This is a post-peer-review, accepted version of an article published in *Mind and Language*. The final version is available online at: https://doi.org/10.1111/mila.12279

Feeling the right way: Normative influences on people's use of emotion concepts

Rodrigo Díaz, Institute of Philosophy, University of Bern

Kevin Reuter, Institute of Philosophy, University of Zurich

Abstract

It is generally assumed that emotion concepts are purely descriptive. However, recent investigations suggest that the concept of happiness includes information about the morality of the agent's life. In this paper, we argue that normative influences on emotion concepts are not restricted to happiness and are not about moral norms. In a series of studies, we show that emotion attribution is influenced by whether the agent's psychological and bodily states fit the situation in which they are experienced. People consider that emotions are not just about feeling in certain ways, but also about feeling the *right* way.

KEYWORDS

Emotion, happiness, morality, fittingness, value, theory of mind

1. INTRODUCTION

There is a strand in our commonsense view of the emotions according to which we are at the mercy of these "irrational forces" (Lutz & White, 1986; Solomon, 1977). However, our ordinary language practices suggest that we consider emotions to be open to normative assessment. Expressions of the form "[Agent] should (not) be [emotion term]" are common,[1] and we

[1] The Corpus of Contemporary American English (COCA) lists 241 hits for the phrase "should be happy", for example, "You should be happy that your daughter is getting married". Comparisons with phrases like "should be willing" (168 hits) and "should be going" (242 hits) indicate its rather frequent use in everyday discourse.

routinely talk about emotions as right or wrong, good or bad, appropriate or inappropriate.[2] The contemporary philosophical literature on the topic (D'Arms & Jacobson, 2000; de Sousa, 2014; Deonna & Teroni, 2012) has established three main normative standards that are used to assess emotions: Moral (e.g., "You shouldn't be envious of your friend because envy is a vice"); prudential (e.g., "You shouldn't be envious of your friend because he might stop being generous if he realizes your envy"); and fittingness or accuracy standards (e.g., "You shouldn't be envious of your friend because he actually is not very wealthy.").

Most scholars have thought that such normative assessments need to follow emotion attribution: It seems that only when we have determined which emotional state a person is in, can we criticize or approve of the emotion. In this paper, we show that normative aspects, fittingness in particular, already play a fundamental role in identifying and attributing emotions. The study of normative influences on people's use of mental state concepts is not new. Previous research has shown that the use of concepts such as deciding, intending, or choosing are shaped by normative considerations (Pettit & Knobe, 2009). More recently, Phillips et al. (2011, 2014, 2017) have argued that even the concept of happiness is in part normative. Here, we will extend and challenge these latest findings by arguing and presenting new evidence suggesting that people attribute happiness *and other emotions* based on whether they *fit* the situation in which they are experienced.

The structure of this paper is as follows. In Section 2, we elaborate in detail what we mean when we say that emotion concepts are normatively shaped. We then present two competing models that aim to account for the influence of normative factors in emotion concepts, the "moral happiness model" and our own account, the "fittingness model". In Section 3, we report the results from two studies that support our model, and in Section 4 present two more studies responding to objections to our claims. Finally, we discuss some of the implications of our results for the philosophy and psychology of emotion.

2. NORMATIVE CONSIDERATIONS AND HOW TO MODEL THEIR IMPACT

2.1 The descriptive and the normative in concepts

[2] The susceptibility of emotions to normative assessments is, of course, inherently tied to normative assessments of people having those emotions. Thus, in virtue of an emotion being appropriate or inappropriate, we can say that an agent should or should not be in that emotional state.

The philosophical literature has discussed at least three different kinds of ways in which concepts may not be purely descriptive. First, there is the class of thick concepts. Thick concepts not only encode descriptive information, but are also normatively laden. For example, *generous* and *thriftless* share the same descriptive content of spending money or other goods in a lavish manner. However, they differ in their normative charge: A person *should be* generous but *should not* be thriftless. Most philosophers agree that the normative aspect of thick concepts evaluates the descriptive content but has no independent normative content itself (Kirchin, 2010; Putnam, 2002). Second, and in contrast, so called dual-character concepts have both a descriptive but also an *independent* normative dimension for categorization (Knobe et al., 2013; Del Pinal & Reuter, 2017; Reuter, 2019). For example, a person can be classified as a father if he fulfils certain descriptive criteria—being a male biological parent. Although we believe that a father *should* take care of his offspring, in its purely descriptive sense, that normative belief is irrelevant for his status as a father. There is, however, also a sense for which such normative information is crucial to whether we consider a man to be a father. Accordingly, a man who does not care about his child is not a (true) father, and vice versa, a man who looks after the child of another person, is in that normative sense a (true) father. Other dual-character concepts include *artist*, *self*, *colleague*, *rock music*, and many more. Third, apart from thick concepts and dual-character concepts, some central philosophical concepts have been argued to be *influenced* by normative considerations. Most famously, it has been shown that normative information has a strong impact on whether we think a person has acted intentionally (Knobe, 2003), or has been the cause of something (Hitchcock and Knobe, 2009). For example, if two people each take a pen from a table, but only one of them is allowed to do so, most people single out the person who *should* not have taken the pen as the cause of the lack of pens on the table.[3]

The structure and normative aspects of the concepts discussed above differ in important ways depending on whether the concept is thick, dual-character, or influenced by normative considerations.[4] In this paper, we do not take a strong stance on whether emotion concepts are constitutively normative or causally influenced by normative information. However, all three classes of concepts share a crucial feature that cannot be captured by purely descriptive information. Importantly, these concepts also have an action-guiding or reason-guiding dimension. Consequently, whether we believe a person to be thriftless depends on whether we

[3] While some people have argued that the concept of cause might be inherently normative (Sytsma *et al.*, 2019), others have merely argued that the application of the concept cause is influenced by normative information (Alicke, 1992).

[4] Note, however, that some studies (Chtuc, 2012) have argued that the concept of happiness belongs to the class a dual-character concepts and hence has a normative dimension for categorization.

believe that person *should* be less extravagant; whether we believe a man to be a father, partially depends on our belief that a man *should* look after his offspring; and whether we believe a person to be the cause of an event often depends on whether we believe that person *should* have acted differently. The main thesis of this paper is that emotion concepts are normative too. In other words, when we apply emotion concepts, normative considerations play a substantial role. But which norms influence the use of emotion concepts? In the following, we will introduce two accounts for the influence of normative considerations on the use of emotion concepts which differ precisely in this respect: The moral happiness and the fittingness model.

