ARTICLE – SPECIAL SECTION

emotion review

Emotion Review
1–9

ISSN 1754-0739
DOI: 10.1177/1754073916679005
er.sagepub.com

Do Emotions Cause Actions, and If So How?

Andrea Scarantino
Department of Philosophy and the Neuroscience Institute, Georgia State University, USA

Abstract

The main purpose of this article is to consider two of the most popular arguments offered in support of the view that emotions do not cause actions. One argument suggests that emotions come after actions and therefore cannot cause them (*argument from timing*). The other argument suggests that emotions are not necessarily followed by actions and therefore cannot cause them (*argument from imperfect association*). I argue that neither of these two arguments is compelling. At the same time, some of the concerns of causation skeptics can help us better understand what a theory of the causal connection between emotions and actions should explain.

Keywords
action tendency, control precedence, emotion, emotion indicators, emotional action, Frijda, James, mood-freezing experiments, probabilistic causation

Introduction

The integration of emotions with actions demands answering a basic question about their relation: Do emotions cause actions? If we were to simply rely on common sense, the answer would be fairly straightforward: of course emotions cause actions! What else could have caused me to strike my coworker, if not my anger? What else could have caused you to run away from the tiger, if not your fear?

Commonsensical as it may be, the view that emotions cause actions has been forcefully attacked from a variety of theoretical perspectives. William James famously argued that common sense is wrong about the direction of causation. We do not strike because we are angry, but rather are angry because we strike. This is because emotions, being perceptions of bodily changes, must come after such changes and consequently cannot cause them (*argument from timing*).

In more recent times, a second argument associated with psychological constructionism has acquired considerable prominence. The argument enlarges the scope of analysis to the full set of manifestations of emotions, including, besides actions, expressions, physiological changes, and experiential changes. The argument holds that, since these manifestations do not cohere tightly, an emotion cannot possibly be their common cause (*argument from imperfect association*).

In this article, I evaluate and ultimately reject these two influential arguments, concluding that neither the timing of the occurrence of, respectively, emotions and actions nor the imperfect association between them warrant the conclusion that emotions do not cause actions. While defending this view, I begin to sketch what sort of causal relation exists between emotions and actions, although most of the pesky details will have to be left for future work.

The Argument from Timing

Emotions as Feelings Violate Causal Precedence

James argued that emotions do not cause actions because they emerge *after* the bodily changes that constitute actions have occurred. On James's view, "*the bodily changes follow directly the* PERCEPTION *of the exciting fact*" in a reflex-like fashion, and "*our feeling of the same changes as they occur* IS *the emotion*" (James, 1884, p. 190). Since bodily changes must occur prior to being perceived, emotions as feelings cannot cause them, on pain of violating the common assumption that causes must precede their effects (*precedence*).

This well-known argument is not of merely historical significance. In modified forms, it continues to orient research in the affective sciences. Consider Baumeister, Vohs, DeWall, and

Corresponding author: Andrea Scarantino, Department of Philosophy and Neuroscience Institute, Georgia State University, P.O. Box 3994, Atlanta, GA 30302-3994, USA. *Email*: ascarantino@gsu.edu

Zhang's (2007) recent rejection of the commonsensical view that "[f]ear makes you flee" or that "anger makes you fight." Baumeister et al. understand emotions as "state[s] of conscious feeling, typically characterized by physiological changes such as arousal," where the feelings are "typically slow to arise and dissipate" and are "heavily saturated with cognitions" (pp. 168–169). Whereas for James emotions are the *raw feelings* that arise upon perceiving bodily changes (e.g., fear is the perception of one's body trembling and running), for Baumeister et al. (2007) emotions are the *labeled feelings* associated with categorizing oneself as having a certain emotion (e.g., fear is the feeling that comes about when we "recognize[. . .] the bodily state as fear" [p. 169]).

Suppose we encounter a tiger. On Baumeister et al.'s (2007) view, we automatically form a tendency to avoid it by virtue of an *affective reaction*, and we start acting on such tendency by fleeing well before we have a conscious fear experience, which would require recognizing one's arousal as fear, a process that takes "at least seconds, more likely minutes." Affective reactions are "twinges of feeling that something is good or bad, of liking or disliking for something," which "arise quite rapidly [and] may dissipate just as quickly" and do not require "elaborate cognitive processing" (p. 169). On this view, when we do feel fear, we are "already fleeing" or we have possibly already "reached safety," from which it follows that fear "would not alter the decision about how to act" and would at best "improve the person's ability to continue and succeed at that line of action" (p. 170).

So how can emotions affect actions, if they occur after them? Baumeister et al. (2007) argue that they do so *indirectly* by stimulating "cognitive processing." Their starting point is Gollwitzer's (1999) proposal that when agents deliberate they do not consider all behavioral options but rather apply previously created if-then rules to simplify the decision problem. Conscious feelings update the if-then rule database, with positive emotions (e.g., happiness, amusement) serving to confirm the rules and negative emotions (e.g., guilt, fear) serving to call the rules into question.

