

Emotions and beliefs about morality can change one another

Monica Bucciarelli^{a,*}, P.N. Johnson-Laird^{b,c}

^a Dipartimento di Psicologia and Center for Logic, Language and Cognition, Università di Torino, Turin 10123, Italy

^b Stuart Professor of Psychology, Emeritus, Princeton University, Princeton, NJ 08540, USA

^c Visiting Scholar, Department of Psychology, New York University, New York, NY 10003, USA

ARTICLE INFO

Keywords:

Beliefs
Deontics
Emotions
Moral assertions
Reasons
Subjective probabilities

ABSTRACT

A dual-process theory postulates that belief and emotions about moral assertions can affect one another. The present study corroborated this prediction. Experiments 1, 2 and 3 showed that the pleasantness of a moral assertion – from loathing it to loving it – correlated with how strongly individuals believed it, i.e., its subjective probability. But, despite repeated testing, this relation did not occur for factual assertions. To create the correlation, it sufficed to change factual assertions, such as, “Advanced countries are democracies,” into moral assertions, “Advanced countries should be democracies”. Two further experiments corroborated the two-way causal relations for moral assertions. Experiment 4 showed that recall of pleasant memories about moral assertions increased their believability, and that the recall of unpleasant memories had the opposite effect. Experiment 5 showed that the creation of reasons to believe moral assertions increased the pleasantness of the emotions they evoked, and that the creation of reasons to disbelieve moral assertions had the opposite effect. Hence, emotions can change beliefs about moral assertions; and reasons can change emotions about moral assertions. We discuss the implications of these results for alternative theories of morality.

1. Introduction

Your actions usually depend on your factual beliefs about what is real and what is possible, and on your deontic beliefs about what is obligatory and what is permissible. Factual assertions, such as:

Some public and private schools play the same sports

are in principle open to empirical test. You discover schools of both sorts do play the same sports, and so the assertion is true (see, e.g., Russell, 1912, for a defense of this “correspondence” theory of truth). In contrast, the parallel deontic assertion:

Some public and private schools ought to play the same sports

is not open to empirical test. The claim's truth depends, not on facts, but on moral principles. A logical positivist might suppose that if assertions are not susceptible to empirical tests, they do not have truth values – they are neither true nor false. But, it is true that one ought not to torture children, even though the claim is not open to empirical test. To argue otherwise on the grounds of a philosophical doctrine is to depart from everyday values and practices.

As we have just illustrated, deontic assertions often concern morality. Yet, no simple way exists to pick out all and only moral assertions from their deontic superset, which includes truths based on laws, rules, and social conventions (Bucciarelli, Khemlani, & Johnson-Laird, 2008). They are not dependent on facts about the world, and so like other

unempirical beliefs, such as those concerning religion, aesthetics, and life after death, individuals can have intense emotions about them (Johnson-Laird, 2006, Ch. 23). What factual and deontic assertions share is that beliefs about them vary in degree – from certainty to impossibility. They also share the capacity to evoke emotional reactions – from liking to loathing. Our aim in the present investigation was to establish the relations between degrees of belief in assertions and emotional reactions to them, where the assertions were factual or moral.

Degrees of belief can be treated as subjective probabilities (Ramsey, 1990; de Finetti, 1937/1964). Hence, in a numerate culture, individuals are prepared to assert that, for instance, they are *90% certain* about a proposition. In a non-numerate culture, they are prepared to say that, for instance, it is *highly likely* or *more than possible*. And we doubt whether any culture lacks an intuitive grasp of likelihood or possibility. No evidence exists contrary to the concept of belief as a subjective probability, and it applies both to factual and to deontic claims. Yet, not all psychologists accept subjective probabilities (cf. Cosmides & Tooby, 1996; Gigerenzer, Hertwig, van den Broek, Fasolo, & Katsikopoulos, 1995). They defend the notion that the only sensible probabilities are those based on the natural frequencies with which events occur. One problem for this view is that individuals in our society are happy to estimate the probabilities of events that have no natural frequencies,

* Corresponding author.

E-mail addresses: monica.bucciarelli@unito.it (M. Bucciarelli), phil@princeton.edu (P.N. Johnson-Laird).

such as the probability that Trump will be impeached (Khemlani, Lotstein, & Johnson-Laird, 2012, 2015). Like people who live in cultures without numbers, they also have a non-numerical notion of probability (Fontanari, Gonzalez, Vallortigara, & Girotto, 2014). Hence, subjective probabilities do not need to be numerical, and so cannot be based only on natural frequencies. The theory of mental models shows how the subjective probability of a unique event can reflect the proportion of models of relevant evidence in which the event occurs, and it can be represented in an iconic model of a magnitude of a sort found in infants, animals, and adults in non-numerical cultures (Khemlani et al., 2012, 2015; see, e.g., Gordon, 2004). We accordingly treat degrees of belief in factual and deontic assertions as subjective probabilities.

Emotions about people, things, or events, also vary, and one reliable contrast is the degree of liking or disliking them. According to the “communicative” theory of emotions (e.g., Johnson-Laird & Oatley, 2016; Oatley & Johnson-Laird, 1987, 1996), emotions are internal communications in the brain that enable organisms to select goals. The process in humans relies on two different mechanisms, one of which is evolutionally older than the other. The older mechanism conveys no internal symbolic structure of significance: an emotional signal simply sets the whole system into a particular mode to prepare for a general course of action or inaction. The mechanism implements innate basic emotions, which have their own distinctive signals, subjective experiences, and facial expressions (Keltner, Ekman, Gonzaga, & Beer, 2003). Cognitive evaluations eliciting emotions depend in part on cultural factors, as shown in the case of disgust (e.g., Rozin, 1996), but they are primitive as shown in the simple elements of music that evoke emotions (Johnson-Laird & Oatley, 2016). Some basic emotions, such as happiness and sadness, can be experienced without any propositional content, whereas others, such as hatred, have a known object. In contrast, the newer mechanism is “propositional” in that it communicates symbolic messages. They combine with basic emotions to yield complex emotions, which often concern mental models of the self and others (Johnson-Laird & Oatley, 2000). Individuals therefore cannot experience complex emotions without an awareness of the evaluations that elicited them. For example, individuals feel remorse, which is sadness about an action or inaction, because they judge themselves to have violated their idealized models of themselves. Unlike anxiety, free-floating remorse would be paradoxical – in the customary sense of the word, it needs something to be remorseful about.

A contrast pertinent to our studies concerns the difference between emotional feelings and emotional evaluations. Studies in the nineteenth century showed that people report an emotional reaction to music (e.g., Downey, 1897). But, they can distinguish between the emotion that music is intended to convey and the emotion that they experience as a result of listening to music. The two often go hand in hand, but listeners can make an emotional evaluation without having a corresponding emotional experience (Krumhansl, 1997). To distinguish between the two in the laboratory is difficult, and we have not attempted to do so in the present studies. Readers should therefore bear in mind that our studies are neutral with respect to the two, and so we use the expression “emotional reaction” to embrace an evaluation of an emotion with, or without, its actual experience.

What might be the relations between beliefs and emotions? You could suppose that the degree to which you believe an assertion is merely the strength of your emotion about it. A moment's thought shows that this hypothesis is wrong. You can be certain that the news will be broadcast at 6 p.m. but have no strong emotion about the matter. Conversely, you may hate the proposal to ban immigrants from entering a country, yet believe it to be highly likely. That belief and emotion can vary independently is clear for matters of fact, but less clear for moral assertions. Many different accounts of their relations exist, and so we consider only three theories that hold the most distinct positions about the matter.

The original Utilitarians identified the good with pleasure. For example, Bentham (1995/1789, p.11) wrote: “Nature has placed mankind

under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do” (see also Mill, 1861, for a similar view). So, if an action is moral then it is desirable. An interpretation of this view that we owe to Jon Baron (p.c., 11-5-2018) is that believing a moral assertion and liking it are two ways of saying much the same thing. Their relation is in effect a tautology, and so any investigation of it is a study in lexical semantics. In fact, however, the relation is not tautological, because it admits counterexamples. For example, it is morally right to pay your income tax, but it need not be pleasurable. Indeed, if *believing* it and *liking* it are two ways of saying the same thing about moral assertions, it ought to be self-contradictory to assert:

I believe that it's right to pay income tax but I don't like doing it.