2.2. The moral happiness model

There is a long-standing philosophical debate around whether the nature of happiness is purely descriptive, or whether it includes moral considerations. On the one hand, *descriptive* views of happiness claim that happiness consists in the possession of certain psychological states, either high pleasure and low pain (*hedonist theories*; Feldman, 2010), being satisfied with one's life (*life satisfaction theories*; Suikkanen, 2011), or both (*hybrid theories*; Sumner, 1996). On the other hand, *normative* views defend that happiness is not only a matter of descriptive psychological states, but also about being a morally good person (Foot, 2001). On the latter view, an immoral person can never be truly happy, even if she feels good and is highly satisfied with her life.

In a series of recent studies, Phillips et al. (2011, 2014, 2017) showed that moral assessments influence people's use of the concept of happiness.[5] They presented participants in their studies with stories about a person who either lives a morally good or a morally bad life, and asked them to judge to what extent they thought that person was happy. In both cases, the person was described as experiencing the psychological states associated with a descriptive notion of happiness: High levels of positive affect, low levels of negative affect, and high life satisfaction. For example:

[5] It is important to note that the notion of happiness Phillips et al. are interested in is a very specific one, which involves a long-lasting disposition to experience certain episodes of emotions, and an appraisal which targets not a particular event, but one's life as a whole. In contrast, we think that our fittingness model could be applied both to occurrent episodes of emotion (like the fear cases of Study 1) and long-lasting emotional dispositions (like the happiness cases in Phillips et al.'s studies).

> 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.

According to a descriptive view of happiness, these psychological states are all that happiness is about. In order to determine whether participants also take into account normative considerations when attributing happiness, they designed two variants of the same vignette, which differed primarily in the moral value of the agent's life (moral vs. immoral). In the moral version of the vignette, the story continued with a morally laudable reason for Garrett to feel this way:

> The reason Garrett feels this way is that when his niece comes over to spend the night, Garrett reads to his niece.

In the immoral variant, Garrett feels good because of a totally immoral reason

> 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.

Participants reading the immoral version of the vignette rated Garrett to be significantly less happy than participants reading the Moral version.[6] Similar results were found across different studies using a large variety of vignettes, indicating the robustness of this effect.

[6] Participants rated their agreement with the statement "[Agent] is happy" on a scale from 1 (completely disagree) to 7 (completely agree). These ratings were higher in the "morally good" condition ($M \approx 6.5$) than in the "morally bad" condition ($M \approx 5.5$).

The authors interpreted their results as showing that the folk concept of happiness is not purely descriptive but also partially normative. In line with the ideas of philosophers such as Philippa Foot (2001), people seem to consider happiness to be not only about *feeling* good, but also about *being* good. We call this the "moral happiness model".

While we agree that Phillips and colleagues have shown the concept of happiness to be underpinned by normative aspects, we believe that this normative aspect does not primarily track the moral value of the agent's life, but rather the value of what she is happy about. Happiness is not about *feeling good* and *being good*, but about *feeling good about something good*. In other words, the normative standards that influence emotion attributions are not moral standards but fittingness standards.

2.3. The fittingness model

It is generally assumed that emotions are intentional states, in the sense that they are about something.[7] We are not simply afraid or sad, but rather afraid of darkness or sad that someone died. Thus, emotions have particular objects they are directed to. But emotions are also related to values. For example, fear is related to danger, and sadness is related to loss. These values are known as the emotions' formal objects (Teroni, 2007) or core-relational themes (S. Lazarus, 1991). We usually feel emotions towards particular objects that exemplify the relevant evaluative property: We are afraid of dangerous animals, and sad about things that constitute a loss for us. In these cases, we say that the emotion is *fitting*. But this is not always the case. For example, some people are afraid of butterflies, despite butterflies not being dangerous. In this case, fear is *unfitting*.

Fittingness standards thus track whether the emotion's particular object (e.g., a mouse) is an instance of the emotion's formal object (e.g., danger) (D'Arms and Jacobson, 2000). Depending on what one takes emotions to be, we can say that, e.g. bodily feelings or cognitive appraisals fit the situation.[8] But this assessment always ultimately refers to whether or not the relevant value can be reasonably ascribed to the target of the emotion.[9] Please note that many

[7] See Shargel, 2015 and Whiting, 2011 for arguments against this claim.

[8] In what follows, we will use the term "situation" to refer to the situation in which the emotion is experienced and which contains the particular object of the emotion.

[9] While many have accepted a connection between emotions and evaluative properties, the nature of this connection is at the heart of a fundamental debate in the philosophy of emotion (Todd, 2014). The main positions in this debate can be classified into two groups, which differ in the answer they give to Plato's long-standing Euthyphro question. Are we afraid (or happy, sad, amused …) of something because it is

theories of emotion (Prinz, 2004; Stephan et al., 2014; Feldman Barrett, 2012) emphasize the importance of the relationship between the psychological and bodily components of emotion and the situation in which those are experienced, although none of them talk explicitly about fittingness and its impact on emotion attribution.

Similar to what the moral happiness model postulated in relation to happiness, we defend a *normative* view of emotion concepts. But, instead of moral standards, we claim that emotion concepts are subject to fittingness standards. On this account, what we call the "fittingness model", when people attribute emotions like happiness or fear, they not only look for certain psychological and bodily states, but also take into account whether the situation exemplifies the relevant value. For example, being happy would not be only about feeling good and satisfied, but feeling good and satisfied towards something good. In other words, people consider that emotions are not just about feeling in certain ways, but also about feeling the *right* way. In the following, we present the results from a series of studies testing the fittingness model.

3. EMPIRICAL STUDIES INVESTIGATING THE NORMATIVE IMPACT ON EMOTION ATTRIBUTION

In this section, we present two studies**.** Study 1 demonstrates that fittingness standards influence the use of emotion concepts other than happiness; the results of Study 2 suggest that the role of moral considerations in the folk concept of happiness can be explained in terms of fittingness. These studies will test the main predictions of the fittingness model in two very different emotions: Fear (Study 1) and happiness (Study 2). Together, they constitute a first test for the adequacy and explanatory power of the fittingness model.

dangerous, or is it dangerous because we are afraid of it? On the one hand, many cognitivists and perceptual theories of emotion (but certainly not all of them) hold an objectivist idea of value in that they believe emotions to grasp (or fail to grasp) mind-independent evaluative properties. On the other hand, sentimentalist theories emphasize the primacy of emotion, and try to explain evaluative properties in terms of what elicit the relevant emotions in normal situations (dispositionalism), or what merits specific emotional responses (fitting-attitude analysis). While value objectivists explain the normative relation between emotion and value in terms of truth or correctness, sentimentalists talk about appropriateness or fittingness. Our claims here remain neutral on this debate, but we use the term fittingness to avoid strong objectivist implications (see also the Discussion section).