In addition, emotions can update the database of affective reactions, creating a new liking or disliking with its attendant approach or avoidance tendencies. These forms of *feedback* on if-then rules and affective reactions represent the two main ways in which labeled feelings indirectly cause actions. For example, recognizing a bodily state as fear produced by a tiger can lead to forming a new negative affective reaction for the meadow in which the tiger was found and, through cognitive elaboration, to a new "if-then" rule ("if invited to a safari, just say no") that will affect future deliberations.[1] I support the thesis that emotions can have an indirect impact on actions via the conscious experiences associated with emotions. Although the details of Baumeister et al.'s (2007) story can be fussed with, there is ample evidence that current feelings can have a causal impact on upcoming deliberations (see Loewenstein & Lerner, 2003). Shedding light on the specific mechanisms by which conscious feelings shape future actions won't be my concern in this article.

What I want to question is the *argument from timing*. In both James's (1884) and Baumeister et al.'s (2007) formulations, the argument states that, since emotions are nothing but the conscious feelings that come about after actions have already been selected and causes must precede their effects, emotions have no direct causal impact on the actions that proximately follow them (except perhaps intervening on an already established "line of action"). Rejecting this line of argument requires either denying that causes must occur before effects, or denying the identification of emotions with the conscious feelings that come about after actions have been selected. Some have suggested that it would be advisable for a theory of causation not to stipulate that backward-in-time causation is by definition impossible, but I will not question precedence in this article.[2]

My strategy will be to argue emotions are not the conscious feelings posited by either James (1884) or Baumeister et al. (2007). One way to do so would be to rehearse the various complaints that feeling theories in general have received from competing research programs, for example with respect to their ability to differentiate between emotions or with respect to their ability to account for the intentionality of emotions. Although I consider these lines of attack effective, I will not further articulate them here, nor explore how feeling theorists have tried to respond to them over the years. There is a more straightforward way to reject James and Baumeister et al.'s proposals, which focuses on their inability to capture the very nature of emotional actions (see also Scarantino, 2014).

Emotional Actions, Action Tendencies, Control Precedence

Most emotional actions are not reflex-like and they result from the regulation of action tendencies whose characteristics are emotion-specific. These simple facts cannot be accounted for if we identify emotions with either the *raw feelings* or the *labeled feelings* that come about *after* emotional actions have been performed. Let us consider how the problem affects the two versions of the *argument from timing* in turn.

For James, emotional actions such as striking follow in a reflex-like fashion from the perception of the exciting stimulus, and the emotion emerges only when the action and the attendant bodily changes are perceived, giving rise to the *raw feelings* with which emotions are identified. The job of bringing emotional actions about is left to *instincts*, where an instinct is "defined as the faculty of acting in such away as to produce certain ends, without foresight of the ends, and without previous education in the performance" (James, 1890, p. 383). For example, "love of man for woman," or "wrath at snakes," or "fear of precipices" are all "instances of the way in which peculiarly conformed pieces of the world's furniture will fatally call forth most particular mental and bodily reactions" (James, 1890, p. 191). Once these instinctual reactions are perceived, the emotion comes about. The problem with this proposal is that it focuses on a tiny sliver of emotional actions—those "fatally" brought about by the stimuli—while missing out on the great majority of emotional actions, which are not reflex-like.

In standard cases, when people are in the grip of love, anger, or fear they display *action tendencies*, which are states of readiness for a goal-oriented sequence of bodily movements. Whether the tendency is manifested hinges on the interplay between the tendency and *rational control*, a general-purpose capacity that adapts the tendency to the specific circumstances at hand, leading either to its flexible manifestation or to its inhibition. The management of action tendencies is one of the facets of the broader phenomenon of *emotion regulation*, which consists of the "processes by which we influence which emotions we have, when we have them, and how we experience and express [our] emotions" (Gross, 2008, p. 497).

The idea that emotions are associated with impulses to action in the form of action tendencies has been noted over and over again in the history of emotion theory, starting with the Aristotelian account of anger as an "impulse, accompanied by pain, to a conspicuous revenge for a conspicuous slight directed without justification towards what concerns oneself or towards what concerns one's friends" (1984, *Rhetoric*, 1378a31–1378b1). What is still fiercely disputed is the nature of the connection between emotions and action tendencies. For now, we can bracket the question of what grounds the emotion–tendency connection, and explore the nature of the action tendencies associated with emotions, turning to Baumeister at al.'s (2007) proposal.

Unlike James, Baumeister et al. (2007) get right the fact that emotional actions are not reflex-like but come about as a result of the regulation of action tendencies. On Baumeister et al.'s (2007) view, the job of bringing action tendencies about is left to *affective reactions*, which, unlike emotions, are alleged to have a *direct* causal impact on actions through the action tendencies they cause.[3] Since affective reactions are twinges of feeling that something is good or bad, they cannot bring about anything other than *generic* approach or avoidance tendencies.

This is precisely the issue, because emotional actions do not emerge from the regulation of generic approach and avoidance tendencies, but from the regulation of action tendencies whose characteristics are reflective of the type of emotion the agent is undergoing. But if the emotions are simply the conscious feelings that emerge after one has started acting, the specificity of the action tendencies from which emotional actions result is left unexplained. Furthermore, there are circumstances in which Baumeister et al.'s (2007) account simply makes the wrong predictions concerning the characteristics of the emotional action tendencies we observe.

