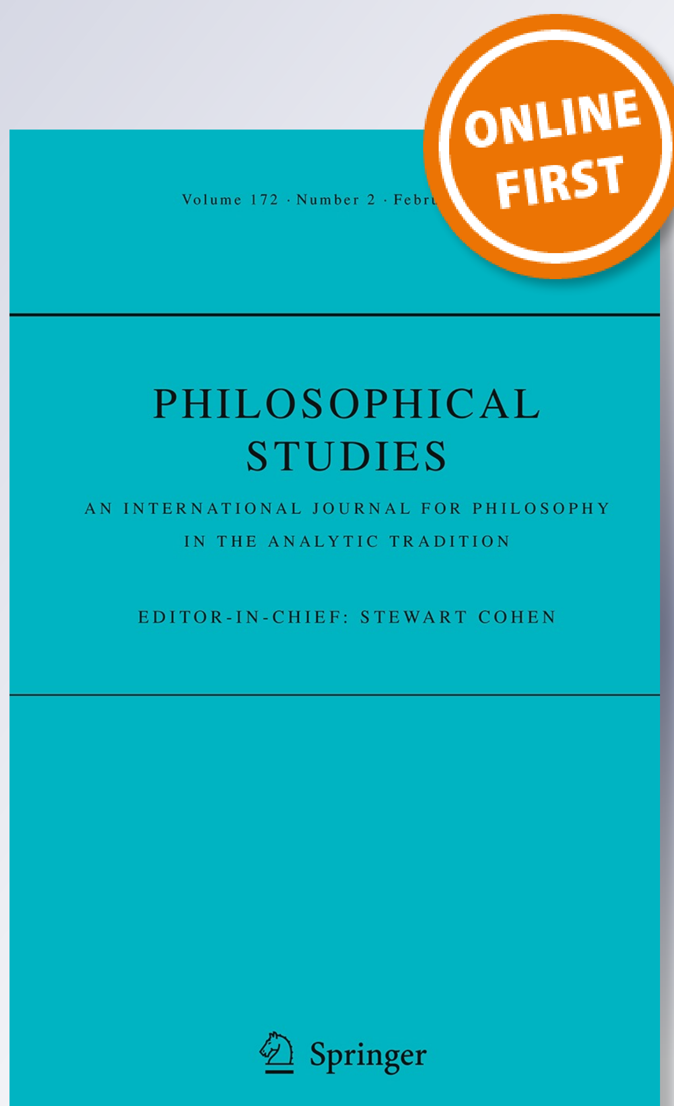
Voodoo dolls and angry lions: how emotions explain arational actions

Andrea Scarantino & Michael Nielsen

Philosophical Studies
An International Journal for Philosophy in the Analytic Tradition

ISSN 0031-8116

Philos Stud
DOI 10.1007/s11098-015-0452-y



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com”.

Springer

Author's personal copy

Philos Stud
DOI 10.1007/s11098-015-0452-y

Voodoo dolls and angry lions: how emotions explain arational actions

Andrea Scarantino · Michael Nielsen

© Springer Science+Business Media Dordrecht 2015

Abstract Hursthouse (J Philos 88(2):57–68, 1991) argues that arational actions—e.g. kicking a door out of anger—cannot be explained by belief–desire pairs. The Humean Response to Hursthouse (Smith in Human action, deliberation and causation. Springer, Netherlands, pp 17–41, 1998; Goldie in The emotions: a philosophical exploration. Oxford University Press, Oxford, 2000a; Goldie in Mind 109(433):25–38, 2000b) defends the Humean model from Hursthouse's challenge. We argue that the Humean Response fails because belief–desire pairs are neither necessary nor sufficient for causing emotional actions. The Emotionist Response is to embrace Hursthouse's conclusion that emotions provide an independent source of explanation for intentional actions. We consider Döring's (Philos Q 53(211):214–230, 2003) feeling-based Emotionist account and argue that it fails to explain arational actions. Finally, we develop our own Emotionist account, grounded in the Motivational Theory of Emotions one of us has developed. On our account, arational actions form a non-homogeneous class, some members of which must be understood as instrumental actions and some members of which must be understood as displacement behaviors of the kind animals display when their motivations are thwarted or in conflict.