3.1. Study 1: Testing the fittingness model

So far, normative influences have mainly been recorded for *happiness* and *love*. However, other more basic emotions like fear and sadness have not been investigated yet. In order to test our model, we examine folk attributions of fear, which is arguably one of the most prototypical examples of an emotion.[10] Our fittingness model claims that attributing an emotion to an agent does not only depend on what she experiences, but also on whether the situation she is in exemplifies the relevant evaluative property. We designed two vignettes in which an agent is described as experiencing the psychological and bodily states typically associated with fear (Scherer and Summerfield, 1983; Shaver *et al.*, 1987), and manipulated the fittingness of the emotion by varying the value of the situation she reacts to (dangerous vs. non-dangerous). Our experiment had two conditions: *Fitting* (dangerous object) versus *Unfitting* (non-dangerous object). If, as the fittingness model predicts, fittingness standards influence the use of emotion concepts, people's attributions of fear should differ between these two conditions.

3.1.1 Method

Two hundred and forty-four participants were recruited on Amazon's Mechanical Turk (AMT) and completed the survey for a monetary payment of $0.30. Five participants gave wrong answers to the control question (see below) and were excluded, leaving a final sample of 239 participants (118 males, 121 females, M_{age} = 38.14 years, SD = 11.169 years, age range 18–71 years).

Participants were randomly assigned to either the fitting or the unfitting condition. In both conditions, participants read a vignette in which the agent is described as experiencing the psychological and bodily states typically associated with fear. For example:[11]

> *It's Monday, and Tom is getting prepared to leave home to go to work. He takes his bag and walks out of his place. When Tom walks down the road, Tom starts*

[10] In the General discussion, we not only elaborate on the scope and possible limitations of our account, but also on why no effect has been found for concepts like *lust* and *unhappiness* (Phillips et al., 2011).
[11] For a complete list of the vignettes used in this and the following studies, see the Supporting information.

trembling. He feels distressed and nervous. His palms are sweating. His muscles contract and he starts walking fast.

In the fitting condition, the vignette continues by describing an event that everyone is likely to consider dangerous:

The reason Tom feels this way is that, when he walked down the road, he came across a big pitbull. The owner was calling him but the dog was unleashed. The pitbull approached Tom quickly and followed him when he walked away.

In the unfitting condition, the reason for Tom's feelings is a clearly non-dangerous object:

The reason Tom feels this way is that, when he walked down the road, he came across a little poodle. An old woman was taking the dog for a walk with a leash. The poodle approached Tom and wagged its tail.

After reading the corresponding vignette, participants rated how much they agreed with the statement "[Agent] is afraid" on a scale from 1 (completely disagree) to 7 (completely agree).

As a control question, participants were asked to rate whether the statement "[Agent] feels distressed and nervous" was true or false. This served to assure that participants across conditions agreed that the main character in the vignette was in the same psychological states we described him or her to be in. Participants that responded "false" to this question were excluded.

Finally, participants rated their agreement with the statement "[Agent] should be afraid". This allowed us to check whether our manipulation influenced participants normative judgments regarding the emotion.

3.1.2 Results

There was a significant difference in fear ratings between the fitting condition (M = 6.68, SD = 0.77) and the unfitting condition (M = 5.39, SD = 1.68): $t(237) = 7.60$, $p < .001$, $d = .986$ (see Figure 1).[12] In the unfitting condition, 31 out of 119 participants denied that the agent was afraid (gave ratings in the middle of the scale or below), while in the fitting condition only 2 out of 120 did not agree that the agent was afraid. There was a significant association between fittingness and whether or not the agent was *considered to be* afraid χ^2 (1) = 29.85, p < .001. Based on the odds ratio, the odds of denying that the subject was afraid were 20.78 times higher in the unfitting condition. This analysis reveals that the effect not only leads people to downgrade their agreement but, in many cases, to no longer ascribe fear to the agent.

3.1.3 Discussion

While in both conditions the agent is described as having the same reactions and feelings (trembling, distress, nervousness, etc.), attributions of fear were different depending on whether their reaction was fitting (dangerous situation) or unfitting (non-dangerous situation). This suggests that, in line what the fittingness model predicts, the value of the object of emotion influences emotion attributions to that agent. In other words, the results suggest that people are more inclined to judge a person to be afraid, if they believe the person *should* be afraid. Note also that it is highly unlikely for moral considerations to have affected people's responses because the scenarios were designed to leave out any moral aspects that could possibly play a role. Thus, it seems implausible to claim that the moral happiness model can be successfully applied to explain the results of Study 1.

3.2. Study 2: Moral happiness, fitting happiness

As we have already mentioned, Phillips et al. (2011, 2014, 2017) have shown that manipulating the moral character of an agent's life (while maintaining her affective states and life satisfaction constant) influences people's attributions of happiness to that agent. While the moral happiness model explains this effect in terms of happiness being not only about feeling good but also *being good* (value of the agent's life), the fittingness model explains it in terms of happiness being not only about feeling good but also *feeling good about something that is good* (value of the object).

[12] Confirming that our manipulation was successful, participants rated fear as being significantly less appropriate in the unfitting condition (M = 1.88, SD = 1.15) than in the fitting condition (M = 5.85, SD = 1.39), t(237) = 23.98, $p < .001$, d = 3.106.

However, in Phillips et al.'s studies one cannot differentiate between these two possibilities. In their vignettes, agents are described as feeling good and satisfied *about the life they are living*, and thus both models make the same predictions.

To overcome this problem, we designed new vignettes in which the agent does not feel good and satisfied about her own (moral/immoral) actions but someone else's (moral/immoral) behavior. This way, we can distinguish between the moral value of the agent's way of living, which is presumably neutral, and what she feels good about. This allows us to test whether the normative influences on happiness attributions can be explained by our fittingness model. That is, in terms of the value of the object of happiness. Presumably, most people will conceive that having a moral (immoral) partner makes positive feelings fitting (unfitting). Thus, we expect to find a difference in happiness attributions between the moral and the immoral condition of our vignettes. This effect could not be explained by the moral happiness model, as the agent is presumably living a morally neutral life in both cases.