Yet, it is not a contradiction, but a common complaint. Moreover, many religions draw a sharp line between what is good and what is pleasant, and urge their followers to do the right thing and eschew pleasure (see, e.g., Corinthians 2). As we will see, an intimate relation exists between them, but it is not a tautological one.

As the Utilitarians recognized, one of their precursors was a philosopher who argued that moral judgments depend on emotions. Hume (1978/1739) wrote:

Morals excite passions, and produce or prevent actions. Reason of itself is utterly impotent in this particular. The rules of morality, therefore, are not conclusions of our reason.

This idea is embodied in the modern *socio-intuitionist* theory. It postulates that intuitions about what is moral or immoral depend on instant feelings of approval or disapproval. No need exists to consider evidence or to make inferences: all that matters are immediate emotional appraisals (Haidt, 2001, 2007; see also Blair, 1995). They occur before reasoning and they yield moral judgments (Shweder & Haidt, 2000; Wilson, 1993). Unlike intuition, conscious reasoning occurs slowly, calls for effort, and includes some steps that are conscious. It comes after moral judgments, and it aims to affect other people (Haidt & Graham, 2007). Hence, if moral judgments depend on emotional evaluations, then belief and emotion should be correlated for moral assertions.

An antithesis is in theories of *moral grammar* (Hauser, 2006a; Mikhail, 2000, 2011). They postulate that emotional reactions are a consequence of moral judgments, not their cause. Like Chomsky's (e.g., Chomsky, 1986) universal grammar:

Ordinary individuals possess a complex moral grammar that enables them to judge the deontic status of actions in a manner roughly analogous to how native speakers intuitively recognize the grammaticality of sentences.

(Mikhail, 2011, p. 309)

The universal moral grammar consists of unconscious rules that guide moral intuitions, but culture sets the values of its parameters. The rules have nothing to do with emotions, which play no causal role in moral judgments (Hauser, 2006b; Mikhail, 2011).

Syntheses of the two preceding extremes can be found in “dual process” theories postulating that moral judgments reflect reasoning or emotions (see, e.g., Baron, 2008, Ch. 16; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Paxton, Ungar, & Greene, 2011). Our theory is also a dual-process one. It proposes that deontic judgments, which include those concerning morality, rely on reasoning (Bucciarelli et al., 2008). But, reasoning itself is a dual process – a hypothesis that the late Peter Wason pioneered (see, e.g., Johnson-Laird & Wason, 1970, for an algorithmic account; Johnson-Laird, 2006): it relies either on intuitions, which make no use of counterexamples, or on deliberations, which do. The theory assumes that reasoning can be intuitive, and that such intuitions can yield moral evaluations and emotions. No special sort of reasoning is needed for deontic topics, just normal everyday inferences (cf. Cushman & Young, 2011). In its principle of

independent systems, our theory postulates that emotions and reasoning rely on separate systems, though emotions are evolutionarily older. Experimental results corroborate this principle: people can judge some actions to be moral or immoral faster than they can evaluate them emotionally, but they can evaluate other actions emotionally before they can judge their morality (Bucciarelli et al., 2008). Emotions and reasoning run in parallel but they interact, and so they should have causal effects on each other.

The socio-intuitionist theory predicts that a change in an emotion about a moral assertion should change the degree of belief in the assertion. Moral grammars predict that a change in a belief about a moral assertion should change the emotion about the assertion. Only the present dual-process theory, however, predicts both these causal relations. It implies that an individual's complex emotion about a moral assertion and degree of belief in the assertion should mutually influence one another (see, e.g., Bucciarelli & Daniele, 2015; Bucciarelli & Johnson-Laird, 2005; Bucciarelli et al., 2008).

The first aim of the present investigation was to examine the relation between degrees of belief in assertions, factual and moral, and the emotional reactions they evoke. We carried out three experiments designed to examine these relations. Its second aim was to establish causal relations from degrees of belief in moral assertions to emotional reactions to them, and the converse causal relations from emotional reactions to degrees of belief. The dual-process theory predicts that a change in one of these factors – either belief or emotion – should cause a change in the other. All our experiments had the approval of the Ethical Committee of the University of Turin.

2. The correlations between emotions and beliefs

2.1. Experiment 1

2.1.1. Method

Our first experiment examined one group of participants' degrees of belief in a set of moral and factual assertions, and another group of participants' emotional reactions to the same moral and factual assertions. The participants were randomly assigned to one of the two groups. In each group, the moral and factual assertions were in two different blocks, with counterbalanced orders of presentation of the two blocks. The assertions in each block were in a different random order for each participant. The participants rated their degrees of belief in an assertion using a five-point Likert scale of subjective probability:

Impossible–improbable–as probable as not–probable–certain.

The participants rated their emotional reactions to an assertion also using a five-point Likert scale from loathing to liking:

I loathe this idea–I don't like this idea–I am indifferent to this idea–I like this idea–I love this idea.

2.1.1.1. Participants. Twenty university students took part in the experiment (19 females and 1 male with a mean age of 23.55 years, $SD = 1.73$). As in all the present experiments, they were taking a course on general psychology at the University of Turin, they gave their informed consent, and they volunteered in exchange for course credits. They were all native speakers of Italian, which was the language of the studies.

2.1.1.2. Procedure. The participants were tested together in a quiet room. The key instructions to the participants in the belief group were: “Your task is to assign a probability to each assertion on a five-point scale. It should reflect how strongly you believe in the assertion.” The key instructions to the participants in the emotion group were: “Your task is to evaluate your emotional reaction to each assertion on a five-points scale”.

2.1.1.3. Materials. To develop the materials, we carried out a preliminary study in which a panel of 13 students from the same

population as those in the experiment proper created factual and moral assertions that differed in both the pleasantness of the emotions they evoked and in the degree to which they were believable. They each had to generate 9 assertions referring to “moral principles” in three categories of believability: strongly believable, fifty-fifty, unbelievable, and in three categories of emotional pleasantness: loathe, indifferent, love. Hence, each member of the panel created one assertion in each of the 9 categories. From the resulting assertions, we chose those that made sense, and the result was a set of 34 moral assertions and a set of 33 factual assertions (see Tables S1 and S2 in the Supplementary Materials). Examples of strongly believable moral assertions from the respective categories of loathing, indifference, and loving, are:

Torturing children should be a capital crime.

Young people should respect their elders.

Prison should re-educate criminals.

Examples of strongly believable factual assertions from the respective categories of loathing, indifference, and loving, are:

Excessive eating leads to gross obesity.

Car-pooling helps to limit pollution.

Humanitarian organizations save a lot of lives.

As these examples illustrate, the moral assertions contained a modal verb (based on “dovere” in Italian, which corresponds to “should” or “ought to” in English) and the factual assertions were generalizations in the present tense.

2.1.2. Results and discussion

Fig. 1 is a scattergram showing the correlation between the mean ratings of the two groups for the moral assertions: one group rated the believability of the assertions on a five-point scale and the other group rated their emotional reactions on a five-point scale. Fig. 2 is a scattergram of the same sort for the two groups' ratings of the factual assertions. As the Figures show, the ratings of beliefs and emotions were reliably correlated for the 34 moral assertions (Kendall's $\tau = 0.68$, $p < .0001$), but not for the 33 factual assertions (Kendall's $\tau = 0.12$, $p > .1$). Because the scales may not be “interval” ones, in which equal distances from one point to another have the same significance over the whole scale, we used Kendall's tau, which is sensitive only to the rank orders of items, not to the magnitudes between them. Its values range

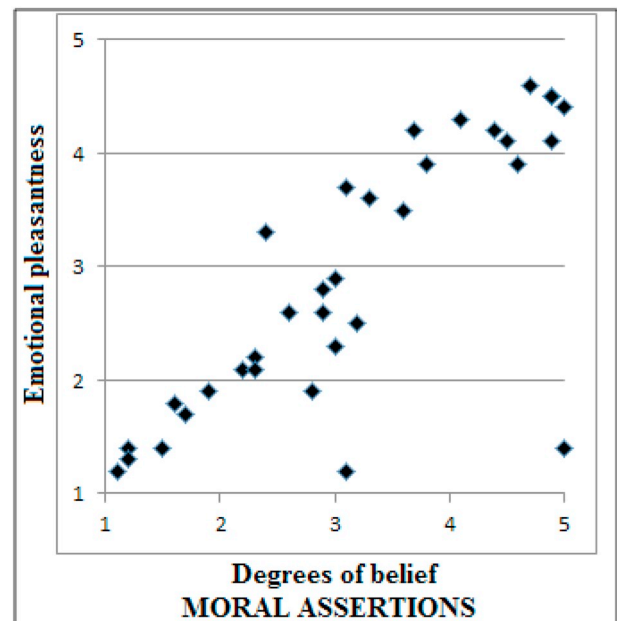


Fig. 1. Scattergram from Experiment 1 of the mean degrees of belief in 34 moral assertions for one group and the mean pleasantness of the emotional reactions to them for the other group.