Suppose that we find ourselves faced with a dangerous tiger or a vicious insult. On Baumeister et al.'s (2007) account, what happens next is that we form an immediate twinge of dislike, which brings about an avoidance tendency. This avoidance tendency will then be regulated, and eventually lead to action if the circumstances allow it. Seconds or minutes later, fear (in the case of the tiger) and anger (in the case of the insult) come about, once the bodily changes involved in coping with the tiger and the insult are recognized as fear and anger respectively. The first and most obvious problem is that the action tendency whose regulation leads to angry actions is *not* an avoidance tendency. It is widely acknowledged that angry actions come about from the regulation of *attack tendencies* (Frjida, 1986), which are a variety of *approach* rather than *avoidance* tendencies. So, contrary to Baumeister et al.'s (2007) prediction, not all stimuli appraised as "bad" lead to avoidance tendencies.

The broader problem is that Baumeister et al.'s (2007) fail to account for the complexities of the appraisal process that precedes the emergence of emotions. Starting with Arnold (1960), appraisal theorists have distinguished between various dimensions of stimulus assessment that take place within milliseconds and go well beyond determining whether the stimulus is "good" or "bad." Good and bad stimuli are further assessed in terms of the type of ego involvement at stake, the responsibility for the goodness and badness at hand, the ability to cope with the stimulus and the expectations concerning the future (Lazarus, 1991).

This fine-grained appraisal process can result not only in approach tendencies with respect to bad stimuli (the anger case), but more generally in tendencies that express not just whether the stimulus is liked or disliked but an emotion-specific appraisal of it. For example, when we appraise a dangerous tiger, we do not merely assess it as bad, but as bad in the specific way that dangerous things are bad. So the tendency that gets formed is not the generic tendency to avoid a bad stimulus but the fear-specific tendency to achieve safety from a dangerous stimulus. This form of avoidance differs from forms of avoidance associated with other "bad" stimuli. When ashamed to have been caught taking a bribe, for example, we form the tendency to hide our flawed self, which is a shame-specific form of avoidance different from the one involved in avoiding a tiger.

In addition, whereas the action tendencies caused by liking or disliking of a stimulus can have various ranges of prioritization, what is distinctive about the action tendencies whose regulation leads to emotional actions is their *control precedence*. The notion of control precedence was introduced by Frjida (1986) to emphasize that emotional action tendencies "clamor for attention and for execution." I have an action tendency to get spicy foods and avoid meats when ordering food at a restaurant, based on my affective reactions to the menu items (I like spicy foods, I dislike meat). But such tendency does not qualify as emotional unless it seizes control of me the way the action tendencies elicited by tigers or insults do. Emotional action tendencies take priority over other action tendencies by interrupting competing processes, by preempting access—in memory, inference, perception, etcetera—to information not related to the emotion's goal and by preparing the body for action.

What makes the action tendency specific to the emotion is in contention. As we will see in a later section, the most popular explanations characterize action tendencies as either effects or as constituents of emotions. Neither option is compatible with Baumeister et al.'s (2007) account, because if the emotion is the conscious feeling that comes about after the tendency has already been manifested, the tendency itself cannot be causally or constitutively affected by the emotion. If so, the fact that the tendency is emotion-specific and prioritized cannot be accounted for, which leaves a gaping explanatory hole at the heart of Baumeister et al.'s (2007) theory.

I conclude that the *argument from timing* must be rejected in either of the two versions we have considered. The identification of emotions with raw feelings (James, 1884) or labeled feelings (Baumeister et al., 2007) misrepresents emotional actions as being, respectively, reflex-like or emerging from the regulation of generic action tendencies, contrary to what we observe. Let us now turn to the second argument causation skeptics have offered to deny the existence of a causal connection between emotions and actions, namely the *argument from imperfect association*.

The Argument From Imperfect Association

Emotion Indicators Show Emotions Can't Be Causes

Causes do not just need to come *before* their effects (*precedence*). They must also be *distinct* from their effects (*distinctness*), and they must have the right type of *association* with their effects (*association*). A paradigmatic violation of *distinctness* occurs when one event is part of another. For instance, the event of writing "rr" (A) is part of the event of writing "Larry" (B; Kim, 1973). Since nothing can cause itself or part of itself, A cannot cause B.

Things get more complicated when it comes to *association*. The classical view that emerged from Hume's account of causation is that the association must be a "constant conjunction": in order for A to cause B it needs to be the case that B is constantly conjoined with A. As Hume (1748/1978, Section VII) put it, "we may define a cause to be an object, followed by another, where all the objects similar to the first are followed by objects similar to the second." On this view, the occurrence of A without B shows that A does not cause B, because if A caused B then B would follow A. This is the idea at the heart of so-called *regularity theories* of causation, which interpret events in a causal relationship as being necessarily connected to one another. It is this classical notion of causation that drives the *argument from imperfect association* recently made popular by psychological constructionists like Jim Russell (2003) and Lisa Barrett (2006).