Our study here is not the first one investigating whether happiness attributions are influenced by factors others than the morality of the agents' life. In Phillips et al.'s (2017) Study 5, no significant effect was found in happiness attributions when the agent's life was described as being bad in non-moral terms, like having a terrible job and living in a desolate area. However, this information—terrible job, desolate area—hardly rules out other reasons for feeling good and satisfied like having romantic and social relationships. Thus, when participants read that an agent "rarely ever leaves his small house, and doesn't do much other than sometimes walk along the roads nearby or watch the cows" (Phillips et al., 2017, additional online material) but is in a great mood every day, they will likely infer that he is happy for other reasons like having a great partner or friend. To address this issue, in the present study we manipulated exactly this aspect of people's life. Arguably, having good relationships is more important than having a good job in order to determine whether feeling good and satisfied is fitting.[13]

3.2.1 Method

Two hundred and forty-two participants were recruited on Amazon's Mechanical Turk (AMT) and completed the survey for a monetary payment of $0.25. Seventeen participants gave wrong answers to the control question (see below) and were excluded, leaving a final sample of 225

[13] We would like to thank a reviewer for this journal for pressing us on this point.

participants (82 males, 142 females, 1 Other, M_{age} = 35.20 years, SD = 11.46 years, age range 19–71 years).

Participants were randomly assigned to either the "moral partner" condition or the "immoral partner" condition. For example, one of the moral partner vignettes reads as follows:

> *Sarah and Tom are married and they recently had a baby. Since then, Sarah's life has changed. She decided to quit her job to take care of the baby. She had a very good job, but she didn't care that much. Sarah spends most of the day cooking, cleaning and taking care of the baby while Tom is working extra hours to compensate financially. Sarah believes that Tom does this because he loves her and he wants the best for her and their baby.*

While in the immoral partner version, it reads:

> *Sarah and Tom are married and they recently had a baby. Since then, Sarah's life has changed. Tom forced her to quit her job to take care of the baby. She had a very good job, but he doesn't care about it at all. Sarah spends most of the day cooking, cleaning and taking care of the baby because Tom doesn't want her to go outside and meet other people. Sarah believes that Tom does this because he loves her, and he wants the best for her and their baby.*

Both the moral and immoral partner versions of the vignette continue with an identical description of the agent's psychological states:

> *Almost every 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.*

After reading the corresponding vignette, participants rated how much they agreed with the statement "[Agent] is happy" on a scale from 1 (completely disagree) to 7 (completely agree). As in Study 1, participants answered a psychological states control question which served as an inclusion criterion. They rated whether the statement "[Agent] feels good and generally experiences a lot of pleasant emotions" was true or false. Finally, participants rated the moral character of both the agent and her partner on a scale from 1 (completely moral person) to 7 (completely immoral person). This allowed us to check whether our manipulation was successful in manipulating the moral character of the partner without manipulating perceptions of the morality of the agent's life.

3.2.2 Results

Happiness ratings in the moral partner condition (M = 6.60, SD = .81) were significantly higher than happiness ratings in the immoral partner condition (M = 5.22, SD = 1.72), $t(223) = 7.85$, $p < .001$, $d = 1.025$ (Figure 2). In the immoral partner condition, 32 out of 105 participants did not agree that the subject was happy (gave ratings in the middle of the scale or below), while in the moral partner condition only five out of 120 did not agree that the subject was happy. There was a significant association between the morality of the object and whether or not the subject was happy; χ^2 (1) = 28.21, p < .001. Based on the odds ratio, the odds of denying that the subject was happy were 10.08 times higher in the immoral object condition.

Participants rated the partners as significantly less moral in the morally bad condition (M = 2.59, SD = 1.59) than in the morally good condition (M = 5.93, SD = 1.11), $t(223) = 18.41$, $p < .001$, $d = 2.436$. There was a marginally significant difference between agents' moral character ratings in the morally good condition (M = 5.58, SD = 1.19) and in the morally bad condition (M = 5.19, SD = 1.68), $t(223) = 2.00$, $p = .052$, $d = 0.268$. One might worry that this incidental manipulation of the agent's moral character explains the recorded effect on happiness attributions. If this were the case, our results could be explained by the moral happiness model. We found a significant correlation between agents' moral character ratings and happiness attributions ($r = .363$, $p < .001$). However, a further regression analysis controlling for agents' moral character ratings ($\beta = .307$, $p < .001$) still showed an effect of condition ($\beta = -.424$, $p < .001$) on happiness attributions. Thus, although agents' moral character ratings were accidentally influenced by our manipulation, the results are not well explained by morality (see also Study 4, in which we further tackle the question of whether factors other than fittingness might explain differences in people's attributions of happiness).

3.2.3 Discussion

While in both conditions the agent is described as having the same psychological states (high pleasure, low pain, and high life satisfaction), attributions of happiness were different depending on whether the agent had a moral or an immoral partner. All things equal, having a moral partner makes happiness fitting (e.g., one should be happy about having a supportive husband, it is good), while having an immoral partner makes happiness unfitting (one should not be happy about having an abusive husband, it is not good). Thus, these results suggest that fittingness considerations influence the use of the concept of happiness. This provides further evidence in support for our fittingness model. While these results do not prove the moral happiness model wrong, they suggest that the effect of morality on happiness attributions recorded in previous studies (Phillips et al., 2011, 2014, 2017) can be accounted for by our fittingness model.

4. OBJECTIONS

Do the empirical studies presented in the previous section clearly demonstrate that our folk concepts of emotions are shaped by fittingness considerations? In this section, we consider two possible alternative explanations for our results. The first objection is that our results in Study 1 are due to the absence of an important descriptive component of emotion: Evaluative judgments. The second objection is that our Study 2 merely suggests but does not demonstrate that the effect was driven by fittingness attributions. Other alternatives, like the morality and/or meaningfulness of the agent's life (Phillips et al., 2017), have not been shown to be false. In this section, we present the results of two studies designed to rule out these alternative explanations of our results.

4.1. Study 3: Fear and evaluative judgments

Whereas some emotion researchers believe emotions to be feelings or perceptions of bodily changes (Hufendiek, 2016; James, 1884; Prinz, 2004), others have argued that emotions are fundamentally cognitive states like appraisals or judgements (Bedford, 1957; Nussbaum, 2001; Solomon, 1976). Others, again, have proposed hybrid theories accepting that neither feelings

nor judgements are sufficient to single out emotions (Barlassina and Newen, 2014; Greenspan, 1981; Hume, 1978).

