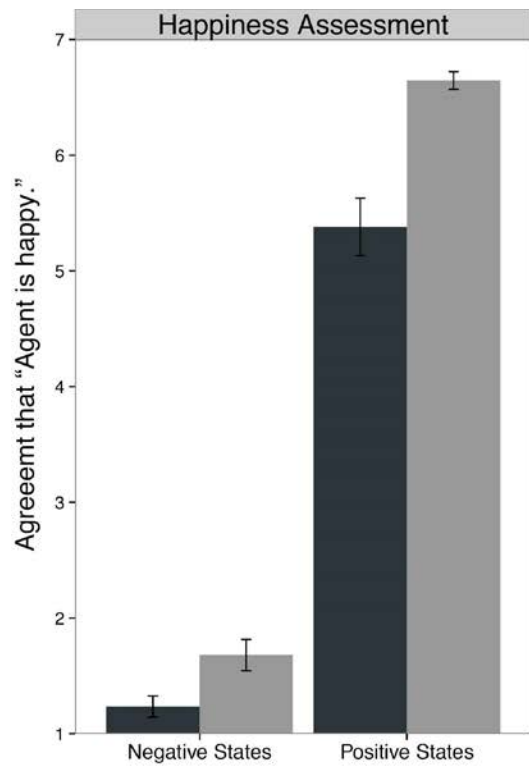
b) **Descriptive happiness model**
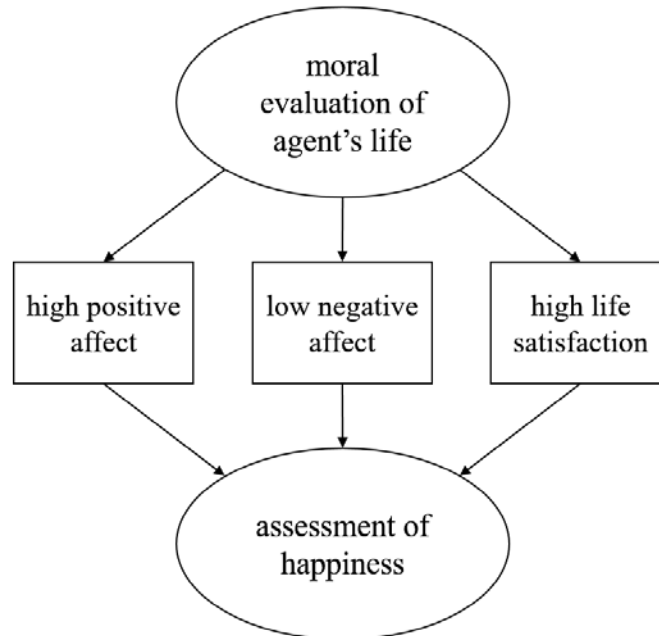
c)

Evaluative happiness model

Appendix

Vignettes from Study 1

<u>Agent 1: morally good condition</u>

Tom always enjoys his job as a janitor at a local community college. What he likes most about his job is how it gives him a chance to see his wife who is attending the community college. Almost every single day Tom feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever feel negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he feels highly satisfied with the way he lives.

The reason Tom feels this way is that every day he takes the handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Each night as he's going to sleep, he thinks about the children he will be helping the next day.

<u>Agent 1: morally bad condition</u>

Tom always enjoys his job as a janitor at a local community college. What he likes most about his job is how it gives him a chance to meet the young female students who are attending the community college. Almost every single day Tom feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever feel negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he feels highly satisfied with the way he lives.

The reason Tom feels this way is that every day he goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. Each night as he's going to sleep, he thinks about the things he will be stealing the next day.

<u>Agent 2: morally good condition</u>

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted. Almost every single day Sarah feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever feel negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she feels highly satisfied with the way she lives.

The reason Sarah feels this way is that she helps the sick children by giving them vitamins that taste like gummy bears. Sarah doesn't really know how many children have been helped by her, but she likes to think about it when she falls asleep at night.