I focus here on Russell's formulation of the critique, but very similar moves can be found in Barrett's work. The "traditional view" rejected by Russell (2003, p. 151) holds that in emotion episodes an "antecedent [stimulus] causes the emotion, which causes all of its various "manifestations," including not only "instrumental actions," but also "subjective feelings," "nonverbal signals," and "autonomic patterns." I call these the *indicators* of emotions, following common practice in measurement theory, to emphasize that they are the elements on the basis of which the presence of an emotion is inferred.

The traditional view has dominated emotion theory for a large part of its history, but Russell's primary target is *basic emotion theory* (BET), which proposes that anger, fear, disgust, happiness, sadness, surprise, and possibly other basic emotions are evolutionary adaptations that provide efficient solutions to fundamental life tasks. These solutions consist of *affect programs* that *cause* coordinated changes in the various indicators of emotions so as to orchestrate an adaptive response to a specific evolutionary problem. But if causes must necessarily be followed by their effects, then basic emotion theory leads to a prediction: whenever the affect program of, say, fear or anger is activated, a cascade of changes in indicators must occur, leading to a high degree of coherence among such indicators. As Russell (2009, p. 1262) puts it, "because the various components stem from a single [causal] entity, they [must] cohere in tight packages."

Basic emotion theorists seem to agree that this is what their theory predicts.[4] For example, Ekman tells us that when a basic emotion occurs, a

> [C]ascade of changes (without our choice or immediate awareness) occurs in split seconds in: the emotional signals in the face and voice; preset actions; learned actions; the autonomic nervous system activity that regulates our body; the regulatory patterns that continuously modify our behavior; the retrieval of relevant memories and expectations; and how we interpret what is happening within us and in the world. (Ekman & Cordaro, 2011, p. 366)

Psychological constructionists have argued on the basis of extensive meta-analyses of the empirical literature that this "cascade" of mandatory changes occurring "in split seconds" is nowhere to be found: the indicators of basic emotions do not cohere in tight packages and they manifest high individual variability. For example, some instances of fear involve changes only in a subset of the indicators, rather than changes in the full range of them (e.g., only experiential and physiological changes, but no changes in instrumental actions or expressions). Moreover, the specific changes in indicators that are activated by fear differ from time to time (e.g., different instances of fear are associated with different instrumental actions, different nonverbal signals, etc.). As Russell (2009, p. 1259) concludes, "in the realm of the emotions, reflex-like consistency is the exception rather than the rule."

So far we have focused on the empirical data. How can we use them to reject basic emotion theory? Enters the *argument from imperfect association*. The argument states that, since causes must necessarily be followed by their effects, and since the indicators of basic emotions lack the high coherence and specificity that would be expected if a basic emotion were their common cause, basic emotions cannot cause the changes in indicators. Thus, fear, anger, happiness, sadness, disgust, and surprise are not "internal entities [with] . . . the power to cause their own components" (Russell, 2015, p. 432). The *argument from imperfect association* plays an essential role in psychological constructionism. It motivates the "emergent view" that "emotions do not cause, but rather *are caused by*, their measured indicators" (Coan & Gonzalez, 2015, p. 213). The idea here is that the instantiation of an emotion is an effect of changes in indicators rather than the other way around, in the same way in which life stress is not the cause but the effect of changes in measured indicators such as job loss, divorce, recent bodily injury, and death in the family.

Different constructionists have proposed different accounts of what causes the changes in indicators, if not the emotion. The details of these alternative causal stories do not matter here. What matters is that, on an emergent model, fear is brought about by the changes in indicators. The reason why

psychological constructionists consider the "emergent view" superior to the "traditional view" is that it can "incorporate the contextual flexibility in emotion" (Coan & Gonzalez, 2015, p. 213) that the empirical data on low coherence and variability of indicators have revealed. Despite its popularity, the *argument from imperfect association* is flawed. First, it targets a version of basic emotion theory that is a priori unreasonable, although basic emotion theorists have largely themselves to blame for having turned such version into the standard bearer of their research program. Second, the argument presupposes an inadequate theory of causation.

The New Basic Emotion Theory

Basic emotion theorists have generally responded to the charge that their theory is incompatible with the empirical data by arguing that there is *more* coherence and specificity in emotion indicators than psychological constructionists have suggested (e.g., Izard, 2007). Although this may well be true, this counterobjection does not go far enough, because it does not vindicate the thesis that the indicators of anger, fear, disgust, happiness, sadness, and surprise have the sort of reflex-like consistency that would be expected if indicator changes were mandatory. So the rejoinder is not fully effective against the *argument from imperfect association*.

A better strategy for basic emotion theorists is to ask whether their theory is committed to the reflex-like consistency of indicators in the first place. The answer is a resounding no. This is because the fundamental life tasks that basic emotions allegedly evolved to solve are highly abstract and require different adaptive responses in different circumstances. Consider the task of reaching safety from dangers. Dangers differ in terms of how serious they are, in terms of how distant they are, and in terms of what is required to avoid them. Some dangers are relatively negligible and quite distant (e.g., entering an area in which a predator attack has occurred in the past), and they demand nothing more than orienting and getting ready for unspecified actions if the danger increases. Some dangers are significant and imminent (e.g., being hit on the head by a large falling object), and they demand specific and reflex-like actions (e.g., instantaneously jumping away from the object's trajectory).