While in Study 2 we provided descriptive information about both Sarah's feelings (she experiences high positive affect and low negative affect) and her appraisal of the situation (she evaluates her life to be satisfactory), Study 1 mostly provided information about the agent's feelings and bodily reactions, with no reference to cognitive states. Thus, authors that defend cognitive theories of emotions are likely to point out that the vignettes in Study 1 lacked an important ingredient when describing the states of the agents. According to their accounts, the essential core of an emotion is an evaluative judgement or appraisal of the right kind:

A subject S is [emotion] about X only if S judges X to be [evaluative property].

For example, Tom is afraid of the dog only if he judges the dog to be dangerous. None of the scenarios we designed for Study 1 specified that the agents were evaluating the situation as dangerous. Cognitivists are therefore likely to question whether the data we collected demonstrate that emotion attributions are indeed influenced by normative considerations. Instead, they could argue that the participants who read the fitting condition are likely to infer that the agent judges the situation to be dangerous. Hence, they will strongly agree with the claim that the agent is afraid. However, participants in the unfitting condition would infer that the agent does not judge the situation to be dangerous, and thus will not (strongly) agree that the agent is afraid. In other words, the difference in fear ratings could be explained by people's inferences about the emotion the agent *is* in, rather than people's views about the emotion she *should* be in.[14]

In order to examine the "cognitivist" challenge, we reran Study 1, albeit with one important modification. Not only did we specify the motions the agent went through, we also mentioned that the agent considered the situation to be dangerous in both the fitting and the unfitting conditions. For example:

[14] Note, that feeling theorists cannot object to our results in a similar fashion, because in both types of cases it was specified that Tom felt nervous and uneasy, and that his hands were sweaty and he was trembling. Furthermore, only the data from those subjects who correctly answered a control question about Tom's feelings were included in the statistical analyses.

When Tom walks down the road, Tom starts trembling. He thinks that the situation is dangerous. He feels distressed and nervous. His palms are sweating and his muscles contract.

If the cognitivist objection holds, the effect recorded in Study 1 should disappear.

4.1.1 Method

Two hundred and forty-two participants were recruited on Amazon's Mechanical Turk (AMT) and completed the survey for a monetary payment of $0.30. Participants were randomly assigned to either the "fitting" or the "unfitting" condition. Thirteen participants answered at least one of the control questions (see below) wrongly and were excluded, leaving a final sample of 229 participants (93 males, 136 females, M_{age} = 34.60 years, SD = 10.625 years, age range 18–77 years).

After reading the corresponding vignette, participants rated how much they agreed with the statement "[Agent] is afraid" on a scale from 1 (completely disagree) to 7 (completely agree). Afterwards, participants answered two control questions about the character's psychological states. The first control question consisted in rating whether the statement "[Agent] feels distressed and nervous" was true or false. The second control question consisted in rating whether the statement "[Agent] thinks that the situation is dangerous" was true or false. This control questions served as inclusion criteria.

4.1.2 Results

There was a significant difference in fear ratings between the fitting condition (M = 6.80, SD = 0.55) and the unfitting condition (M = 5.93, SD = 1.35), $t(227) = 6.36$, $p < .001$, $d = .842$ (Figure 3).[15] Seventeen out of 113 participants in the unfitting condition denied that the agent was afraid

[15] We would like to point out that the difference between the two conditions was slightly less pronounced compared to the original data in Study 1. This is because the influence of our fittingness manipulation was different depending on the vignette. A factorial ANOVA using "fittingness" and "vignette" as fixed factors and fear ratings as the dependent variable revealed an interaction between the fittingness of the emotion and the vignette presented F (1, 225) = 16.17, $p < .001$, $d = .570$. Fear ratings in the unfitting condition were significantly higher in the first vignette (poodle case: M = 6.28) than in the second vignette (cleaning case: M = 5.53). We believe this is the case because people consider the dangerousness of a poodle to be relative to the person evaluating it, whereas they think that getting your office cleaned is objectively not dangerous. Thus, fear is always considered unfitting in the cleaning case, but not necessarily in the poodle Case. This explains why we did not observe the same reduction in fear ascriptions in the latter.

(gave ratings in the middle of the scale or below), while in the fitting condition, only 1 out of 116 did not agree that the agent was afraid. There was a significant association between fittingness and whether or not the subject was afraid; χ^2 (1) = 15.90, $p < .001$. Based on the odds ratio, the odds of denying that the subject was happy were 20.36 times higher in the unfitting condition.

4.1.3 Discussion

In Study 3, we specified that the agent judges the situation as dangerous. Despite this addition, whether the situation is in fact dangerous or not still had a significant effect on fear ascriptions. Thus, it seems like, regardless of whether the folk concepts of emotion pick out feelings or cognitive appraisals, none of these theories can account for the effect based merely on the descriptive features of emotions.

4.2. Study 4: Fittingness, morality, and meaningfulness

In our Study 2, we showed that folk attributions of happiness are not only affected by the moral value of actions performed by the agent one is ascribing happiness to (Phillips et al. 2011, 2014, 2017), but can also be affected by the moral value of the actions of someone else who is part of the agent's social environment (e.g., his romantic partner). We interpreted this effect in terms of happiness attributions being affected by whether the agent's psychological states fit the situation she is in. However, we did not directly test this, as we included no explicit measure of fittingness. Thus, someone could argue that the effect we found is not actually driven by fittingness but rather by another variable that was not measured.

In their most recent paper on the role of morality in the folk concept of happiness, Phillips and colleagues (2017) suggest that the folk concept of happiness might not only involve moral considerations but also consideration regarding the meaningfulness of the agent's life. While it seems intuitive that the effect recorded in Study 2 is not driven by perceptions of the morality of the agent's life (it is not the agent but the partner who does immoral acts), it could be that the effect is driven by perceptions of meaningfulness.[16] On this view, the life of a woman whose romantic partner treats her badly lacks meaningfulness, and this makes participants think that she is not fully happy. In order to test whether the results of Study 2 are due to perceptions

[16] We owe this objection to a reviewer for this journal.

of morality, meaningfulness, or fittingness, we reran this study including explicit measures of these three potential mediators of the effect.