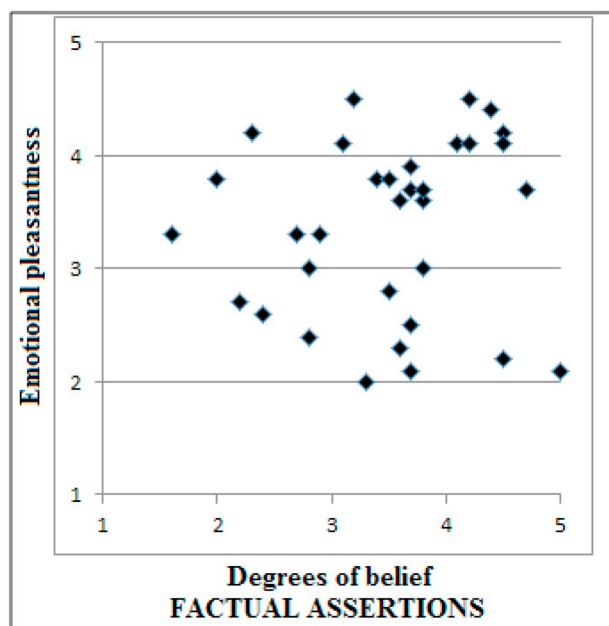


Fig. 2. Scattergram from Experiment 1 of the mean degrees of belief in 33 factual assertions for one group and the mean pleasantness of the emotional reactions to them for the other group.

from -1 (perfect inverse correlation) through 0 (no correlation) to 1 (perfect correlation).

The difference between the value of tau for factual assertions and the value of tau for moral assertions was highly reliable. For more than ten pairs in a correlation, tau's distribution approximates to the normal variate, z . The proportion of cases in the normal distribution for a tau of at least 0.1 yielding a tau of at least 0.67 equals $0.000000013/0.206$ ($p < .00000006$). Likewise, a resampling procedure confirmed the difference: 100,000 reassignments of the factual rankings yielded the chance probability that from over 20,000 cases with at least a tau of 0.1 none of them yielded a correlation at least as large as 0.67 ($p < .00005$).

The two outliers in Fig. 1 with low emotional ratings and higher degrees of belief are for the assertions:

According to ISIS, Muslims should behead non-Muslims.
According to Hitler, Jews should be exterminated.

In reality, they make factual claims about the beliefs of individuals, and so it was a mistake to include them in the set of moral assertions. When we dropped them from the analysis, the correlation between believability and emotion was even greater (Kendall's $\tau = 0.82$, $p < .0001$). Tables S1 and S2 in Supplementary Materials present the participants' mean ratings of belief and emotion for each moral assertion and for each factual assertion.

The participants' ratings for degrees of belief in the 34 moral assertions had a high level of agreement (Kendall's coefficient of concordance, $W = 0.78$, $p < .001$), and so too did their ratings for emotions for the moral assertions ($W = 0.70$, $p < .001$). Their ratings for the 33 factual materials were not quite so much in agreement, but they were reliable both for degrees of belief ($W = 0.58$, $p < .001$) and for their emotions ($W = 0.47$, $p < .001$).

The results show that if individuals believe a moral assertion then they tend to like it; likewise, if they do not believe a moral assertion then they tend to dislike it. The converse relations hold too. But, no such relations occur with factual assertions. One possible artifact is that the range of emotions was greater for moral assertions (1.0 to just over 4.5) than for factual assertions (2.0 to just over 4.5). Perhaps, if the

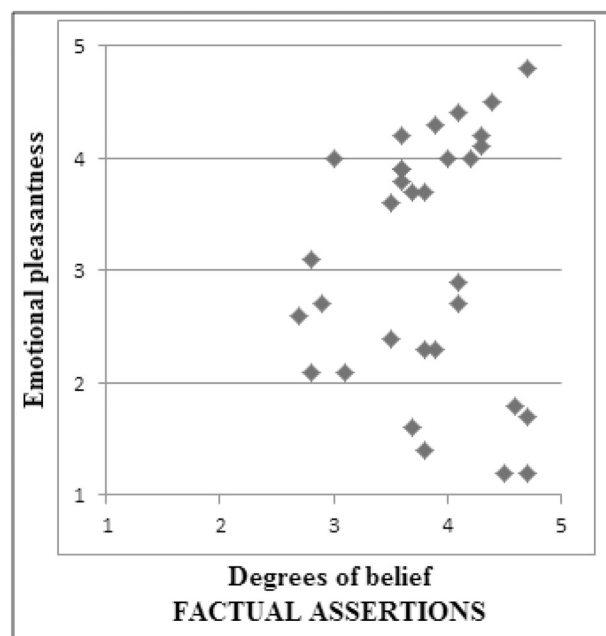


Fig. 3. Scattergram from a replication of Experiment 1 of the participants' mean degrees of belief in 32 factual assertions and the mean pleasantness of their emotional reactions to them.

factual materials varied over a broader range of pleasantness, then they too would elicit a correlation.

To test this conjecture, we carried out a replication of Experiment 1, using factual materials selected to elicit a broader range of emotions (for a total of 32 factual assertions), and using a within-participants design in which 20 university students rated the factual assertions for both belief and emotion. Fig. 3 is the resulting scattergram showing the lack of correlation between the two sorts of rating with a wider range in the ratings of emotions (from 1.0 to just under 5.0). No reliable correlation occurred between the two ratings (Kendall's $\tau = 0.09$, $p > .4$).

Another potential artifact is that the moral and factual assertions had different contents, and the choice of factual assertions may have happened to yield contents that did not yield a correlation between belief and emotion. Our next experiment therefore examined the issue in a more stringent way. The moral and factual assertions had almost identical contents.

2.2. Experiment 2

2.2.1. Method

The participants acted as their own controls: they rated their degrees of belief in moral and factual assertions and their emotional reactions to them. The two sets of assertions were based on the same contents. Here are two examples of the matched pairs, in which the moral assertion occurs first:

- 1) Some public and private schools should offer the same opportunities.
Some public and private schools offer the same opportunities.
- 2) Immigrants who do not find a job ought to go back to their country.
Immigrants who do not find a job tend to go back to their country.

Each participant rated one member of each of 30 such pairs of assertions. In one block they rated their beliefs, in another block they rated their emotions, and the order of the two blocks was counterbalanced. Half of the assertions were moral and half of them were factual. The participants were assigned at random to one of two sets of counterbalanced blocks of assertions, and half of them had the moral

assertions first and half of them had the factual assertions first. The order of the assertions in all blocks was randomized for each participant.

2.2.1.1. Participants. Twenty university students from the same population as before took part in the experiment (12 females and 8 males with a mean age of 24.55 years, $SD = 3.49$).

2.2.1.2. Procedure. The procedure and instructions were based on those in Experiment 1 except that each participant received the instructions for rating both believability and emotional reactions.

2.2.1.3. Materials. The assertions were constructed from the moral assertions in Experiment 1 by adding a parallel factual version that was plausible. But, in order to have assertions that varied equally in believability in both versions, we made slight edits to the original materials. We also added 11 new pairs of assertions in order to cover some topics not included in the original set. Table S3 in Supplementary Materials states the full set of 30 matched pairs of assertions. The materials were assembled into two different sets of assertions: one set had the odd numbered moral assertions (in Table S3) and even numbered factual assertions (in Table S3), and the other set had the opposite assignments.

2.2.2. Results and discussion

Fig. 4 is a scattergram showing the correlation between the mean ratings for the degrees of belief in the 30 moral assertions on a five-point scale and the mean ratings for their emotional reactions on a five-point scale. Fig. 5 is a scattergram of the same sort for the factual assertions. As the two Figures show, the two sets of ratings correlated significantly for the moral assertions (Kendall's $\tau = 0.82$, $p < .0001$), but not for the factual assertions (Kendall's $\tau = 0.23$, $p = .085$).