Keywords Emotion · Humeanism · Arational actions

A. Scarantino (✉)
Department of Philosophy and Neuroscience Institute, Georgia State University, Atlanta, GA, USA
e-mail: ascarantino@gsu.edu

M. Nielsen
Department of Philosophy, Columbia University, New York, NY, USA

Published online: 07 February 2015

Author's personal copy

1 Introduction

One of the central challenges of contemporary philosophy of emotions is to make sense of emotional actions. The literature has focused in particular on what Hursthouse (1991) has called *arational actions*. These are emotional actions like kicking a table out of anger or gouging the eyes out in someone's picture out of hatred. Hursthouse argued that arational actions are intentional, not done for a reason, and best explained by being in the grip of an emotion rather than by a belief and desire pair. If that's right, at least some emotional actions are counterexamples to the Humean model of explanation, according to which belief and desire pairs cause and explain all intentional actions.

Hursthouse's seminal paper has generated two main lines of response in the literature. The *Humean Response,* advocated for instance by Smith (1998) and Goldie (2000a, b), has been to defend the view that belief and desire pairs are an essential part of the explanation of emotional actions of all kinds. We argue that the Humean Response fails, because the Humean toolbox is not equipped to capture what makes actions emotional in the first place. What we call the *Emotionist Response,* offered for instance by Döring (2003, 2007), has been to embrace the conclusion that emotions can provide an independent source of explanation for intentional actions. We argue that Döring's account also fails to explain arational actions, but is on the right track in giving a central explanatory role to emotions.

Finally, we present our own Emotionist account, which is grounded in the Motivational Theory of Emotions (MTE) that one of us has developed (Scarantino 2014), but is compatible with a broader range of emotion theories. MTE aims to define emotions in terms of how they motivate us to act (Frijda 1986, 2007; Kovach and De Lancey 2005), rather than in terms of how they feel (e.g. Damasio 1994, 1999; Prinz 2004) or in terms of the evaluations they entail (Nussbaum 2001, Solomon 2003). The core insight of MTE is that emotions are prioritized action tendencies endowed with intentional content.

We will show how this theory can ground a general taxonomy of emotional actions that explains arational actions in terms of the interplay between prioritized action tendencies and rational control. On the view we propose, arational actions do not form a unified category: some of them are straightforwardly *instrumental* emotional actions that pursue the goals of the emotions that cause them, whereas others are *displaced* emotional actions, namely actions whose goals are redirected, for reasons we will investigate, away from the original goals of the emotions that cause them.

2 Hursthouse on arational actions

According to the dominant Humean model, a belief and desire pair always explains intentional actions. This pair is often referred to as a *motivating reason* (Smith 1994) and it is assumed to cause the intentional action that it explains. Consider John's moving his finger against a light switch. Humeans would explain it in terms

Author's personal copy

of John's desire to illuminate the room and his belief that moving his finger against the switch is a means for illuminating the room (Smith, 1998). To generalize,

Basic Humean Model: For any intentional action A, agent A-ed because she desired to X and believed that A-ing would be a means to X.

Beliefs and desires are distinguished by Humeans in terms of *direction of fit* (Searle 1983). The direction of fit of a mental state specifies whether the mental state provides a representation that must "fit" the world (mind-to-world direction of fit) or a representation that the world must be changed to "fit" (world-to-mind direction of fit). Beliefs are the paradigmatic mind-to-world attitudes because they aim at being true, whereas desires are the paradigmatic world-to-mind attitudes because they aim at being realized.

Hursthouse's contention is that there is a class of emotional actions that the Basic Humean Model can't explain. Here are some of her examples:

[1] Rumpling someone's hair (out of love)
[2] Gouging holes in someone's picture (out of hatred)
[3] Kicking a door (out of anger)
[4] Jumping up and down (out of joy)
[5] Rolling around in one's dead wife's clothes (out of grief)
[6] Covering one's face in the dark (out of shame)

Hursthouse argues that most—but "not necessarily all"—tokens of action types [1]–[6] have three things in common. First, they are intentional actions. Second, they are not done for a (motivating) reason, understood as a belief and desire pair. According to Hursthouse, actions [1]–[6] generally involve a desire but not a belief to the effect that the action performed is a means to a valued end.[1] Third, the actions would not have been performed had the agent not been in the grip of an emotion. Hursthouse concludes that actions with these three characteristics should be called *arational* in order to emphasize that they are *not done for a reason* on account of the missing belief.