4.2.1 Method

One hundred and fifty-six participants were recruited on Amazon's Mechanical Turk (AMT) and completed the survey for a monetary payment of $0.30. Seventeen participants gave wrong answers to the control questions (see below) and were excluded, leaving a final sample of 139 participants (83 males, 56 females, M_{age} = 37.10 years, SD = 11.89 years, age range 22–70 years).

Participants were randomly assigned to either the "moral partner" condition or the "immoral partner" condition. The vignettes were taken verbatim from the supportive/abusive husband scenarios from Study 2 (see Methods section of Study 2 for the full wording of the vignettes). After reading the vignette, participants rated how much they agreed with the statement "Sarah is happy" on a scale from 1 (completely disagree) to 7 (completely agree). Afterwards, they rated how much they agreed to the following statements on a scale from 1 (completely disagree) to 7 (completely agree): "Sarah lives a morally good life" (morality), "Sarah's life is meaningful" (meaningfulness), and "Feeling good and satisfied is fitting given the situation Sarah is in" (fittingness). Order of presentation was randomized. As in in the rest of our studies, participants answered psychological states control questions which served as inclusion criteria. They rated whether the statements "Sarah feels good and generally experiences a lot of pleasant emotions" and "Sarah doesn't like to look after the baby"[17] were true or false.

4.2.2 Results

Happiness ratings in the moral partner condition (M = 6.52, SD = .77) were significantly higher than happiness ratings in the immoral partner condition (M = 5.34, SD = 1.61), $t(137) = 5.57$, $p < .001$, $d = 0.94$. In the immoral partner condition, 17 out of 68 participants did not agree that the agent was happy (gave ratings in the middle of the scale or below), while in the moral partner condition only 2 out of 71 did not agree that the subject was happy. There was a significant association between the morality of the object and whether or not the agent was happy; χ^2 (1)

[17] Some participants in the study might infer that Sarah is not truly satisfied but just accepting a situation that she does not really like. This control question addresses this worry.

= 14.48, $p < .001$. Based on the odds ratio, the odds of denying that the agent was happy were 11.5 times higher in the immoral partner condition.

For all fittingness, meaningfulness and morality ratings, there were significant differences between the moral partner condition ($M_{Fittingness} = 6.13$, $M_{Meaningfulness} = 6.23$, $M_{Morality} = 6.06$) and the immoral partner condition ($M_{Fittingness} = 4.87$, $M_{Meaningfulness} = 5.29$, $M_{Morality} = 5.63$), all ps $< .05$.[18] In order to examine which of these variables is driving the recorded effect on happiness attributions, we conducted a multiple mediation analysis introducing morality, meaningfulness, and fittingness ratings as possible mediators of the relationship between condition and happiness ratings. We found a significant indirect effect of condition on happiness attributions through fittingness ratings, b = .37, 95% CI [-.76, -.06]. Neither morality nor meaningfulness significantly mediated the effect of condition on happiness attributions (Figure 4).

4.2.3 Discussion

Replicating the results of Study 2, participants were more willing to attribute happiness to agents who have a moral partner than agents that have an immoral partner. Here, we also tested whether the effect is mediated by perceptions of the morality of the agent's life, perceptions of the meaningfulness of the agent's life, or perceptions of fit between the agent's psychological states and the situation she is in. Our analysis showed that the effect of the condition on happiness attributions was *only* significantly mediated by fittingness ratings. This suggests that our fittingness model does a better job than the moral happiness model in explaining the normative influences we found in this study.

It is important to note that, although the indirect effect of condition on happiness attributions through meaningfulness ratings was not significant, we believe that meaningfulness (and even morality) could play an indirect role in happiness attributions. Whether someone's life is meaningful could determine whether it is a good life, and thus a life that is worthy of happiness. In other words, meaningfulness could be part of happiness' fitting conditions. Similar to how a Pitbull's aggressive behavior makes fear fitting, the meaningfulness of someone's life could make happiness fitting. An interesting avenue for future research is thus to investigate

[18] There were significant correlations between fittingness ratings and meaningfulness ratings ($r = .672$, $p < .001$), meaningfulness ratings and morality ratings ($r = .527$, $p < .001$), as well as between morality ratings and fittingness ratings ($r = .571$, $p < .001$).

which factors people consider relevant to determine whether someone's life is good, and thus worthy of happiness.

5. GENERAL DISCUSSION AND CONCLUSION

In a series of studies, we tested our fittingness model for the impact of normativity on the use of emotion concepts. In particular, we tested people's attributions of fear (Study 1 and Study 3) and happiness (Study 2 and Study 4). In line with the fittingness model's predictions, the results of our studies showed that whether or not the situation was dangerous or good had a significant impact on attributions of fear and happiness, respectively. People seem to consider that being in a particular emotional state is not only a matter of having certain psychological and bodily states, but also a matter of whether those psychological and bodily states fit the situation in which they occur. Our first two studies served to test the main prediction of the fittingness model (Study 1) and its aptness to explain some of the previous results regarding normative influences on the folk concept of happiness (Study 2). Studies 3 and 4 were designed to tackle possible alternative explanations of our results. The results of Study 3 suggest that the effect recorded in Study 1 is not due to participants making inferences about the descriptive psychological states the agent has. And the results of Study 4 provide evidence that the effect on happiness attributions we recorded in Study 2, is not due to other factors like the morality or the meaningfulness of the agent's life. Thus, it seems that our results truly reflect that emotion concepts are shaped by normative fittingness considerations.

In this paper, we only tested the fittingness model for two emotions: Fear and happiness. It is therefore an open question whether other affective states are similarly influenced by normative judgements. Many other emotions like sadness, envy, anger, and so forth, may fail to represent the world correctly and thus can be judged to be unfitting given the situation the agent is in. Thus, we predict that the fittingness model applies to a wide range of affective states including sadness, envy and anger. Such a wide scope of application of our model does not, however, seem to square with previous results by Phillips et al. (2011, 2014). In their studies, effects were only found for *happiness* and *love*, but not for *lust* and *unhappiness*.