The difference between the tau for factual assertions and the tau for moral assertions was highly reliable. The proportion of cases in the normal distribution with a tau of at least 0.21 yielding a tau of at least 0.83 equals $0.0000003/0.052$ ($p < .000001$). The resampling procedure confirmed the difference: 100,000 reassignments of the factual ranks showed from over 6000 cases with a tau of at least 0.21 none of them yielded a correlation at least as large as 0.83 ($p < .0005$).

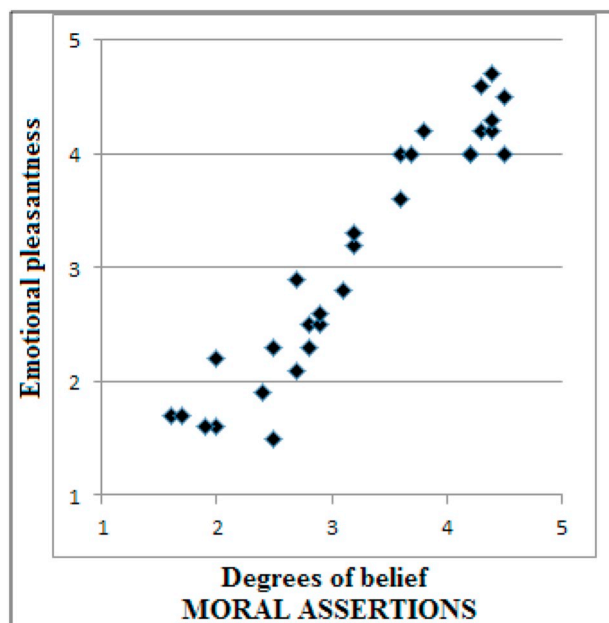


Fig. 4. Scattergram from Experiment 2 of the participants' mean degrees of belief in 30 moral assertions and the mean pleasantness of their emotional reactions to them.

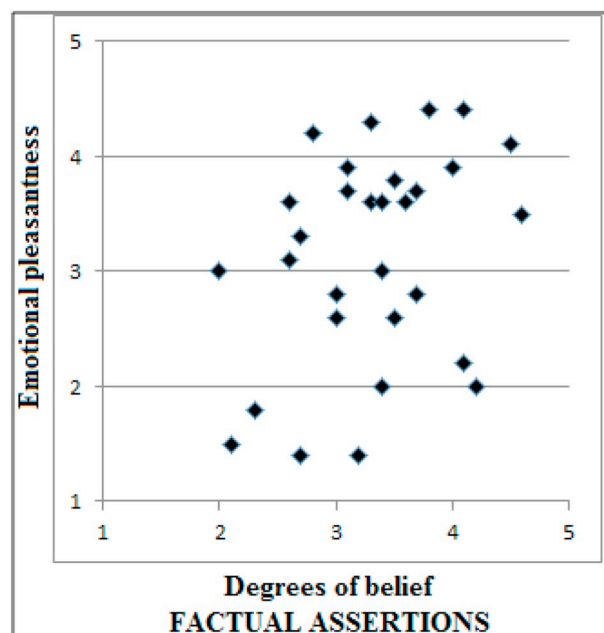


Fig. 5. Scattergram from Experiment 2 of the participants' mean degrees of belief in 30 factual assertions and the mean pleasantness of their emotional reactions to them.

Table S3 in Supplementary Materials presents the pairs of moral and factual assertions and the participants' mean ratings for belief and emotions. We examined the degree to which the participants agreed in their ratings. The concordances were reliable for the believability of the 30 moral assertions (Kendall's $W = 0.50$, $p < .001$) and for their emotional reactions ($W = 0.65$, $p < .0001$). Their ratings for the 30 factual materials were not so much in agreement, but they were reliable for believability ($W = 0.38$, $p < .001$) and for emotional reactions ($W = 0.58$, $p < .0001$).

The results replicated those from Experiment 1, but in this case the two sorts of assertion have almost identical contents, and the participants rated both their pleasantness and their believability. We conclude that a genuine difference exists between moral and factual assertions. Only the moral assertions yield a robust correlation between the believability of an assertion – its subjective probability – and the emotion it evokes – from loathing to loving. Because emotions do not correlate with beliefs for factual assertions, we pursued their investigation no further.

2.3. Experiment 3

The experiment concerned only moral assertions and so it allowed us to assess the critical correlation for a larger set of moral assertions about a wider set of topics.

2.3.1. Method

The experiment included all the assertions in the previous two studies and some additional assertions. The participants rated their degrees of belief in and their emotional reaction to each of 48 moral assertions in two blocks of trials in counterbalanced orders.

2.3.1.1. Participants. Twenty university students from the same population as before took part in the experiment (12 females and 8 males with a mean age of 23.00 years, $SD = 1.38$).

2.3.1.2. Material and procedure. The materials consisted of moral assertions from Experiments 1 and 2, and 9 additional assertions in order to cover a set of comprehensive topics: the 48 moral assertions are

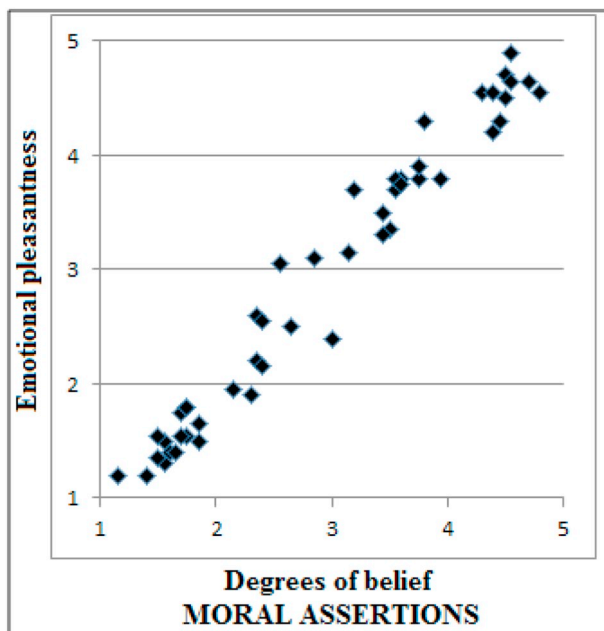


Fig. 6. Scattergram from Experiment 3 of the participants' mean degrees of belief in 48 moral assertions and the mean pleasantness of their emotional reactions to them.

presented in Table S4 in Supplementary Materials. The procedure and instructions were identical to those in the previous experiment.

2.3.2. Results and discussion

Fig. 6 is a scattergram showing the correlation between the mean of the participants' degrees of belief in the 48 moral assertions on a five-point scale and the mean of their emotional reactions on a five-point scale. As the Figure shows, the two judgments were very highly correlated (Kendall's $\tau = 0.90$, $p < .0001$). Table S4 in Supplementary Materials presents the moral assertions and the participants' mean ratings of belief in them and mean emotional reactions to them. The concordances over the participants' degrees of belief in the 48 moral assertions (Kendall's $W = 0.61$, $p < .0001$) and over their emotional reactions ($W = 71$, $p < .0001$) were highly reliable.

The present experiment corroborated the prediction that the degrees to which people believe moral assertions correlate with their emotional reactions to them.

3. The causal relations for moral assertions between beliefs and emotions

Our two final studies aimed, first, to change the participants' emotional reactions to moral assertions and to see whether it changed their degrees of belief in them (Experiment 4), and, second, to change participants' degrees of belief in moral assertions and to see whether it changed their emotional reactions to them (Experiment 5).

Previous studies have shown that emotions can cause changes in moral judgements. Individuals exposed to a disgusting smell make more severe moral judgments than control participants (e.g., Schnall, Haidt, Clore, & Jordan, 2008). Positive and negative moods that music induces also alter moral judgments (e.g., Pastötter, Gleixner, Neuhauser, & Bäuml, 2013). But, our study also aimed to demonstrate the converse relation, and so it would have been inappropriate to use a procedure in which emotions are general instead of reactions to particular assertions. That is because the dual-system theory does not predict that changes in the believability of a moral assertion should affect a person's general mood. The appropriate procedure therefore calls for an examination of how changes in emotions to a particular assertion affect its

believability, and vice versa.

3.1. Experiment 4: Changes in emotions to moral assertion CHANGE beliefs in them

A feasible way to change individuals' emotions about a moral assertion is to ask them to recall an autobiographical memory that evokes a pleasant memory or else an unpleasant memory pertinent to the moral assertion. Mills and D'Mello (2014) showed that such memories can induce positive or negative emotions. The dual-process theory predicts that if individuals like a moral assertion more, they should believe it more; but, if they dislike the assertion more, they should believe it less.