<u>Agent 2: morally bad condition</u>

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted. Almost every single day Sarah feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever feel negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she feels highly satisfied with the way she lives.

The reason Sarah feels this way is that she poisons the sick children by giving them vitamins that have pesticides inside of them. Sarah doesn't really know how many children have died because of her, but she likes to think about it when she falls asleep at night.

Agent 3: morally good condition

Garrett is a forty-five year old man who lives in the suburbs of Atlanta with his family. One of his favorite things is when his niece comes over to play with his kids. Almost every single day Garrett feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever feel negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he feels highly satisfied with the way he lives.

The reason Garrett feels this way is that when his niece comes over to spend the night, Garrett reads to his niece. After that, sometimes they all talk about what they did that day and what they want to be when they grow up. Afterward, Garrett likes to think about what happened that night as he falls asleep

Agent 3: morally bad condition

Garrett is a forty-five year old man who lives in the suburbs of Atlanta with his family. One of his favorite things is when his niece comes over to play with his kids. Almost every single day Garrett feels good and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever feel negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he feels highly satisfied with the way he lives.

The reason Garrett feels this way is that when his niece comes over to spend the night, Garrett pressures her to have sex with him. After that, they all talk about what they did that day and what they want to be when they grow up. Afterward, Garrett likes to think about what happened that night as he falls asleep.

Vignettes from Study 2

<u>Morally good condition</u>

Maria is the mother of three children who all really love her. In fact, they couldn't imagine having a better mom. Maria usually stays pretty busy taking care of her children. She often finds herself rushing from one birthday party to the next, and is always going to pick up some groceries or buy school supplies. While Maria has been preoccupied with her children, she still makes an effort to be nice and spend time with her old friends. Almost every night she ends up helping her children or planning something for her children's future.

Day to day, Maria usually feels excited and really enjoys whatever she is doing. When she reflects on her life, she also feels great. She can't think of anything else in the world that she would want to spend her time doing and feels like the success she's had is definitely worth whatever sacrifices she has made.

<u>Morally bad condition</u>

Maria wants to live the life of a celebrity in L.A. In fact, she has even started trying to date a few famous people. Maria usually works hard to become popular. She often finds herself rushing from one social gathering to the next, and is always going to pick up some alcohol or a dress. Maria is so preoccupied with becoming popular that she is no longer concerned with being honest or nice to her old friends unless they know someone famous. Almost every night she ends up drinking or partying with famous people she wants to be like.

Day to day, Maria usually feels excited and really enjoys whatever she is doing. When she reflects on her life, she also feels great. She can't think of anything else in the world that she would want to spend her time doing and feels like the success she's had is definitely worth whatever sacrifices she has made.

Vignettes from Study 3

Morally bad condition

Maria is a mother of three children, who wants to live the life of a celebrity in L.A. In fact, she has even started trying to date a few famous people. Maria usually works hard to become popular. She often finds herself rushing from one social gathering to the next and is always going to pick up some alcohol or a dress.

Maria is completely preoccupied with becoming popular. She is no longer concerned with being honest or nice to her old friends unless they know someone famous. Almost every night she ends up drinking or partying with famous people she wants to be like.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Morally evil condition

Maria is part of a violent terrorist group called White Nation. From an extremely young age, she has believed that only white people are good and that all black people should be killed or locked away in prisons. Every day, she is busy working with the other members of her group to make sure this happens. They are currently plotting to blow up a number of civil rights museums and black churches around the country.