In contrast to love, we believe that lust is better characterized as a desire instead of an emotion. Whereas emotions do aim at representing the world, and so can be assessed on whether they accurately represent the world (whether they are fitting or not), desires do aim at being fulfilled, and so can be assessed in terms of whether they are satisfied or not. As lust seems

to behave more like a desire than an emotion, our model does not predict a normative influence on lust attributions. Phillips and colleagues also failed to find a normative influence on the state of unhappiness. Initially, one might think that our account clearly predicts an effect for unhappiness. However, states of unhappiness are quite likely to be interpreted differently. While "unhappy" is quite a common word in English (only three times less common than sad), people hardly express themselves by using the "feeling unhappy" phrase (20 times less common than "feeling sad"),[19] suggesting that (at least in some contexts) unhappiness is not considered an affective state. This seems to be also reflected by dictionaries, for example, the Cambridge English Dictionary, and the Collins English Dictionary, which often define unhappiness in terms of dissatisfaction and less in terms of unpleasant feelings. Thus, if many people read the term "unhappy" to be synonymous with "dissatisfied", then a normative effect might be strongly extenuated.

Besides the cases of lust and unhappiness, there are likely to be other limitations to the scope of our model: First, moods and biological drives are less obviously states that represent the world to be such-and-so. Accordingly, there may be no situations in which these states can be properly considered fitting or unfitting. Second, pains and other bodily sensations may represent one's body to be in a certain state, but may still not be open to normative judgements, because bodily states are usually private and hence, not accessible to other persons. Third, it is a matter of debate whether people consider more complex emotions like boredom to be states with an evaluative content. Certainly, more research is needed to make more reasoned predictions about such states. Fourth, while we have shown that fear can be influenced by normative considerations, phobias might not display the same effect. Phobias are anxiety disorders and therefore less likely to be open to normative assessments than "normal" cases of fear. Accordingly, if a person has a phobia of mice, we do expect people to attribute high levels of fear to that person, because people will consider the relation between the experiential states of fear and its intentional object to be hard-wired into the agent's system. Thus, even if there is a tendency to attribute less fear when the object of fear is not dangerous, this might as well be overridden by information regarding the phobic nature of a person's fear.

It is important to note that our studies do not put into question that emotions are constituted by bodily feelings, by appraisals, or by a combination of both. In fact, based on previous research, we assumed that people take those descriptive elements to be constitutive

[19] We found the following number of hits on the Corpus of Contemporary American English (COCA): "feeling sad" (96 times), feeling unhappy (5 times); "feel sad" (329 times), "feel unhappy (20 times).

of emotions. Instead, what the results of our studies show is that bodily feelings and appraisals do not always suffice to identify emotions. In order to determine the specific emotion at hand, people rely on whether those bodily feelings and appraisals fit the situation in which they are experienced. In other words, whether those descriptive elements are normatively adequate. Note that this puts pressure on any account that posits emotions to be non-intentional states (Morag, 2017; Whiting, 2011): If emotions and the situation in which they are experienced are only causally related (if anything), it seems difficult to explain why the value of the situation makes a difference in emotion attribution across cases where the reaction of the agent is the same.

But why do fittingness standards influence emotion attribution? There are different ways to account for our results in terms of people's conception of emotions. First, it could be that emotions are considered to reliably track, or be caused by, evaluative properties in the environment. As Jesse Prinz states, "Emotions … are reliably caused by relational properties that pertain to well-being" (Prinz, 2004, p. 66). Thus, only reactions that fit the situation are considered to be cases of the relevant emotion. Second, the normative aspect of folk concepts of emotions could be reflecting the "embeddedness" of emotions, an understanding of emotions as co-dependent upon extra-bodily aspects, or as "situated" processes (Stephan et al., 2014). Along these lines, Lisa Feldman Barrett (2012, 2014) has already claimed that emotion categorization is the result of meaningfully relating changes in the body to changes in the environment. In light of our account, this "meaningful relation" could be interpreted in terms of fittingness. Finally, some have gone further and claimed that external processes are not only intimately related to, but *constitutive* of emotions (Krueger and Szanto, 2016). Accordingly, the normativity of folk concepts of emotion could be "explained away": The situation is always relevant for emotion attribution just because it is part of the emotion.

To conclude, we would like to connect our fittingness model with some results in the literature regarding the impact of context in emotion recognition in faces. In an influential study, Carroll and Russell (1996, Study 1) presented participants with a picture of a face posing the stereotypical "fear" expression paired with a written description of either (1) a dangerous situation or (2) an offensive situation. Participants in the "danger" condition tended to "recognize" the facial expression as fear, while most participants in the "offense" condition thought the face expressed anger. Our fittingness model provides a straightforward explanation for this effect: Facial expression being the same, people were largely guided by selecting the emotion that was fitting in each situation. In the offensive situation, only anger was fitting, and thus participants attributed anger despite being presented with a face displaying a prototypical

fear expression. The results we presented in this paper nicely complement these previous results (see also Aviezer *et al.*, 2008, Study 1). In our studies, the information that was available for participants went beyond facial expressions. Instead, we provided participants with detailed information about the psychological and bodily features associated with the emotions we investigated. Just as the studies by Carroll and Russell challenge the claim that facial expressions are sufficient to identify emotions, our studies question whether bodily feelings or cognitive appraisals alone make emotions intelligible.

We hope to have paved the way for more experimental as well as theoretical work on emotions to charter the normative aspects of emotions as well as explain why emotions are in part governed by normativity. Further studies along these lines could inform long-standing debates and foster communication between philosophers and researchers in other disciplines interested in the study of emotion.

ACKNOWLEDGEMENTS

This work was supported in part by the Swiss National Science Foundation (grant number 100012_169484, Kevin Reuter). We would like to thank Florian Cova, and the audiences in Bern, Hertfordshire, Madrid, and Osnabrück, for their comments on early drafts of this paper.

REFERENCES

Alicke, M. D. 1992: Culpable causation., *Journal of Personality and Social Psychology*, 63(3): 368–78. doi: 10.1037/0022-3514.63.3.368.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M. and Bentin, S. 2008: Angry, disgust, or afraid? Studies on the malleability of emotion perception., *Psychological Science*, 19(7): 724–32. doi: 10.1111/j.1467-9280.2008.02148.x.

Barlassina, L. and Newen, A. 2014: The Role of Bodily Perception in Emotion: In Defense of an Impure Somatic Theory, *Philosophy and Phenomenological Research*, 89(3): 637–78. doi: 10.1111/phpr.12041.

Barrett, L. F. 2012: Emotions are real., *Emotion*, 12(3): 413–29. doi: 10.1037/a0027555.

Barrett, L. F. 2014: The Conceptual Act Theory: A Précis, *Emotion Review*, 6(4): 1–20. doi: 10.1177/1754073914534479.

Bedford, E. 1957: Emotions, *Proceedings of the Aristotelian Society*, 57: 281–304.