Hursthouse's explanation of why someone rumpled her lover's hair [1] is that the person was in love and just desired/wanted/felt she had to rumple the loved one's hair. Her explanation of why Jane gouged holes in Joan's picture [2] is that Jane was in the grip of hatred and just desired/wanted/felt she had to gouge holes in Joan's picture. And so on.

Although an emotion and a desire are both mentioned in these examples, Hursthouse's view is that "the mere fact that [one] was in [the] grip [of emotion] explains the action as much as anything else does" (p. 59). This suggests that emotions alone are *sufficient* for explaining arational actions. To generalize,

Hursthouse's Emotionist Model: For any arational emotional action A in [1]–[6], agent A-ed because she was in the grip of some emotion E.

[1] Hursthouse allows for cases in which the relevant beliefs are instantiated. For example, I may rumple my child's hair with the belief that this will be a way to express my love to her. This is why on her view some tokens of types [1]–[6] are amenable to a standard Humean explanation.

Author's personal copy

Hursthouse defends this model by arguing that no suitable belief and desire pairs can ground a Humean explanation of actions [1]–[6]. She focuses on two candidate belief–desire pairs in particular. The first is that the actions in question are caused by the desire to express an emotion E and by the belief that by A-ing one expresses E. The second is that they are caused by the desire for pleasure and by the belief that A-ing would bring pleasure. We think Hursthouse is right to reject these two candidate belief–desire pairs because we often engage in actions [1]–[6] without any independent desire to express an emotion or any belief that we will get pleasure out of it.

Despite her misgivings about Humeanism with respect to *arational* emotional actions, Hursthouse accepts the Basic Humean Model for *garden-variety* emotional actions. For example, she thinks that running from a bull out of fear can be explained by a desire for self-preservation and a belief that running is a means of self-preservation, or that slapping someone out of anger can be explained by a desire to punish someone and the belief that slapping that person is a means of punishing him.

Hursthouse deserves credit for having raised awareness of the limitations of the Humean model, but we find her overall critique only partially convincing. First, Hursthouse only considers two candidate belief–desire pairs, and not even the most plausible ones at that, as Humeans were quick to point out. What anti-Humeans need is a principled argument against the necessity and/or sufficiency of belief and desire pairs for the explanation of actions [1]–[6]. Second, Hursthouse concedes too much to Humeanism. Our view is that the Humean Model is unsuitable for explaining not only arational actions [1]–[6], but also garden-variety emotional actions. Third, Hursthouse does not clarify what emotions are, treating them as unexplained explainers. We think that emotional actions can best be understood on the basis of a viable general theory of emotions that does not treat emotions as primitives.

3 Smith strikes back: the expanded Humean model

Smith (1998) argues that Hursthouse does not consider an obvious belief and desire pair in her analysis of emotional actions [1]–[6]. Take the action of rolling around in one's dead wife's clothes out of grief. According to Smith, this action can be explained by the desire to roll around in one's dead wife's clothes and by the belief that one can do so just by doing what one is doing. The latter belief is indeed unusual, because it identifies means and ends, but it is a belief nevertheless. The same general recipe applies to all other cases of allegedly arational actions, which turn out to be actions done for a reason: the agent desires to perform a certain action—gouging holes in someone's picture, kicking a door, etc.—and believes that she can do so just by doing what she is doing.

This being said, Smith acknowledges that his account is "distinctly unsatisfying" when it comes to understanding actions [1]–[6] because "it provides us with very little illumination" (p. 160). What is left unexplained, Smith thinks, is the *origin* of

Author's personal copy

the relevant desires. Why would anyone desire to roll around in one's dead wife's clothes or gouge holes in a picture?

Smith attempts to provide more "illumination" by offering the following model:

> *Smith's Expanded Humean Model*: For any emotional action A in [1]–[6], agent A-ed because she desired to A and believed that she could A just by doing what she was doing, and she was in the grip of emotion E.

Smith stays neutral on what emotions are, but argues that being in the grip of an emotion "by definition" involves being disposed "to think, and to desire, and to do, all sorts of things". For example, being in the grip of grief makes one disposed to "cry, dwell on memories of the loved one, seek out things that remind [him] of the loved one and hold them close, and so on and so forth" (p. 160).