Maria is completely preoccupied with her goal, and only talks with white people who share her beliefs. Almost every night, she ends up talking with the other white supremacists about how to take back their country and eliminate or imprison everyone who is not white.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Vignettes from Study 4a

Agent 1: Morally good condition with positive psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Sarah stays incredibly busy all day helping the many sick children who depend on her for their medication. She often wonders whether she will be able to help all of them. Every day she carefully measures how much medicine she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been helped by her, but she thinks about it when she falls asleep at night.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 1: Morally bad condition with positive psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Sarah stays incredibly busy all day poisoning the many sick children who depend on her for their medication. She often wonders whether she will be able to poison all of them. Every day she carefully measures how much poison she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been killed by her, but she thinks about it when she falls asleep at night.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 2: Morally good condition with positive psychological states

Maria is the mother of three children who all really love her. In fact, they couldn't imagine having a better mom. Maria usually stays pretty busy taking care of her children. She often finds herself rushing from one birthday party to the next, visiting friends important to her children.

While Maria has been preoccupied with her children, she still makes an effort to be nice and spend time with her old friends. Almost every night she ends up helping her children or planning something for her children's future.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or

loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

<u>Agent 2: Morally bad condition with positive psychological states</u>

Maria is a mother of three children, who wants to live the life of a celebrity in L.A. In fact, she has even started trying to date a few famous people. Maria usually works hard to become popular. She often finds herself rushing from one social gathering to the next and is always going to pick up some alcohol or a dress.

Maria is completely preoccupied with becoming popular. She is no longer concerned with being honest or nice to her old friends unless they know someone famous. Almost every night she ends up drinking or partying with famous people she wants to be like.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

<u>Agent 3: Morally good condition with positive psychological states</u>

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Every day Tom takes the autistic and handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Tom always wonders whether he could be doing more to help these students get through their day without struggling, and to fit in with their peers. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

<u>Agent 3: Morally bad condition with positive psychological states</u>

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Every day after school Tom goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. He occasionally puts stolen items in black students' lockers so that they're blamed for the thefts. Tom always wonders whether he could be stealing more. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 4: Morally good condition with positive psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, she asks him to read long books to her late at night so that she can fall asleep. Garrett always wonders if he could be doing more and debates what kind of effect his actions will have on his niece's life. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 4: Morally bad condition with positive psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, he pressures her to have sex with him late at night. Garrett always wonders what kind of effect his action will have on his niece's life, but then reminds himself that he's the only reason she's not homeless. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Vignettes from Study 4b

Agent 1: Morally good condition with positive psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

The reason Sarah feels this way is that she stays incredibly busy all day helping the many sick children who depend on her for their medication. She often wonders whether she will be able to help all of them. Every day she carefully measures how much medicine she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been helped by her, but she thinks about it when she falls asleep at night.

Agent 1: Morally bad condition with positive psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

The reason Sarah feels this way is that she stays incredibly busy all day poisoning the many sick children who depend on her for their medication. She often wonders whether she will be able to poison all of them. Every day she carefully measures how much poison she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been killed by her, but she thinks about it when she falls asleep at night.

Agent 1: Morally good condition with negative psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Almost every single day Sarah is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that she would ever experience positive emotions like joy or excitement. When Sarah thinks about her life, she always comes to the same conclusion: she is highly unsatisfied with the way she lives.

The reason Sarah feels this way is that she stays incredibly busy all day helping the many sick children who depend on her for their medication. She often wonders whether she will be able to

help all of them. Every day she carefully measures how much medicine she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been helped by her, but she thinks about it when she falls asleep at night.

Agent 1: Morally bad condition with negative psychological states

After going to nursing school for several years, Sarah got a job at the children's hospital and sees many different children each day. This is the job she has always wanted.

Almost every single day Sarah is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that she would ever experience positive emotions like joy or excitement. When Sarah thinks about her life, she always comes to the same conclusion: she is highly unsatisfied with the way she lives.

The reason Sarah feels this way is that she stays incredibly busy all day poisoning the many sick children who depend on her for their medication. She often wonders if she will poison all of them. Every day she carefully measures how much poison she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been killed by her, but she thinks about it when she falls asleep at night.