Some dangers are also significant and fairly imminent (e.g., being attacked by a wolf running towards us from a short distance), and they can be dealt with successfully by a nonspecific range of actions that require some degree of planning and bodily control (e.g., finding a tree and climbing it, getting a long stick hidden behind a bush and keeping the wolf at bay with it, reaching for an arrow and shooting the wolf with it). In other words, in order to successfully tackle the problem it evolved to solve, the affect program of fear must operate in a reflex-like fashion in limiting cases (e.g., the large falling object case), but most of the time it needs to allow for a great deal of flexibility in responding. If it did not, its evolutionary selection would be utterly implausible in the presence of widely diverse embodiments of the same life task.

In light of these reflections, I have formulated and defended in previous publications a New BET, which differs from traditional BET in two main ways (Scarantino, 2014, 2015). First, the *New BET* replaces Ekman's prediction that the activation of an affect program causes a cascade of reflex-like changes in indicators with the prediction that the activation of an affect program causes the activation of an action tendency with control precedence. Second, the New BET rejects the terminology so far endorsed by basic emotion theorists, who have used ordinary folk terms like "fear" and "anger" to designate basic emotions. This is a mistake, because these are folk categories that have emerged to facilitate communication among English speakers and apply to widely diverse array of emotions, just like the folk categories of memory or intelligence apply to an assorted variety of types of memory and intelligence.

Two choices are possible to mark the discontinuity with ordinary language. The first is the creation of a neologism. The second is the addition of a modifier to the folk term. Both options have advantages and disadvantages, which I will not discuss here. For the purposes of this article, I will add the qualifier *basic* to fear, anger, disgust and so on, to clarify that I am focusing on such emotions *as prescriptively defined by basic emotion theorists*, without suggesting that all things we call "fear," "anger," "disgust," and so on in English fit the profile of a basic emotion.

The main advantage of the New BET is that, besides being more evolutionarily plausible than traditional BET, it is compatible with the empirical data unveiled by psychological constructionists. The New BET predicts that the indicators of the same basic emotion will correlate but weakly, because basic emotions by design allow for flexibility in responding. This flexibility will be reduced the more prototypical the elicitors become, and reach reflex-like consistency when failure to act rigidly but with maximal speed has major fitness costs (see Scarantino, 2015).

The *argument from imperfect association*, however, still stands. Critics may grant that basic emotions could bring about action tendencies with control precedence rather than a cascade of reflex-like changes in indicators, but add that this is precisely why we should not consider basic emotions to be causes, since causes must necessarily be followed by their effects. This rejoinder reveals that the *argument from imperfect association* does not just have basic emotion theory in its sights. It aims to undermine the "traditional view" of emotions writ large. The argument in effect holds that *any* theory that takes emotions to cause changes in indicators—for example, traditional BET, the New BET, innumerable other nonemergent theories of emotions—must be rejected because it is incompatible with the empirical data on lack of reflex-like consistency of indicators.

To properly reject the argument, then, we need to step back from basic emotion theory and ask a more general question about causation: what would it take for an emotion to cause changes in its indicators?

Probabilistic Causation: From Constant Conjunction to Probability Raising

I mentioned earlier that the causal relation must respect at least three principles: *Precedence, distinctness* and *association*. So far, we have worked under the assumption that *association* is equivalent to *constant conjunction*, as proposed by Humean

regularity theories of causation. In recent years, a great many theorists of causation have questioned this assumption, and embraced probabilistic approaches instead (see Hitchcock, 2012).

At the heart of probabilistic approaches is the notion of *association as probability raising*: in order for A to cause B, it needs to be the case that the occurrence of A increases the probability of B. As Edwards and Bagozzi put it in a helpful discussion of causal modeling in measurement theory, "the cause increases the likelihood of the effect but does not guarantee that the effect will occur" (2000, p. 157). This quote makes it clear that, as long as A raises B's probability, A can cause B even if A and B are not constantly conjoined, in the sense that B does not always occur when A does.

Two considerations in particular have undermined regularity theories of causation. The first is the realization that causes do not operate in isolation, but in the context of a broad set of other conditions. As a result, a cause that is followed by an effect in certain circumstances is not followed by the same effect in other circumstances. As Mackie (1974) pointed out, when we say that the striking of a match caused a fire, we do not mean that the striking was on its own sufficient for the fire, but rather than the striking of the match, in conjunction with the dryness of the materials, the presence of oxygen, the absence of a nearby sprinkler, and a great many other conditions, was sufficient for the fire. The striking of the match qualifies for Mackie as an INUS condition: it is an Insufficient but Nonredundant part of an Unnecessary but Sufficient condition. In other words, the striking alone is insufficient for the fire (since oxygen and other factors must also be present/absent), but it is a nonredundant part (the other factors without the striking would have been insufficient for the fire) of a condition that is unnecessary (a short-circuit might also have started the fire), but sufficient for the fire.