One may interpret Smith's Expanded Humean Model as an implicit recognition of the fact that emotional actions [1]–[6] cannot be adequately explained without mentioning emotions. But Smith does not equate failure to mention the emotions underlying such actions with failure to adequately explain them, despite his claim that very little illumination is provided unless the emotions are mentioned. The Humean view is that what makes [1]–[6] *actions* in the first place is that a belief and desire pair causes them. And adequately explaining an action simply amounts to spelling out the belief and desire pair that caused it.

Humeans, in effect, posit two types of action explanations:

> *Constitutive Humean explanations* are given in terms of belief and desire pairs and they provide understanding of why a certain sequence of bodily movements is an intentional action rather than a mere happening.

> *Non-constitutive non-Humean explanations* are not given in terms of belief and desire pairs, but whatever additional understanding of the action they provide presupposes a constitutive Humean explanation in terms of beliefs and desires.

On this view, there is no competition between Humean and non-Humean explanations, because the latter always presuppose the former. Furthermore, the proposed schema offers a clear recipe for increasing the understanding provided by a Humean explanation, namely adding details on the origin of the beliefs and desires that caused action. Smith acknowledges that pragmatic reasons will generally favor adding such details in the case of arational actions

For instance, we will provide a better explanation of the action of rolling around in one's dead wife's clothes if we also mention that the husband is grieving. But Smith makes it clear that belief and desire pairs still occupy a central place in the order of explanation, because grief only sheds light on the causal history of the belief and desire pair that is causally responsible for rolling around in one's dead wife's clothes. Grief alone, on the other hand, cannot tell us why rolling around is an action rather than a mere happening.

Kovach and De Lancey (2005) have criticized Smith's account by arguing that the Humean explanation is superfluous rather than explanatorily central. If grief is by definition a state that disposes us "to think, and to desire, and to do, all sorts of

Author's personal copy

things", we may explain why a person rolls around in his dead wife's clothes by simply mentioning his grief. Now, although we agree that actions done out of grief can be explained in terms of grief exclusively, we find the Humean explanation worse than superfluous: it is positively misleading.

What makes it misleading is its commitment to the view that emotional actions are caused by belief and desire pairs. We will argue (Sect. 7) that if emotional actions were caused by the sorts of belief and desire pairs that Humeans posit, they would not manifest the characteristics that make them emotional. So not only do Humean explanations provide very little illumination of emotional actions, as Smith himself acknowledges, but they also cannot tell us what makes actions done out of emotions actions in the first place. We will consequently offer a non-Humean and yet constitutive explanation of emotional actions in terms of the emotions that cause them. Such explanation is in direct competition with Humean explanations, because it explicitly denies that emotions explain actions by explaining belief and desire pairs, proposing a more direct causal path from the emotion to the emotional action.

4 Goldie's hybrid theory and the role of wishes

Goldie (2000a, b) attempts to deal with the problem of arational actions by introducing further amendments to the Basic Humean Model. He first distinguishes between two types of actions done out of emotions: *reasoned actions* and *genuine expressions*. Reasoned actions are emotional actions such as jumping over a gate in fear of a bull (Goldie 2000b, p. 26) or doing "some sort of harm to the person who is the object of your anger" (p. 28). Like Hursthouse, Goldie thinks that the Basic Humean Model adequately explains these sorts of garden-variety emotional actions.

Genuine expressions of emotions, on the other hand, are not done as a means to some further end. Genuine expressions correspond to Hursthouse's arational actions, but Goldie aims to reject Hursthouse's "revisionary view that we have here a class of intentional action that cannot be explained by a belief and a desire" (p. 27). His view is that, with some additional amendments, the Expanded Humean Model can explain arational actions while preserving a central role for beliefs and desires.

Goldie's worry with respect to Smith's model is that it posits desires that Goldie does not find "primitively intelligible" in light of the emotions being ascribed—e.g. the desire to roll around in someone's clothes or the desire to gouge out the eyes in a photo. In order for desires to be primitively intelligible, it should be the case that they "cannot be explained in virtue of anything else other than the emotion which they are a part of" (p. 28).