3.1.1. Method

The participants carried out three tasks for each of six moral assertions presented in a different random order to each of them. First, they rated their degrees of belief in each assertion using a 21-point Likert scale of subjective probability ranging from *impossible* to *certain* on 5 major points, with four intervening points between adjacent labels to make a more sensitive measure likely to detect changes. Second, after rating all the assertions in this way, the participants recalled an episode from their lives that concerned a moral assertion. Third, after this task for first assertion, they rated again their degree of belief in the assertion. They repeated the second and third tasks for each of the remaining assertions. For half the participants, the autobiographical episode for the first three assertions was for a pleasant memory, and for the second three assertions it was for an unpleasant memory; and for the other half of the participants, the assignment of the two sorts of episode was in the opposite order. The participants were assigned at random to one of the two preceding orders.

3.1.1.1. Participants. Thirty-two students from the same population as before took part in the experiment (16 females and 16 males with a mean age of 25.53 years, $SD = 5.97$).

3.1.1.2. Materials and procedure. We chose six moral assertions varying in their degree of believability according to Experiments 1, 2 and 3. These six assertions (see Table S5 in Supplementary Materials) were assigned to two sets in order to counterbalance their assignment to the pleasant and unpleasant memory tasks. In set 1, assertions 1, 2, 3, were assigned to the unpleasant memory task, and assertions 4, 5, and 6, were assigned to the pleasant memory task. In set 2, the assignments were swapped around.

The participants were tested individually in a quiet room. The key instructions for rating the believability of assertions, both en masse before the memory task and individually after each memory task, were as follows:

Your task is to assign a probability to each assertion on a scale ranging from *Impossible* to *Certain*. This probability should reflect the degree to which you believe the assertion.

The key instructions for the two memory tasks were as follows:

For each assertion, your task is to write down a specific episode in your life in which what happened was an event related to the assertion. Do not recall episodes that occurred repeatedly in a period of life: describe in details a specific episode that is located at a particular time in a particular place. The episode must be very pleasant (very unpleasant). If you don't have such a memory, then imagine such an episode that is a very pleasant (unpleasant) episode and write it down.

3.1.2. Results and discussion

Two independent judges checked that each memory was of a specific episode, that it was relevant to the moral assertion, and that it evoked the appropriate emotion. They agreed in their coding on 97% of

Table 1

Mean ratings of the believability on a five-point scale (with 95% confidence intervals) before and after the pleasant memory task and before and after the unpleasant memory task in Experiment 4.

	Rating before	Rating after	Overall
Pleasant memory	3.16 [2.81, 3.51]	3.51 [3.16, 3.86]	3.33 [2.99, 3.67]
Unpleasant memory	3.03 [2.68, 3.38]	2.86 [2.52, 3.20]	2.95 [2.62, 3.28]
Overall	3.05 [2.92, 3.18]	3.15 [3.00, 3.30]	3.10 [2.96, 3.24]

items (Cohen's $K = 0.90$, $p < .001$). For the final score, they discussed each item on which they disagreed, until they reached agreement. They judged that the participants performed the pleasant memory task correctly on 91% of trials and the unpleasant memory task correctly on 84% of trials. The most common failure was to recall an episode in the unpleasant memory task but for the negation of the moral assertion, e.g., for the assertion, *you should be generous in your life*, a participant recalled an unpleasant memory in which she was not generous. Such failures may explain the smaller effects of unpleasant memories (see Table 1). We excluded from analysis the trials in which the participants failed the memory task.

The order of the two sorts of memory task had no reliable effect on the ratings of the assertions either after the pleasant memory task (Mann-Whitney test, $z = 0.03$, $p > .25$, Cliff's $\delta = 0.24$), or after the unpleasant memory task (Mann-Whitney test, $z = 1.72$, $p > .08$, Cliff's $\delta = 0.15$). Hence, we combined the results from the two groups of participants for the subsequent analyses.

Fig. 7 presents the ratings of belief for each of the six moral assertions before and after the two memory tasks. As it shows, there was an interaction in the believability of the assertions: it tended to increase after a pleasant memory and to decrease after an unpleasant memory (Wilcoxon test by participants, $z = 3.7$, $p < .0001$, one tail, Cliff's $\delta = 0.70$; and all six contents showed the predicted interaction (Binomial test, $p < .02$, one tail). The mean believability of the assertions increased reliably after the memory evoking a pleasant emotion (Wilcoxon test by participants: $z = 4.18$, $p < .00002$, Cliff's $\delta = 0.22$; and by-materials, Binomial, $p < .025$), and decreased reliably after the memory evoking an unpleasant emotion (Wilcoxon, $z = 2.30$, $p < .015$, Cliff's $\delta = 0.08$, but not reliably in a by-materials analysis, Binomial test, $p > .1$). With hindsight, we might have detected greater effects of the memory tasks if we had chosen moral assertions with mean degrees of belief in Experiments 1, 2 and 3. Those at one or other end of the scale leave less room for an effect on ratings.

Overall, the effects of autobiographical memories tended to be

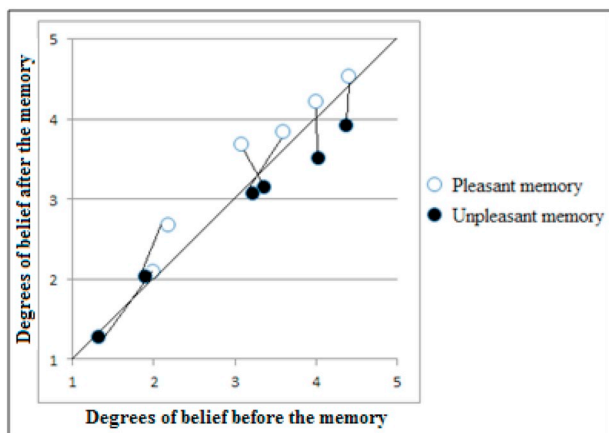


Fig. 7. For Experiment 4, the participants' mean degrees of belief in each of the six moral assertions before and after the recall of pleasant or unpleasant memories related to them. The lines connecting pairs of black and white circles pick out the same moral assertions in the two memory conditions.

larger for memories that were pleasant than for memories that were unpleasant, but the difference was not reliable (Wilcoxon test, $z = 1.54$, $p > .12$, two tail, Cliff's $\delta = 0.25$). Table 1 presents the mean ratings of the believability of the six moral assertions before and after the two sorts of autobiographical memory (the detailed results for the six assertions are in Table S5 in Supplementary Materials).

The results corroborated the prediction. The effects were modest, but robust: a change to the emotion that a moral assertion evokes changed its believability. A common experience in everyday life is that the recall of an episodic memory can evoke emotional feelings, and not just cognitive evaluations of emotions appropriate to the occasion (see also Mills & D'Mello, 2014). Nevertheless, skeptics might argue that the consequent shifts in believability were the result of the demand characteristics of the experiment (e.g., Nichols & Maner, 2008), which the participants inferred was to increase their belief in a moral assertion after remembering a pleasant personal event, and to decrease their belief in a moral assertion after remembering an unpleasant personal event. Such an effect is rather implausible, because it would call for participants:

- To remember their initial rating prior to the recall task, when they had rated all six assertions on a 21-point Likert scale.
- To infer that because in one condition they were asked to recall a pleasant memory they should shift their rating upwards, and because in another condition they were asked to recall an unpleasant memory they should shift their rating downwards.
- To make the required shift in ratings.

The inference in the second step above depends on a belief in the particular hypothesis under test: a pleasant emotional experience related to a moral assertion ought to increase its believability, and an unpleasant emotional experience related to a moral assertion ought to decrease its believability. If individuals believed that the relation should occur for any sort of proposition, the demand characteristics of Experiment 1 and 2 should have led to a correlation between beliefs and emotions for factual assertions. But, no such correlation occurred.

A more plausible alternative hypothesis is that the contents of the autobiographical memories, not the emotions that they evoked, were the cause of the change in degrees of belief. We postpone our discussion of this possibility and one concerning response bias until the General discussion.

3.2. Experiment 5. Changes in beliefs about moral assertions CHANGE emotions to them

A potential way to change individuals' degrees of belief in an assertion – its subjective probability – is to get them to create their own reasons for believing, or else for disbelieving, the assertion. Nothing is more persuasive than a proof, but reasons can also be effective in changing beliefs (e.g., Slusher & Anderson, 1996). The present experiment therefore used this procedure to determine whether changes in degrees of belief in moral assertions also changed emotional reactions to them.