Agent 2: Morally good condition with positive psychological states

Maria is the mother of three children who all really love her. In fact, they couldn't imagine having a better mom. Maria usually stays pretty busy taking care of her children. She often finds herself rushing from one birthday party to the next, visiting friends important to her children.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

The reason Maria feels this way is that while Maria has been preoccupied with her children, she still makes an effort to be nice and spend time with her old friends. Almost every night she ends up helping her children or planning something for her children's future.

Agent 2: Morally bad condition with positive psychological states

Maria wants to live the life of a celebrity in L.A. In fact, she has even started trying to date a few famous people. Maria usually works hard to become popular. She often finds herself rushing from one social gathering to the next and is always going to pick up some alcohol or a dress.

Almost every single day Maria is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Maria thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

The reason Maria feels this way is that Maria is so preoccupied with becoming popular that she is no longer concerned with being honest or nice to her old friends unless they know someone famous. Almost every night she ends up drinking or partying with famous people she wants to be like.

<u>Agent 2: Morally good condition with negative psychological states</u>

Maria is the mother of three children who all really love her. In fact, they couldn't imagine having a better mom. Maria usually stays pretty busy taking care of her children. She often finds herself rushing from one birthday party to the next, visiting friends important to her children.

Almost every single day Maria is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that she would ever experience positive emotions like joy or excitement. When Maria thinks about her life, she always comes to the same conclusion: she is highly unsatisfied with the way she lives.

The reason Maria feels this way is that while Maria has been preoccupied with her children, she still makes an effort to be nice and spend time with her old friends. Almost every night she ends up helping her children or planning something for her children's future.

<u>Agent 2: Morally bad condition with negative psychological states</u>

Maria wants to live the life of a celebrity in L.A. In fact, she has even started trying to date a few famous people. Maria usually works hard to become popular. She often finds herself rushing from one social gathering to the next and is always going to pick up some alcohol or a dress.

Almost every single day Maria is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that she would ever experience positive emotions like joy or excitement. When Maria thinks about her life, she always comes to the same conclusion: she is highly unsatisfied with the way she lives.

The reason Maria feels this way is that Maria is so preoccupied with becoming popular that she is no longer concerned with being honest or nice to her old friends unless they know someone famous. Almost every night she ends up drinking or partying with famous people she wants to be like.

<u>Agent 3: Morally good condition with positive psychological states</u>

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

The reason Tom feels this way is that every day Tom takes the autistic and handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Tom always wonders whether he could be doing more to help these students get through their day without struggling, and to fit in with their peers. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Agent 3: Morally bad condition with positive psychological states

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

The reason Tom feels this way is that every day after school Tom goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. He occasionally puts stolen items in black students' lockers so that they're blamed for the thefts. Tom always wonders whether he could be stealing more. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Agent 3: Morally good condition with negative psychological states

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Almost every single day Tom is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that he would ever experience positive emotions like joy or excitement. When Tom thinks about his life, he always comes to the same conclusion: he is highly unsatisfied with the way he lives.

The reason Tom feels this way is that every day Tom takes the autistic and handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Tom always wonders whether he could be doing more to help these students get through their day without struggling, and to fit in with their peers. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Agent 3: Morally bad condition with negative psychological states

Tom is a janitor at a local community college. His job is very busy, but he meets many different people attending the college.

Almost every single day Tom is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that he would ever experience positive emotions like joy or excitement. When Tom thinks about his life, he always comes to the same conclusion: he is highly unsatisfied with the way he lives.

The reason Tom feels this way is that every day after school Tom goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. He occasionally puts stolen items in black students' lockers so that they're blamed for the thefts. Tom always wonders whether he could be stealing more. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the college.