The second consideration that opened the door to probabilistic accounts is that the world may be *indeterministic*, namely such that the past and the laws of nature do not completely "fix" the present and the future. On this view, not only the striking of the match is not *individually* sufficient for the fire, but there may not even be a set of conditions inclusive of the striking that are *jointly* sufficient for the fire. What we can say in an indeterministic world is that the striking, in combination with a great many other conditions, is sufficient for *raising the probability of the fire* (rather than for the fire itself). The occurrence of the fire given the complete set of background conditions of the world may be an irreducibly chancy event.

Rejecting *association as constant conjunction* in favor of *association as probability raising* allows us to finally dispose of the *argument from imperfect association*. The argument's main flaw is that it wrongly presupposes that, if indicators' changes had emotions as their common cause, such indicators would show high coherence and specificity, because effects necessarily follow their causes. On the contrary, on the probabilistic theory of causation I endorse effects do not necessarily follow their causes but have their probability raised by them. If so, an emotion can be the common cause of changes in indicators even if such indicators lack reflex-like consistency, as long as the emotion increases the probabilities of such changes. For example, fear can cause running (and other indicator changes) by being causally or constitutively connected to an action tendency that probabilistically potentiates an open range of avoidant behavioral options that include running. Just like cancer is not necessitated by smoking but can still be an effect of smoking by virtue of the fact that smoking raises the probability of getting cancer, so running is not necessitated by fear but can still be an effect of fear by virtue of the fact that fear raises the probability of running.

I conclude that the empirical data on the variability of indicators fail to undermine the claim that emotions are causes of changes in indicators in the probabilistic sense of causation. Since this probabilistic sense has now replaced regularity theories of causation in medicine, physics, biology, neuroscience, economics, behavioral sciences, management sciences, and so on, it seems that causation skeptics need to go back to the drawing board, and provide evidence that emotions are not probabilistic causes of changes in indicators. Until they do, belief in the causal powers of emotions with respect to their indicators, including actions, can stand.

How Are Emotions and Action Tendencies Connected?

The picture I have outlined so far is that emotions are causes of actions, in combination with rational control/regulation, by virtue of their intimate connection with action tendencies, which can be understood as providing a "general direction for behavior by selectively potentiating [i.e., changing the probabilities of] coherent sets of behavioral options" (Gallistel, 1980, p. 322). As I have mentioned earlier in the article, there are different ways to explain this intimate connection. The following three proposals have been especially prominent:

(a) Emotions *cause* action tendencies.
(b) Emotions *are* action tendencies.
(c) Emotions *contain* action tendencies as proper parts.

Selecting among these proposals is not easy, because it requires taking position on which general theory of emotions is most suitable. This is not a task I can tackle in this article. My objective here is more modest, namely showing that all three options are compatible with the claim that emotions cause actions, give and take a few caveats.

The New BET endorses (a), holding that basic emotions, understood as affect programs elicited by appraisals, cause prioritized action tendencies. So the *New BET* provides a straightforward path for emotion–action causation: emotions cause action tendencies, which in turn cause actions in combination with rational control. Since causation is transitive—if A causes B and B causes C, A causes C—it follows that emotions cause actions. Other theories like Frjida's (1986) endorse (b) instead, suggesting that emotions *are* action tendencies caused by appraisals. This proposal also allows for a causal path from emotions to actions, in the sense that emotions cause actions by

being action tendencies, which in turn cause actions in combination with rational control.[5] Still other theories, for instance Scherer's (2009) component process model, endorse (c), proposing that action tendencies are just one of the components of emotions, alongside expressions, experiential changes, and autonomic changes. In this case, there is neither an affect program nor a "gestalt" appraisal causally responsible for the emergence of the action tendency (Moors, 2014). Rather, the action tendency is caused, along with the other elements of the emotion, by a set of "molecular" appraisals that have independent causal powers with respect to emotion components.

Scherer's model also allows for a causal path from emotions to actions, because emotions can cause actions, in combination with rational control, by virtue of the action tendencies they contain. In this case, emotions would cause actions by having one of their proper parts cause actions. Note that this is not a violation of *distinctness*. An event cannot cause itself or part of itself, but there is no obstacle to part of an event causing another event (e.g., a political campaign can cause public harm through the behavior of some of its members). However, if we consider actions themselves, rather than merely action tendencies, to be component parts of emotions, as some theorists have suggested (e.g., Moors 2014), the causal pathway from emotions to actions is effectively blocked.[6] This is because emotions cannot cause actions if actions are their proper parts, on pain of violating distinctness.

A Last Stand: Emotions as Indirect/Illusory Causes?

A causation skeptic may attempt a last stand, and suggest that, since emotions need to cooperate with rational control/regulation to bring about actions, emotions themselves are only *indirect* causes of action. An argument to this effect was presented by Baumeister et al. (2007). In an earlier section, I rejected their claim that emotions are *indirect causes* because, unlike affective reactions, they occur too late in the causal order to affect action tendencies. My objection was that this proposal fails to explain the distinctive characteristics of emotional action tendencies, in particular their being emotion-specific and prioritized.