3.2.1. Method

The participants carried out three tasks for each of six moral assertions presented in different random orders. First, they rated their emotional reactions to the six assertions using a 21-point Likert scale ranging from *I loathe this idea* to *I love this idea* on 5 major points, but with four intervening points between adjacent labels to make a more sensitive measure. Second, the participants had to think up a reason for believing, or disbelieving, a moral assertion. Third, after creating such a reason, they then rated their emotional reaction to it again. They repeated the second and third tasks for each of the remaining assertions. For half the participants, the first three reasons they had to think of were for believing an assertion, and the second three reasons were for

disbelieving an assertion; for the other half of the participants, the two sorts of reason were in the opposite order. The participants were assigned at random to one of the two preceding orders.

3.2.1.1. Participants. Thirty-two university students from the same population as before took part in the experiment (30 females and 2 males with a mean age of 22.66 years, $SD = 1.41$).

3.2.1.2. Materials and procedure. We chose six moral assertions from Experiments 1, 2 and 3, in which the participants made ratings of pleasantness around the mid-point of the Likert scale, e.g.:

One should be dishonest with dishonest people.
Parents should avoid having arguments in front of their children.

The rationale was to allow a detectable change in emotional reactions after the participants had thought of a reason to change their degrees of belief. The six assertions (see Table S6 in Supplementary Materials) were assigned to two sets in order to counterbalance their assignment to reasons to believe, and reasons to disbelieve. In set 1, assertions 1, 2, 3, were assigned to the task of thinking of reasons to believe them, and assertions 4, 5, and 6, were assigned to the task of thinking of reasons to disbelieve them. In set 2, the assignments were swapped around.

The participants were tested individually. The key instructions for rating their emotional reactions to the assertions, both en masse before thinking of reasons and individually after thinking of a reason, were as follows:

Your task is to evaluate your emotional reaction to each assertion on a scale ranging from *I loathe this idea* to *I love this idea*.

The key instructions for the two tasks of thinking of reasons were as follows:

For each assertion, your task is to write down a specific reason to believe (disbelieve) in what the assertion states. If you can't think of such a reason, then imagine a reason that someone could have for believing (disbelieving) in the assertion. Write the reason down.

3.2.2. Results

Two independent judges checked that the participants produced a genuine reason rather than a paraphrase of an assertion, and that they obeyed the request to create one to believe, or to disbelieve, the assertion. The judges agreed in their coding on 99.5% of items (Cohen's $K = 0.92$, $p < .001$). They discussed each of the handful reasons on which they disagreed, until they reached agreement. They judged that the participants thought of reasons to believe the assertions on 98% of trials, and that they thought of reasons to disbelieve the assertions also on 98% of trials. We excluded the remaining trials from analysis.

The order of the two sorts of reasons had no reliable effect on emotional reactions to the assertions either after thinking of reasons to believe the assertions (Mann-Whitney test, $z = 1.72$, $p = .09$, Cliff's $\delta = 0.36$), or after thinking of reasons to disbelieve the assertions (Mann-Whitney test, $z = 0.28$, $p = .78$, Cliff's $\delta = 0.03$). Hence, we combined the results from the two groups of participants for the subsequent analyses.

Fig. 8 presents the ratings of emotional reactions to each of the six moral assertions before and after the two sorts of reasons. As Fig. 8 and Table 2 below show, a reliable interaction occurred in the emotional reactions to the moral assertions: they tended to increase after the participants thought of reasons to believe assertions and to decrease after they thought of reasons to disbelieve them (Wilcoxon test by participants, $z = 2.92$, $p < .003$, Cliff's $\delta = 0.46$; and by-materials all six showed the predicted interaction, Binomial test, $p < .025$, one tail). The emotional reactions tended to become more positive after the participants thought of a reason to believe an assertion. The effect was not reliable by participants (Wilcoxon test: $z = 1.47$, $p = .14$, Cliff's

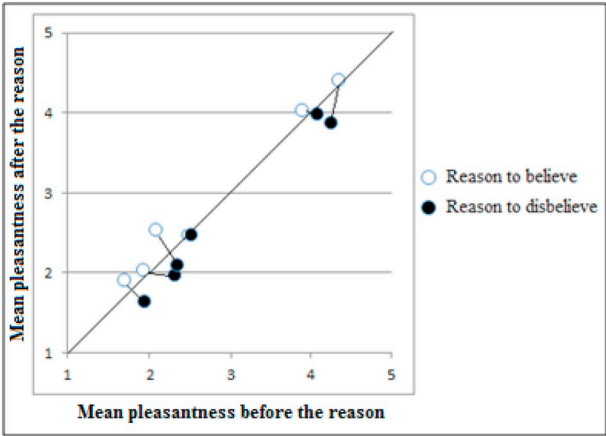


Fig. 8. In Experiment 5, the mean pleasantness of the participants' emotional reactions to the moral assertions for each of the six moral assertions before and after the participants thought of reasons to believe, or to disbelieve, the as-
sertions.

Table 2
Mean ratings in Experiment 5 of the emotional reactions on a five-point scale (with 95% confidence intervals) before and after the participants thought of reasons to believe, or to disbelieve, the six moral assertions.

	Rating before	Rating after	Overall
Reason to believe	2.75 [2.45, 3.05]	2.93 [2.62, 3.24]	2.85 [2.57, 3.13]
Reason to disbelieve	2.94 [2.65, 3.23]	2.67 [2.40, 2.94]	2.80 [2.53, 3.07]
Overall	2.85 [2.68, 3.02]	2.80 [2.66, 2.94]	2.82 [2.68, 2.96]

$\delta = 0.12$), but it was reliable by materials (Wilcoxon test: $z = 1.99$, $p = .025$, Cliff's $\delta = 0.17$). The emotional reactions tended to become more negative after the participants thought of a reason to disbelieve an assertion, and the effect was reliable both by participants (Wilcoxon, $z = 3.06$, $p = .002$, Cliff's $\delta = 0.19$) and by materials (Binomial, $p < .02$).

Table 2 presents the mean emotional reactions to the six moral as-
sertions before and after the two sorts of reasons (the results for each of the six assertions are in Table S6 in the Supplementary Materials). The results corroborated the predicted interaction. But, a reason to believe tended to increase the pleasantness of moral reaction less than a reason to disbelieve tended to increase their unpleasantness. This result, however, parallels previous observations that affirmation has less effect than denial on moral assertions (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). The conjecture that the demand char-
acteristics of the experiment are responsible for the results seems un-
likely (see our discussion of the results of Experiment 4).

4. General discussion

When a group of participants rated their degrees of belief in a set of
assertions and another group of participants rated their emotional re-
actions to the same assertions, there was a striking and hitherto un-
reported interaction. The degrees of belief correlated with the emo-
tional reactions for assertions about morality, but not for assertions
about matters of fact (Experiment 1). The difference might have re-
flected the different contents of the moral and factual assertions. So, we
used the same contents both in moral assertions, such as:

Advanced countries should be democracies.

and in corresponding factual assertions:

Advanced countries are democracies.

Participants saw just one member of each of these pairs, and beliefs correlated with emotional reactions, but again only for moral assertions and not for factual assertions (Experiment 2). The crucial correlation also occurred for a comprehensive set of 48 different moral assertions (Experiment 3).

Correlations cannot establish causality, but our subsequent studies did so and in both directions. In particular, autobiographical memories of pleasant events about a moral assertion increased the degree to which participants believed it, whereas such memories of unpleasant events decreased the degree to which participants believed it (Experiment 4). Hence, a change in emotions can cause a change in the degree to which moral assertions are believable. Skeptics might argue, as did an anonymous reviewer, that the result reflects only the demand characteristics of the experiment. But, in the discussion of Experiment 4's results, we explained that this account is rather unlikely: why should participants infer that a pleasant memory of events in their lives should call for an increase in the believability of a moral assertion? It seems that they must base their inference on the hypothesis under investigation: a pleasant emotion ought to increase their belief in a moral assertion. Even if they made such an inference, it is likely to have occurred sufficiently long after their initial ratings of believability that they would no longer have an accurate memory of six values on a 21-point scale.