Goldie's positive proposal is to add *wishes* to the Humean toolbox. Wishes are special types of desires, in the sense that when "I wish for something, I desire that thing, and I also imagine, or am disposed to imagine, the desire to be satisfied" (Goldie 2000b, p. 28). Consider Jane's scratching out the eyes in a picture of Joan. Goldie thinks, contra Smith, that the desire to scratch out the eyes in Joan's picture is *not* primitively intelligible in light of Jane's hatred. What *is* primitively

Author's personal copy

intelligible is rather Jane's desire to scratch out *Joan's eyes*. But Jane does not try to satisfy this desire because she recognizes certain "civilizing constraints". As a result, Jane only *wishes* to scratch out Joan's eyes: she desires to scratch out Joan's eyes and she imagines satisfying this desire. The desire component of this wish is primitively intelligible in light of Jane's hatred.

We note two major problems with Goldie's account. First, it does not shed light on *why* the genuine expressions of emotion occur. As Kovach and De Lancey (2005) have pointed out, "not all imaginings are cases of active pretending" (p. 119). Jane could easily imagine scratching out Joan's eyes, and do nothing about it. It seems that in order to lead to action, Joan's wish would need to cause an actual desire to scratch out the eyes in the picture, and then combine with a belief that one can scratch out the eyes in the picture just by doing what one is doing. This makes Goldie's wish-based account explanatorily deficient, and in demand of further integration with the sorts of desires posited by Smith.

Second, Goldie's account does not succeed at its intended purpose, namely rejecting Hursthouse's attack of Humeanism. As Döring (2003) also notes, Goldie admits that genuine expressions of emotion that lack a symbolic connection between what one desires and what one does—e.g. kicking a table out of anger—are intentional, and yet a belief and desire pair cannot explain them. In this case and others, Goldie acknowledges that "the explanation is more likely to be along Hursthouse's lines, as due simply to being in the grip of the emotion" (p. 31).

Finally, Goldie differs from Smith in holding that "although genuine expressions of emotion which are actions can…be explained by a belief and a desire, they cannot *adequately* be explained in this way" (p. 26). For Goldie, what is required to achieve an adequate explanation is the mention of an emotion and of the wish that is primitively intelligible in light of it. But if belief and desire pairs are *insufficient* for adequately explaining even the subset of expressive actions to which Goldie's model applies, Hursthouse's revisionary view that there is a class of intentional actions that cannot be explained by a belief and a desire is largely confirmed rather than rejected.

The failures of Smith and Goldie's accounts to provide explanations that are both adequate and properly Humean suggest the opportunity to explore non-Humean and emotion-based explanations of emotional actions. We turn to this task now.

5 Doring's feeling-based account of emotional motivation

Emotionist models hold that adequate explanations of emotional actions can be formulated in terms of the emotions causing them, without invoking belief and desire pairs. Döring (2003) has developed the most detailed Emotionist model to date. Her objective is to replace Humean belief–desire models of emotional action explanation with an emotion-based model built around the idea that emotions exhibit "a distinctive motivational force" (p. 220).

Although Döring does not provide a full-fledged theory of emotions, her working account is that "an emotion is an occurrent conscious state, with a certain affect, and

Author's personal copy

with a certain kind of intentional content" (p. 220).[2] A state with intentional content, on Döring's view, is directed at a "target" construed in evaluative terms. For example, hatred towards one's rivals has one's rivals as a target, and it evaluates them as awful. Grief about one's mother's death has one's mother's death as a target, and it evaluates her death as a loss. According to Döring, what turns a certain evaluation into an emotion is "its feeling dimension, which is also called its 'affect'" (p. 223). Affect plays a key role in Döring's explanations of emotional actions, because "[i]t is the emotion's affect which gives it motivational force, rather than any desire being 'part' of it" (p. 224).

Döring thinks of desires as representations with a world-to-mind direction of fit: their content describes how the world is to be changed. She rejects the view that desires are parts of emotions because she thinks that emotions do not always contain representations of how the world must be changed. This leads her to distinguish between two ways in which emotions motivate actions. In some cases, the emotion's representational content "fails to provide an end for action" (p. 224). For example, grief about one's mother's death represents the death as a loss, but "it is…impossible to change the world in such a way that it fits the emotion" (p. 220).