As it turns out, Baumeister et al. also rely on a second notion of indirect causation. They acknowledge "the obvious fact that powerfully felt emotions can *directly* [emphasis added] cause people to cry, smile, scream" (2007, p. 171), but note that emotional actions do not come about with the same automaticity and therefore are only indirectly caused by emotions. This line of reasoning suggests that emotions are indirect causes because they are *individually insufficient* for actions. Now, Baumeister et al. (2007) are free to label causes that are individually insufficient for effects as *indirect*, but they should acknowledge that hardly any of the things we commonly call causes would qualify as direct on their account. For instance, the striking of a match would be an indirect cause of the fire, because the striking is individually insufficient for the fire. As we have seen, lots of other conditions such as the presence of oxygen or the absence of a sprinkler must be present for the fire to come about.

Furthermore, if indeterminism is true, *all* causes become indirect *sensu* Baumeister et al. (2007), including the cries, smiles, and screams that seemingly follow emotions automatically. This is because emotions, even in combination with other conditions, would not be sufficient for such effects, but at best sufficient for dramatically raising the probability of such effects, which would remain irreducibly chancy. Summing up, what we call "the cause" of an effect is commonly just a condition of special pragmatic interest in a larger set of conditions whose joint occurrence is followed by the effect. As our interests shift, so may our designation of "the cause." For example, we may refer to the presence of oxygen rather than the striking of the match as "the cause" of the fire when reporting on a fire that took place in a controlled environment, say a NASA testing room, that was supposed to be oxygen-free.

Similarly, we can refer to either emotion or rational control/regulation as "the cause" of emotional action, depending on our pragmatic interests, but we should not lose sight of the fact that neither emotions nor rational control/regulation are individually sufficient for the effect, and that they are both nonredundant conditions for the effect (with the possible exception of reflex-like emotional actions, which do not involve rational control). To emphasize this point, my recommendation is that we refer to "the cause" of emotional actions as the "interplay" between emotion and rational control/regulation.

In a section entitled "Mood Freezing: Exposing Illusory Causation by Emotion," Baumeister et al. (2007) raise the bar of their critique, suggesting that emotions may be *illusory* rather than *indirect* causes of actions. In mood-freezing experiments, subjects are given a placebo pill and told that it will preserve their emotions' hedonic status for some time—negative for emotions such as sadness and anger, positive for emotions such as happiness and awe—no matter what they do. This appears to produce changes in behavior. The main evidence for such changes concerns sadness. Whereas sad people tend to manifest helping behaviors, such behaviors are allegedly thwarted by mood-freezing pills. Baumeister et al. (2007) infer that "[i]t is not that sadness automatically or directly triggers a behavioral response of helping." Rather, "people resort to helping as a stratagem to make them feel better," from which they conclude that "emotion is the result, not the cause, of behavior" (p. 178).

This conclusion does not follow. I have various technical concerns about the mood-freezing paradigm, concerning primarily the problematic assumption that the mood-freezing procedure has no confounding effects on the mental state of subjects, but I will set them aside for now. Let us suppose for the sake of argument that mood-freezing experiments deliver what they promise, namely evidence that, once people are in a certain mood/emotion, what they do depends entirely on whether they believe their actions will make them feel better. This would show what we already knew, namely that rational control/regulation plays a key role in the causation of emotional actions. What the mood-freezing experiments do not show is that "emotion is the result, not the cause, of behavior."

On Baumeister et al.'s (2007) view, sadness is the conscious feeling that comes about when we recognize our bodily state as

one of sadness. This feeling is painful, and so we have reason to regulate it by choosing behaviors like helping which have the elimination or reduction of the painful feeling of sadness as an effect. This provides no support to the thesis that the painful conscious feeling of sadness is the effect of the behavior of helping. In fact, it could *not* be the effect, because the feeling of sadness comes *before* the behavior of helping. To assume that the latter causes the former is to violate causal *precedence*.

Baumeister et al. (2007) seem to be confusing the true claim that a reduction in the feeling of sadness is an effect of helping behavior with the false claim that the feeling of sadness we are trying to regulate is an effect of the behavior of helping. This would be like saying that, since eating reduces the negative feeling of hunger, hunger is an effect of eating. On the contrary, although the reduction of hunger is what eating aims for, eating itself is an effect of hunger. I conclude that mood-freezing experiments do not provide any evidence that there is something illusory about the causation of actions by emotions, and at best confirm the crucial role played by rational control/regulation in the causation of emotional actions.

Conclusion

In this article, I have considered and rejected two of the most popular arguments to the effect that emotions do not cause actions. The *argument from timing* holds that emotions cannot cause actions because they are the feelings that occur after actions. I have argued that thinking of emotions as feelings of this sort misrepresents the nature of emotional actions, portraying the latter as either reflex-like or as emerging from generic rather than specific and prioritized action tendencies. The *argument from imperfect association* holds that emotions cannot cause either actions or other changes in indicators because such changes do not necessarily follow emotions. I have argued that this conception of causation is untenable, and outlined an alternative probabilistic theory of causation in light of which emotions can cause indicator changes, including changes in actions, by virtue of raising their probabilities.