At first sight, an alternative explanation seems more plausible. The contents of the autobiographical memories, not the emotions they evoked, cause the shift in the ratings of believability. This conjecture (due to Jon Baron, p.c. 11-5-2018 and to an anonymous reviewer) is that people tended to recall cases that turned out well when they had to recall a pleasant memory, and cases that turned out badly when they had to recall an unpleasant memory. The good outcomes increased the believability of the moral assertion and the bad outcomes decreased it, but as a matter of an intellectual rather than emotional reaction. However, consider a moral assertion that you happen to disbelieve, such as "You should behave as you feel like doing". Suppose, like one of our participants, you recall that once when you were dancing with a group of friends who avoided another bunch of kids, but you felt like dancing with them, and so you did. Later, you discovered that they were from a community of children with psychological problems, and so you felt happy that you had danced with them. The events but not the emotion are supposed to have increased your belief in the principle. There are several oddities with this account. First, given the same events, you might not have felt happy – so why would the events alone increase your belief in the principle? Second, you know that you could have danced with the kids for reasons other than the somewhat dubious moral principle – you could have done so out of kindness towards strangers. Third, the events in your memory are supposed to change your degree of belief in a principle whose truth is deontic not a matter of fact. A long tradition going back to Hume (1778/1739) is that such a change is impossible, because facts don't affect deontic principles – what is the case doesn't imply what *ought* to be the case (see Baron, 2008, Ch. 16; cf. Elqayam, Wilkinson, Thompson, Over, & Evans, 2017). Your memory of dancing with the kids probably led you to relive the emotions you felt when you found out about them (Conway & Pleydell-Pearce, 2000; Raes, Hermans, Williams, & Eelen, 2006), yet they are supposed to have no effect on your belief in the moral principle. Two independent judges evaluated whether the participants reported autobiographical memories consistent with the moral principle and elicited the appropriate emotional reaction. Those memories that didn't were rejected from the analysis of the results. In sum, the memory task is likely to have induced appropriate emotional reactions, and even perhaps actual feelings (Mills & D'Mello, 2014). They seemed to have caused a shift in belief in the moral assertions. Yet, we do concede that contents of the memories alone might have caused the shift. It seems unlikely, but it is possible.

A variant on this hypothesis about Experiment 4 is that the emotions in the experiment created a response bias. As the reviewer suggested, an

increase in positive feelings (or positive contents) could increase a participant's belief in the moral assertions, whereas an increase in negative feelings (or negative contents) could decrease the participant's belief in them. Insofar as this hypothesis concerns emotions, it is indistinguishable from the dual-process theory's predictions.

A converse causal relation occurred in our final experiment. The creation of a reason to believe a moral assertion increased the pleasantness of the emotions it evoked, whereas the creation of a reason to disbelieve the assertion increased the unpleasantness of the emotions it evoked (Experiment 5). The idea that demand characteristics of the experiment led to its results is implausible for the same sort of reasons that applied to Experiment 4. Likewise, it is difficult to envisage some factor other than the content of the participants' reasons leading to a change in the emotional reaction to the assertions. Their reasons could have evoked an emotional reaction of their own, which in turn caused the change in the emotions that the assertions themselves evoked. But, this hypothesis in itself would corroborate the dual-process theory's prediction that beliefs about moral assertions, which include reasons for them, can evoke emotional reactions.

In sum, emotions and beliefs correlate for moral assertions, and a change in one can cause a change in the other. The main theoretical problem is to explain these results. They should hardly surprise Utilitarians. As we mentioned in the Introduction, one interpretation of their views (Jon Baron, p.c.) is that it is tautological to predict that if you believe a moral assertion then you will like it. And this interpretation implies that our experiments are studies in semantics, which corroborate the existence of tautologies depending on the meanings of words (contra to Quine, 1953; cf. Quelhas, Rasga, & Johnson-Laird, 2017). But, the degrees to which participants believed the moral assertions varied from certain to impossible. An assertion that they rated *as probable as not* is hardly a tautology, and it tended to occur with an emotional reaction of *indifference*. The hypothesis of a tautological interpretation cannot explain this aspect of an overall correlation in ratings on scales. Moreover, the negation of a tautology is a self-contradiction, and yet there is nothing contradictory about an assertion, such as:

It's not true that if you ought to pay your taxes, then you will like doing so.

Indeed, it is common for people to violate their own moral principles for the sake of pleasure – a failing that many religions censure. Such failings would be impossible if the relation between the two were tautological. The corroboration of the relation is therefore a discovery of a systematic but contingent dependency – one that the Utilitarians believed for communities as a whole – rather than a psycholinguistic confirmation of a synonymy.

The socio-intuitionist theory of morality, as we mentioned at the outset, postulates that moral judgments are based solely on the emotions that events evoke (Shweder & Haidt, 2000; Wilson, 1993). These emotions can affect degrees of believability, but the converse is impossible on this account: the degree of believability of a moral assertion cannot affect the emotions it evokes. In short, emotions first; then beliefs. The causal effect of emotions on beliefs in moral assertions in Experiment 4 is consistent with the theory, but the converse causal effect in Experiment 5 is inconsistent with the theory.

Moral grammars, as we also mentioned at the outset, postulate that a set of innate rules unconsciously guides individuals' evaluative intuitions about moral assertions (Hauser, 2006b; Mikhail, 2011). The degree of believability of a moral assertion can affect the emotion it evokes, but the converse is impossible: the emotion evoked by a moral assertion cannot affect its believability. In short, beliefs first; then emotions. The causal effect of beliefs on emotions from moral assertions in Experiment 5 is consistent with the theory, but the converse causal effect in Experiment 4 is inconsistent with the theory.

The present dual-process theory postulates that reasoning underlies moral reactions, either rapid, largely unconscious, and intuitive

inferences, or else slower, more conscious, deliberations (Bucciarelli et al., 2008). Basic emotions depend on simple cognitive evaluations, which can prepare individuals for generic reactions, but complex emotions, which incorporate propositional content unique to human beings, are needed to evaluate moral assertions (Oatley & Johnson-Laird, 1987, 2014). Emotions and reasoning are therefore separate systems that run in parallel but that can affect each other. Beliefs can elicit emotions and change them, and emotions can elicit beliefs and change them. In short, either beliefs or emotions first, and then emotions or beliefs. What is unique to deontic assertions is a strong cultural consensus about the pleasing nature of what is right and the hateful nature of what is wrong. As our results showed, there is no such reliable consensus about matters of fact. That the consensus about morality is cultural is illustrated in such cases as human sacrifice, cannibalism, suicide, abortion, and polygamy. In some societies, these practices are morally right; in other societies, they are morally wrong. Hence, beliefs in moral assertions can affect the emotions they evoke, and the emotions they evoke can affect beliefs in them (Bucciarelli et al., 2008; see also Paxton et al., 2011). The theory elucidates – indeed, predicts – the phenomena in our experiments. An open question is whether these phenomena occur only for moral assertions, or else for other deontic assertions about social conventions, rules, and manners. Another open question is whether emotions, and in particular complex ones with propositional content, can cause a change in the content of a person's beliefs about a deontic assertion.

In conclusion, moral assertions yield a correlation between what is good and what is pleasant, and between what is bad and what is unpleasant. To enhance the morality of an assertion can make it more pleasant; and to diminish its morality can make it less pleasant. Conversely, to increase the pleasantness of a moral assertion can make it morally better; and to decrease its pleasantness can make it morally worse. Hence, emotions and reasoning, which depend on separate systems, can exert causal effects on each other. A major nexus concerns morality, for which societies establish a consensus reflected in the Epicurean maxim that what is good is pleasant.

Acknowledgments

The Supplementary Materials and the data reported in this paper are archived at the following database: <https://osf.io/g34j6/>. We thank Bruno Bara, Mo Johnson-Laird, Keith Oatley, and Sangeet Khemlani, for helpful advice. And we thank Jon Baron and two anonymous reviewers for helpful comments on earlier drafts of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.actpsy.2019.102880>.