Agent 4: Morally good condition with positive psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

The reason Garret feels this way is that a few years ago, his niece moved in with him after her parents passed away. Almost every night, she asks him to read long books to her late at night so that she can fall asleep. Garrett always wonders if he could be doing more and debates what kind of effect his actions will have on his niece's life. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Agent 4: Morally bad condition with positive psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

The reason Garret feels this way is that a few years ago, his niece moved in with him after her parents passed away. Almost every night, he pressures her to have sex with him late at night. Garrett always wonders what kind of effect his action will have on his niece's life, but then reminds himself that he's the only reason she's not homeless. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Agent 4: Morally good condition with negative psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

Almost every single day Garrett is in a bad mood and generally experiences a lot of unpleasant

emotions. In fact, it is very rare that he would ever experience positive emotions like joy or excitement. When Garrett thinks about his life, he always comes to the same conclusion: he is highly unsatisfied with the way he lives.

The reason Garret feels this way is that a few years ago, his niece moved in with him after her parents passed away. Almost every night, she asks him to read long books to her late at night so that she can fall asleep. Garrett always wonders if he could be doing more and debates what kind of effect his actions will have on his niece's life. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Agent 4: Morally bad condition with negative psychological states

Garrett is a forty-five year old man who lives in the suburbs of Atlanta.

Almost every single day Garrett is in a bad mood and generally experiences a lot of unpleasant emotions. In fact, it is very rare that he would ever experience positive emotions like joy or excitement. When Garrett thinks about his life, he always comes to the same conclusion: he is highly unsatisfied with the way he lives.

The reason Garret feels this way is that a few years ago, his niece moved in with him after her parents passed away. Almost every night, he pressures her to have sex with him late at night. Garrett always wonders what kind of effect his action will have on his niece's life, but then reminds himself that he's the only reason she's not homeless. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Vignettes from Study 5

Agent 1: Non-morally good condition – No psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah has become very wealthy after things went unexpectedly well when she invested the money she earned from her new job. Not that long ago, Sarah did not have the money to take vacations, go to restaurants or buy nice clothes. She, also had to live in a small shared apartment with people didn't know before moving in. Now that Sarah has become wealthy, she can take vacations whenever she wants, regularly eats at her favorite restaurants and has an entirely new wardrobe. She also now lives in a beautiful house in a nice part of town.

Agent 1: Non-morally bad condition – No psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah has become very poor after things went unexpectedly badly when she invested the money she earned from her new job. Not that long ago, Sarah had the money to take vacations, go to restaurants and buy nice clothes. She, also got to live in a beautiful house in a nice part of town. Now that Sarah has become poor, she can no longer take vacations, never eats at her favorite restaurants and has an entirely old wardrobe. She also now lives in a small shared apartment with people she didn't know before moving in.

Agent 1: Morally good condition – No psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah stays incredibly busy all day helping the many sick children who depend on her for their medication. She often wonders whether she will be able to help all of them. Every day she carefully measures how much medicine she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been helped by her, but she thinks about it when she falls asleep at night.

Agent 1: Morally bad condition – No psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah stays incredibly busy all day poisoning the many sick children who depend on her for their medication. She often wonders whether she will be able to poison all of them. Every day she carefully measures how much poison she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been killed by her, but she thinks about it when she falls asleep at night.

Agent 1: Non-morally good condition – Psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah has become very wealthy after things went unexpectedly well when she invested the money she earned from her new job. Not that long ago, Sarah did not have the money to take vacations, go to restaurants or buy nice clothes. She, also had to live in a small shared apartment with people didn't know before moving in. Now that Sarah has become wealthy, she can take vacations whenever she wants, regularly eats at her favorite restaurants and has an entirely new wardrobe. She also now lives in a beautiful house in a nice part of town.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 1: Non-morally bad condition – Psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah has become very poor after things went unexpectedly badly when she invested the money she earned from her new job. Not that long ago, Sarah had the money to take vacations, go to restaurants and buy nice clothes. She, also got to live in a beautiful house in a nice part of town. Now that Sarah has become poor, she can no longer take vacations, never eats at her favorite restaurants and has an entirely old wardrobe. She also now lives in a small shared apartment with people she didn't know before moving in.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 1: Morally good condition – Psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah stays incredibly busy all day helping the many sick children who depend on her for their medication. She often wonders whether she will be able to help all of them. Every day she carefully measures how much medicine she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been helped by her, but she thinks about it when she falls asleep at night.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 1: Morally bad condition – Psychological Info