Carroll, J. M. and Russell, J. A. 1996: Do Facial Expressions Signal Specific Emotions? Judging Emotion from the Face in Context, *Journal of Personality and Social Psychology*, 70(2): 205–18. doi: 10.1037/0022-3514.70.2.205.

Chtuc, V. 2012: Emotions as Dual Character Concepts, *Unpublished manuscript*,.

D'Arms, J. and Jacobson, D. 2000: The Moralistic Fallacy: On the 'Appropriateness' of Emotions, *Philosophy and Phenomenological Research*, 61(1): 65–90. doi: 10.2307/2653403.

Deonna, J. A. and Teroni, F. 2012: *The Emotions : A Philosophical Introduction*. New York: Routledge. doi: 10.1016/B978-0-12-415956-3.00001-4.

Feldman, F. 2010: *What Is This Thing Called Happiness?* Oxford: Oxford Sholarship Online. doi: 10.1093/acprof:oso/9780199571178.001.0001.

Foot, P. 2001: *Natural Goodness*. New York: Oxford University Press. doi: 10.1093/0198235089.001.0001.

Greenspan, P. S. 1981: Emotions As Evaluations, *Pacific Philosophical Quarterly*, 62(2): 158–69. doi: 10.1111/j.1468-0114.1981.tb00054.x.

Hitchcock, C. and Knobe, J. 2009: Cause and Norm, *Journal of Philosophy*, 106(11): 1–45. doi: 10.1177/03063127067078012.

Hufendiek, R. 2016: *Embodied Emotions : A Naturalist Approach to a Normative Phenomenon*. New York: Routledge.

Hume, D. 1978: *A Treatise of Human Nature*. Edited by P. Nidditch. Oxford: Oxford University Press.

James, W. 1884: What is an Emotion ?, *Mind*, 9(34): 188–205. doi: 10.1093/mind/LI.202.200.

Kirchin, S. 2010: The Shapelessness Hypothesis, *Philosopher's Imprint*, 10: 1–28.

Knobe, J. 2003: Intentional Action and Side Effects in Ordinary Language, *Analysis*, 63(3): 190–94. doi: 10.1111/1467-8284.00419.

Knobe, J., Prasada, S. and Newman, G. E. 2013: Dual Character Concepts and the Normative Dimension of Conceptual Rrepresentation, *Cognition*, 127(2): 242–57. doi: 10.1016/j.cognition.2013.01.005.

Krueger, J. and Szanto, T. 2016: Extended emotions, *Philosophy Compass*, 11(12): 863–78. doi: 10.1111/phc3.12390.

Lutz, C. and White, G. M. 1986: The Anthropology of Emotions, *Annual Review of Anthropology*, 15(1): 405–36. doi: 10.1146/annurev.an.15.100186.002201.

Morag, T. 2017: The Tracking Dogma in the Philosophy of Emotion, *Argumenta*, 2: 343–63. doi: 10.23811/410.arg2017.mor.

Nussbaum, M. 2001: *Upheavals of Thought*. Cambridge: Cambridge University Press. doi: 10.5840/gfpj200223213.

Pettit, D. and Knobe, J. 2009: The pervasive impact of moral judgment, *Mind and Language*, 24(5): 586–604. doi: 10.1111/j.1468-0017.2009.01375.x.

Phillips, J., Freitas, J. De, Mott, C., Gruber, J. and Knobe, J. 2017: True happiness: The role of morality in the folk concept of happiness, *Journal of Experimental Psychology: General*, 146(2): 165–81. doi: 10.1037/xge0000252.

Phillips, J., Misenheimer, L. and Knobe, J. 2011: The Ordinary Concept of Happiness (and Others Like it), *Emotion Review*, 71(3): 929–37. doi: 10.1177/1754073911402385.

Phillips, J., Nyholm, S. and Liao, S.-Y. 2014: The Good in Happiness, in T. Lombrozo, J. Knobe, and S. Nichols (eds) *Oxford Studies in Experimental Philosophy*,. Oxford: Oxford University Press.

Prinz, J. 2004: *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.

Putnam, H. 2002: *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA: Harvard University Press.

Reuter, K. 2019: Dual character concepts, *Philosophy Compass*, 14(1). doi: 10.1111/phc3.12557.

S. Lazarus, R. 1991: *Emotion and Adaptation*. New York: Oxford University Press.

Scherer, K. and Summerfield, A. 1983: Cross-national research on antecedents and components of emotion: A progress report, *Social Science Information*, 22(3): 355–85. doi: 10.1177/053901883022003002.

Shargel, D. 2015: Emotions without objects, *Biology and Philosophy*, 30(6): 831–44. doi: 10.1007/s10539-014-9473-8.

Shaver, P., Schwartz, J., Kirson, D. and O'Connor, C. 1987: Emotion knowledge: Further exploration of a prototype approach, *Journal of personality and social psychology*, 52(6): 1061–86. doi: 10.1037/0022-3514.52.6.1061.

Solomon, R. C. 1976: *The Passions*. New York: Anchor Press.

Solomon, R. C. 1977: The Rationality of the Emotions, *The Southwestern Journal of Philosophy*, 8(2): 105–14. doi: 43155160.

de Sousa, R. 2014: Emotion, in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*,. Metaphysics Research Lab, Stanford University. doi: 10.1111/1467-9973.00225.

Stephan, A., Walter, S. and Wilutzky, W. 2014: Emotions beyond brain and body, *Philosophical Psychology*, 27(1): 65–81. doi: 10.1080/09515089.2013.828376.

Suikkanen, J. 2011: An improved whole life satisfaction theory of happiness, *International Journal of Wellbeing*, 1(1): 149–66.

Sumner, L. 1996: *Welfare, Happiness, and Ethics*. New York: Oxford University Press.

Sytsma, J., Bluhm, R., Willemsen, P. and Reuter, K. 2019: Causal Attributions and Corpus Analysis, in E. Fischer and M. Curtis (eds) *Methodological Advances in Experimental Philosophy*,. London: Bloomsbury.

Teroni, F. 2007: Emotions and formal objects, *Dialectica*, 61(3): 395–415. doi: 10.1111/j. 1746-8361.2007.01108.x.

Todd, C. 2014: Emotion and value, *Philosophy Compass*, 9(10): 702–12. doi: 10.1111/phc3.12167.

Whiting, D. 2011: The feeling theory of emotion and the object-directed emotions, *European Journal of Philosophy*, 19(2): 281–303. doi: 10.1111/j.1468-0378.2009.00384.x.