References

- Baron, J. (2008). *Thinking and deciding* (4th ed.). New York: Cambridge University Press.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5, 323–370. <https://doi.org/10.1037/1089-2680.5.4.323>.
- Bentham, J. (1789). An introduction to the principles of morals and legislation. In J. H. Burns, & H. L. A. Hart (Eds.). *The collected works of Jeremy Bentham* (Oxford: Clarendon Press 1996).
- Blair, R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, 57, 1–29. [https://doi.org/10.1016/0010-0277\(95\)00676-P](https://doi.org/10.1016/0010-0277(95)00676-P).
- Bucciarelli, M., & Daniele, M. (2015). Reasoning in moral conflicts. *Thinking & Reasoning*, 21, 265–294. <https://doi.org/10.1080/13546783.2014.970230>.
- Bucciarelli, M., & Johnson-Laird, P. N. (2005). Naïve deontics: A theory of meaning, representation, and reasoning. *Cognitive Psychology*, 50, 159–193. <https://doi.org/10.1016/j.cogpsych.2004.08.001>.
- Bucciarelli, M., Khemlani, S., & Johnson-Laird, P. N. (2008). The psychology of moral reasoning. *Judgment and Decision making*, 3, 121–139. [https://doi.org/10.1016/0010-0277\(95\)00676-P](https://doi.org/10.1016/0010-0277(95)00676-P).
- Chomsky, N. (1986). *Knowledge of language: Its nature, origins and use*. New York, NY: Praeger.
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107, 261–288. <https://doi.org/10.1037/0033-295X>.
- Cosmides, L., & Tooby, J. (1996). Are human good intuitive statisticians after all? Rethinking some conclusions of the literature on judgment under uncertainty. *Cognition*, 58, 1–73. [https://doi.org/10.1016/0010-0277\(95\)00664-8](https://doi.org/10.1016/0010-0277(95)00664-8).
- Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science*, 35, 1052–1075. <https://doi.org/10.1111/j.1551-6709.2010.01167.x>.
- Downey, J. E. (1897). A musical experiment. *The American Journal of Psychology*, 9, 63–69. <https://doi.org/10.2307/1411476>.
- Elqayam, S., Wilkinson, M. R., Thompson, V. A., Over, D. E., & Evans, J. S. B. T. (2017). Utilitarian moral judgment exclusively coheres with inference from is to ought. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.01042>.
- de Finetti, B. (1937). La prevision: Ses lois logiques, ses sources subjectives. In Ann. Inst. H. Poincaré, 7, pp 1–68. Translated into English in H. E. Kyburg Jr, & H. E. Smokler (Eds.), *Studies in subjective probability*. New York: Wiley, 1964.
- Fontanari, L., Gonzalez, M., Vallortigara, G., & Girotto, V. (2014). Probabilistic cognition in two indigenous Maya groups. *Proceedings of the National Academy of Science*, 111, 17075–17080. <https://doi.org/10.1073/pnas.1410583111>.
- Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (1995). “A 30% chance of rain tomorrow”: How does the public understand probabilistic weather forecast? *Risk Analysis*, 25, 623–629. <https://doi.org/10.1111/j.1539-6924.2005.00608.x>.
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, 306, 496–499. <https://doi.org/10.1126/science.1094492>.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–2108. <https://doi.org/10.1126/science.1062872>.
- Haidt, J. (2001). The emotional dog and its rational tail. *Psychological Review*, 108, 814–834. <https://doi.org/10.1037/0033-295X.108.4>.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316, 998–1002. <https://doi.org/10.1126/science.1137651>.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20, 98–116. <https://doi.org/10.1007/s11211-007-0034-z>.
- Hauser, M. D. (2006a). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Harper Collins.
- Hauser, M. D. (2006b). The liver and the moral organ. *Social Cognitive and Affective Neuroscience*, 1, 214–220. <https://doi.org/10.1093/scan/nsl026>.
- Hume, D. (1978). *A treatise of human nature*. (2nd ed.). Oxford: Oxford University Press. (Originally published 1739).
- Johnson-Laird, P. N. (2006). *How we reason*. Oxford: Oxford University Press.
- Johnson-Laird, P. N., & Oatley, K. (2000). Cognitive and social construction in emotions. In M. Lewis, & J. M. Haviland-Jones (Eds.) *Handbook of Emotions*, (2nd ed.). (pp. 458–475). New York: Guilford Press.
- Johnson-Laird, P. N., & Oatley, K. J. (2016). Emotions in music, literature, and film. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.). *Handbook of Emotions*. (4th ed.) (pp. 82–97). New York: Guilford Press.
- Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, 1, 134–148. [https://doi.org/10.1016/0010-0285\(70\)90009-5](https://doi.org/10.1016/0010-0285(70)90009-5).
- Keltner, D., Ekman, P., Gonzaga, G. C., & Beer, J. (2003). Facial expression of emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.). *Handbook of affective sciences* (pp. 415–432). New York: Oxford University Press.
- Khemlani, S., Lotstein, M., & Johnson-Laird, P. N. (2012). The probabilities of unique events. *PLoS One*, 7(10), e45975. <https://doi.org/10.1371/journal.pone.0045975>.
- Khemlani, S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naïve probability: Model-based estimates of unique events. *Cognitive Science*, 39, 1216–1258. <http://dx.doi.org/10.1111/cogs.12193>.
- Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, 51, 336–353. <https://doi.org/10.1016/j.brainres.2009.05.054>.
- Mikhail, J. (2000). *Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in a theory of justice*. Ph.D. dissertation. Cornell University.
- Mikhail, J. (2011). *Elements of moral cognition*. Cambridge, MA: Harvard University Press.
- Mills, C., & D'Mello, S. (2014). On the validity of the autobiographical emotional memory task for emotion induction. *PLoS One*, 9(4), e95837. <https://doi.org/10.1371/journal.pone.0095837>.
- Nichols, A. L., & Maner, J. K. (2008). The good subject effect: Investigating participant demand characteristics. *Journal of General Psychology*, 135, 151–165. <https://doi.org/10.3200/GENP.135.2.151-166>.
- Oatley, K. J., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion*, 1, 29–50. <https://doi.org/10.1080/02699938708408362>.
- Oatley, K., & Johnson-Laird, P. N. (1996). The communicative theory of emotions: Empirical tests, mental models, and implications for social interaction. In L. L. Martin, & A. Tesser (Eds.). *Striving and feeling: Interactions among goals, affect, and self-regulation* (pp. 363–393). Mahwah, NJ: Lawrence Erlbaum Associates.
- Oatley, K. J., & Johnson-Laird, P. N. (2014). Cognitive approaches to emotions. *Trends in Cognitive Sciences*, 18, 134–140. <https://doi.org/10.1016/j.tics.2013.12.004>.
- Pastötter, B., Gleixner, S., Neuhauser, T., & Bäuml, K.-H. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition*, 126, 373–377. <https://doi.org/10.1016/j.cognition.2012.11.003>.
- Paxton, J. M., Ungar, L., & Greene, J. D. (2011). Reflection and reasoning in moral judgment. *Cognitive Science*, 36, 163–177. <https://doi.org/10.1111/j.1551-6709.2019.102880>.

- 2011.01210.x.
- Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2017). A priori true and false conditionals. *Cognitive Science*, 41, 1003–1030. <https://doi.org/10.1111/cogs.12479>.
- Quine, W. V. O. (1953). Two dogmas of empiricism. In *From a logical point of view*. (pp. 20–46). Cambridge, MA: Harvard University Press. (Originally published, 1951.)
- Raes, F., Hermans, D., Williams, J. M. G., & Eelen, P. (2006). Reduced autobiographical memory specificity and affect regulation. *Cognition and Emotion*, 20, 402–429. <https://doi.org/10.1080/02699930500341003>.
- Ramsey, F. P. (1990). Truth and probability. In Mellor, D.H. (Ed.) *F.P. Ramsey: Philosophical papers*. Cambridge: Cambridge University Press. (pp. 52–109) (Originally published in 1926).
- Rozin, P. (1996). Towards a psychology of food and eating: From motivation to module to model to marker, morality, meaning, and metaphor. *Current Directions in Psychological Science*, 5, 18–24. <https://doi.org/10.1111/1467-8721.ep10772690>.
- Russell, B. (1912). *Problems of philosophy, reprinted at Oxford: Oxford University press 1971 (1912)*.
- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096–1109. <https://doi.org/10.1177/0146167208317771>.
- Shweder, R. A., & Haidt, J. (2000). The cultural psychology of the emotions: Ancient and new. In M. Lewis, & J. Haviland (Eds.). *Handbook of emotions*. (2nd ed.) (pp. 397–414). New York, NY: Guilford.
- Slusher, M. P., & Anderson, C. A. (1996). Using causal persuasive arguments to change beliefs and teach new information: The mediating role of explanation availability and evaluation bias in the acceptance of knowledge. *Journal of Educational Psychology*, 88, 110–122. <https://doi.org/10.1037/0022-0663.88.1.110>.
- Wilson, J. Q. (1993). *The moral sense*. New York, NY: Free Press.