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah stays incredibly busy all day poisoning the many sick children who depend on her for their medication. She often wonders whether she will be able to poison all of them. Every day she carefully measures how much poison she will give each child, and then watches to make sure the child swallows it. Sarah doesn't really know how many children have been killed by her, but she thinks about it when she falls asleep at night.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

Agent 2: Non-morally good condition – No psychological Info

Tom is a janitor at a local school in small town near to where he lives.

Tom has an incredibly easy and engaging job. Each day, Tom spends his hours at work walking around the school. He has a lot of different jobs that he can help out with, so he gets to choose whatever he wants to do each day. This means that Tom gets to learn about different skills, and talk to many different people throughout the day. When the weather is nice, Tom spends the day outside and works on the grounds around the school. Other days, Tom helps out in the school kitchen or cleans the halls while he talks with students.

Agent 2: Non-morally bad condition – No psychological Info

Tom is a janitor at a local school in small town near to where he lives.

Tom has an incredibly hard and tedious job. Each day, Tom spends his at hours at work inside bathrooms or gym locker rooms at the school. He is responsible for cleaning toilets and scrubbing old showers which are often covered in mold and stains that are very difficult to remove. The chemicals he uses to clean are extremely strong, so they can easily cause rashes and always smell terrible. If Tom finishes cleaning the toilets and showers, his next job is to scrub the inside of the lockers, which are small and hard to reach into, but are also extremely dirty.

Agent 2: Morally good condition – No psychological Info

Tom is a janitor at a local school in small town near to where he lives.

Every day, Tom takes the autistic and handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Tom always wonders whether he could be doing more to help these students get through their day without struggling, and to fit in with their peers. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the school.

<u>Agent 2: Morally bad condition – No psychological Info</u>

Tom is a janitor at a local school in small town near to where he lives.

Every day after school Tom goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. He occasionally puts stolen items in black students' lockers so that they're blamed for the thefts. Tom always wonders whether he could be stealing more. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the school.

<u>Agent 2: Non-morally good condition – Psychological Info</u>

Tom is a janitor at a local school in small town near to where he lives.

Tom has an incredibly easy and engaging job. Each day, Tom spends his hours at work walking around the school. He has a lot of different jobs that he can help out with, so he gets to choose whatever he wants to do each day. This means that Tom gets to learn about different skills, and talk to many different people throughout the day. When the weather is nice, Tom spends the day outside and works on the grounds around the school. Other days, Tom helps out in the school kitchen or cleans the halls while he talks with students.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

<u>Agent 2: Non-morally bad condition – Psychological Info</u>

After going to nursing school for several years, Sarah got a job at the children's hospital.

Sarah has become very poor after things went unexpectedly badly when she invested the money she earned from her new job. Not that long ago, Sarah had the money to take vacations, go to restaurants and buy nice clothes. She, also got to live in a beautiful house in a nice part of town. Now that Sarah has become poor, she can no longer take vacations, never eats at her favorite restaurants and has an entirely old wardrobe. She also now lives in a small shared apartment with people she didn't know before moving in.

Almost every single day Sarah is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that she would ever experience negative emotions like sadness or loneliness. When Sarah thinks about her life, she always comes to the same conclusion: she is highly satisfied with the way she lives.

<u>Agent 2: Morally good condition – Psychological Info</u>

Tom is a janitor at a local school in small town near to where he lives.