Acknowledgements

I would like to thank Moritz Mueller, Agnes Moors, Julien Deonna and an anonymous referee for their excellent comments on a previous version of this paper. The comments really made a difference! I also want to thank the Templeton foundation and Al Mele for funding in the Spring 2016 semester under the Philosophy and Science of Self-Control Project. The opinions expressed in this article are my own and do not necessarily reflect the views of the John Templeton Foundation.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1 Baumeister et al. (2007) add that emotions can also indirectly cause actions by means of *anticipated feedback*, in the sense that emoters can anticipate how they will feel in correspondence with various behavioral options, and act so as to either pursue or avoid such feelings prior to experiencing them. For example, a husband can predict that he will feel guilty upon cheating on his wife, and decide not to cheat in order not to feel guilty. This is an effect of *expected* rather than *experienced* emotions, and so it differs in kind from the effects of conscious feelings described in the text.

2 For example, Hitchcock (2008, p. 319) has argued that "there are solutions to the general field equations of general relativity" that allow the state of an object at a spatio-temporal region A to both cause and be caused by the state of the object at another spatio-temporal region B, in violation of precedence.

3 Baumeister et al. (2007) are not always consistent in their understanding of what it takes to be a direct cause of actions. In other parts of their article, they consider a direct cause to be one that is automatically followed by its effects, the way surprise is automatically followed by the widening of the eyes. I consider this second notion of direct causation later on in the article.

4 Various qualifications are added to this prediction. Ekman tells us for instance that basic emotions may blend (which would make the combined changes in indicators harder to forecast), that they form families (which may allow for some degree of variability among members of, say, the anger family or the fear family), and that the mandatory changes can be regulated (e.g., facial changes can be regulated within less than a second, changes in cardiac activity can be regulated after several seconds, etc.). These caveats are neither well motivated nor sufficient to account for the empirical data (Scarantino, 2015). As I argue later in the article, the prediction of mandatory activation of indicators must be straightforwardly rejected, rather than qualified, for evolutionary reasons.

5 Other theorists have proposed that emotions are feelings of action tendencies (e.g., Deonna & Teroni, 2016). Since such feelings come before actions, proposals of this sort are also compatible with giving a causal role to emotions.

6 I owe the clarification of this point to conversations with Agnes Moors.

References

Aristotle. (1984). *The complete works of Aristotle: The revised Oxford translation* (J. Barnes, Ed.). Princeton, NJ: Princeton University Press.

Arnold, M. B. (1960). *Emotion and personality Vol. 1*. New York, NY: Columbia University Press.

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*, 28–58.

Baumeister, R. F., Vohs, K. D., DeWall, C. N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection rather than direct causation. *Personality and Social Psychology Review*, *11*, 167–203.

Coan, J. A., & Gonzalez, M. Z. (2015). Emotions as emergent variables. In L. F. Barrett & J. A. Russell (Eds.), *The psychological construction of emotion* (pp. 209–226). New York, NY: The Guilford Press.

Deonna, J., & Teroni, F. (2016). How to get bodily feeling into emotional experience in the right way. *Emotion Review*. Advance online publication. doi:10.1177/1754073916639666

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155–174.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*, 364–370.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Gallistel, C. R. (1980). *The organization of action: A new synthesis*. Hillsdale, NJ: Erlbaum.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

Gross, J. (2008). Emotion regulation. In M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of emotions* (pp. 497–512). New York, NY: Guilford Press.

Hitchcock, C. (2008). Causation. In In S. Psillos & M. Curd (Eds.), *The Routledge companion to philosophy of science* (pp. 317–326). London, UK: Routledge.

Hitchcock, C. (2012). Probabilistic causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/win2012/entries/causality-probabilistic/

Hume, D. (1748/1978). *An enquiry concerning human understanding*. Oxford, UK: Oxford University Press.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, *2*, 260–280.

James, W. (1884). What is an emotion? *Mind*, *34*, 188–205.

James, W. (1890). *The principles of psychology*. New York, NY: Holt.

Kim, J. (1973). Causes and counterfactuals. *The Journal of Philosophy*, *70*, 570– 572.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Loewenstein, G., & Lerner, J. (2003). The role of affect in decision making. In R. J. Davidson, H. H. Goldsmith, & K. R. Scherer (Eds.), *The handbook of affective science* (pp. 619–642). Oxford, UK: Oxford University Press.

Mackie, J. L. (1974). *The cement of the universe*. Oxford, UK: Oxford University Press.

Moors, A. (2014). Flavors of appraisal theories of emotion. *Emotion Review*, *6*, 303–307.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145–172.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion*, *23*, 1259–1283.

Russell, J. A. (2015). The greater constructionist project. In L. Barrett & J. A. Russell (Eds.), *The psychological construction of emotion* (pp. 429–447). New York, NY: Guilford Press.

Scarantino, A. (2014). The motivational theory of emotions. In D. Jacobson & J. D'Arms (Eds.), *Moral psychology and human agency* (pp. 156–185). New York, NY: Oxford University Press.

Scarantino, A. (2015). Basic emotions, psychological construction and the problem of variability. In L. Barrett & J. Russell (Eds.), *The psychological construction of emotion* (pp. 334–376). New York, NY: Guilford Press.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, *23*(7), 1307–1351.