Every day, Tom takes the autistic and handicapped students from class to class and talks with them about how they are doing and whether they are experiencing any problems. Tom always wonders whether he could be doing more to help these students get through their day without struggling, and to fit in with their peers. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the school.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 2: Morally bad condition – Psychological Info

Tom is a janitor at a local school in small town near to where he lives.

Every day after school Tom goes from locker to locker and steals belongings from the students and re-sells these belongings to buy himself alcohol. He occasionally puts stolen items in black students' lockers so that they're blamed for the thefts. Tom always wonders whether he could be stealing more. Occasionally, Tom hears stories about how hard it is for these students, who end up getting bullied or having to leave the school.

Almost every single day Tom is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Tom thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 3: Non-morally good condition – No psychological Info

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

The area that Garrett now lives in is incredibly convenient to get to and close to everything including excellent restaurants and beautiful parks. Whenever Garrett wants to get anywhere, even to a sports arena, it takes him just a few minutes to get to the subway that connects him to all of the other parts of town. Because of this, Garrett doesn't spend much time at home, and often goes out to eat at new restaurants or see some of the many shows that come through his town.

Agent 3: Non-morally bad condition – No psychological Info

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

The area that Garrett now lives in is incredibly inconvenient to get to and far away from everything other than a few houses and some farmers' fields. Whenever Garrett wants to get anywhere, even just to a grocery store, it takes him at least a forty-five minutes to get to the

highway that connects him to areas that have stores. Because of this, Garrett rarely ever leaves his small house, and doesn't do much other than sometimes walk along the roads nearby or watch the cows.

<u>Agent 3: Morally good condition – No psychological Info</u>

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, she asks him to read long books to her late at night so that she can fall asleep. Garrett always wonders if he could be doing more and debates what kind of effect his actions will have on his niece's life. Often, they also talk about what she did that day and the feelings she's had since her parents died.

<u>Agent 3: Morally bad condition – No psychological Info</u>

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, he pressures her to have sex with him late at night. Garrett always wonders what kind of effect his action will have on his niece's life, but then reminds himself that he's the only reason she's not homeless. Often, they also talk about what she did that day and the feelings she's had since her parents died.

<u>Agent 3: Non-morally good condition – Psychological Info</u>

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

The area that Garrett now lives in is incredibly convenient to get to and close to everything including excellent restaurants and beautiful parks. Whenever Garrett wants to get anywhere, even to a sports arena, it takes him just a few minutes to get to the subway that connects him to all of the other parts of town. Because of this, Garrett doesn't spend much time at home, and often goes out to eat at new restaurants or see some of the many shows that come through his town.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

<u>Agent 3: Non-morally bad condition – Psychological Info</u>

Garrett is a forty-five year old man, who moved away from where he grew up after finishing

high school.

The area that Garrett now lives in is incredibly inconvenient to get to and far away from everything other than a few houses and some farmers' fields. Whenever Garrett wants to get anywhere, even just to a grocery store, it takes him at least a forty-five minutes to get to the highway that connects him to areas that have stores. Because of this, Garrett rarely ever leaves his small house, and doesn't do much other than sometimes walk along the roads nearby or watch the cows.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 3: Morally good condition – Psychological Info

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, she asks him to read long books to her late at night so that she can fall asleep. Garrett always wonders if he could be doing more and debates what kind of effect his actions will have on his niece's life. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.

Agent 3: Morally bad condition – Psychological Info

Garrett is a forty-five year old man, who moved away from where he grew up after finishing high school.

A few years ago, his niece moved in with him after her parents passed away. Almost every night, he pressures her to have sex with him late at night. Garrett always wonders what kind of effect his action will have on his niece's life, but then reminds himself that he's the only reason she's not homeless. Often, they also talk about what she did that day and the feelings she's had since her parents died.

Almost every single day Garrett is in a great mood and generally experiences a lot of pleasant emotions. In fact, it is very rare that he would ever experience negative emotions like sadness or loneliness. When Garrett thinks about his life, he always comes to the same conclusion: he is highly satisfied with the way he lives.