Jane's hatred for Joan may also not provide an end for action because "there may be nothing Jane can do to change the world in such a way that Joan no longer appears awful to her" (224).[3] Döring's proposal is that, when the representational content of an emotion does not provide an end for action, affect causes arational emotional actions. For example, the affect associated with grief may lead one to roll around in one's dead wife's clothes, and the affect associated with Jane's hatred may lead her to scratch out the eyes in a picture of Joan.

In other cases, an emotion's content *does* provide an end for action. On Döring's account, affect motivates action in these cases as well. Fear of the aggressive looking Jane represents Jane as dangerous, and this content "provides…the end of avoiding danger" (p. 224). The action that results—e.g. running away—is a means to that end. To generalize,

> *Döring's Emotionist Model*: For any arational emotional action A in [1]–[6], the agent A-ed because she was in the grip of some emotion E whose representational content did not provide an end for action but whose affect motivated her to A. When representational content does provide an end to action, a garden-variety emotional action ensues.

We consider this model inadequate. The central problem with it is that it remains mysterious why the feeling dimension of emotions should be intrinsically motivating and thus capable of explaining emotional actions. A great many feelings do not motivate. For example, one can experience the feeling of a ceiling fan blowing air through one's hair, but this feeling need not motivate one to act at all. Döring may reply that this is not an *emotional* feeling and therefore not a candidate

[2] This understanding of emotions is influenced by Perceptualist theories of emotion, according to which emotions are a species of "affective perceptions" (Damasio 1994; Prinz 2004; Scarantino 2014).

[3] We take this to be a misinterpretation of the end of hatred, which is to hurt the object of one's hatred, rather than to stop seeing the object of one's hatred as awful. We will disregard this point in what follows.

Author's personal copy

for providing motivation on her view. Fair enough, but we still want to know what distinguishes emotional from non-emotional feelings, such that the former necessarily motivate and the latter do not. And here, Döring's account is silent.

Döring seems to have replaced one form of primitivism—Hursthouse's claim that arational actions occur because one is in the grip of an emotion—with another form of primitivism—emotional actions occur because one is in the grip of a feeling (or affect). To say that Jane gouged holes in Joan's picture out of a feeling of hatred, or that someone rolled around in his dead wife's clothes out of a feeling of grief, simply substitutes one unexplained explainer (the emotion) for another (the feeling associated with the emotion).[4]

6 What is a theory of emotional action supposed to explain?

So far we have discussed emotional actions and given various examples of them, but we have yet to clarify what makes an action emotional in the first place. It is time to get clear on this matter, because adequately explaining emotional actions demands explaining both why they are actions (rather than happenings) and why they are emotional (rather than non-emotional).

We consider two markers of emotionality to be especially important (see Scarantino 2014 for further discussion).

6.1 Impulsivity

Emotional actions manifest impulsivity in two related senses. First, they typically involve a "sense of urge", comprising "both an expectation of gain after completing [the action], and haste to fulfill it" (Frijda 2010, p. 571). There is a sense of urge in running out of fear of a bull, as there is a sense of urge in gouging out the eyes in a picture out of hatred. In some cases, the urge to act does not abate even after the goal of the action has been achieved. As DeLancey (2002) notes, some emotional actions are *postfunctional*: they continue even after their ostensive goals have been achieved, as if endowed with inertial force. DeLancey gives the example of someone who, enraged at a vicious dog responsible for attacking his daughter, continues shooting the dog long after it is clearly dead, or of someone who is terrified of a rabid dog and continues to behave in a panicked way even after safety has clearly been achieved.

Emotional actions also typically involve the "use of only part of the available cues that might indicate the adequacy of action" (Frijda 2010, p. 571). Call this

[4] Another problematic consequence of Döring's view is that it seems to preclude the possibility of being motivated by unconscious emotions. This is because, on Döring's view, affects and feelings are necessarily conscious. This shortcoming, however, may not be fatal. On the one hand, the debate on the very existence of unconscious emotions is unresolved (Hatzimoysis 2007; Berridge and Winkielman 2003, Winkielman and Berridge 2004). On the other hand, affect-based accounts can in principle deal with unconscious emotions if they allow affects themselves to be unconscious (Prinz 2004; Lacewing 2007).

Author's personal copy

partial informational access. We distinguish three sources of partial informational access. First, some emotions appear to be *cognitively impenetrable* (Fodor 1983): the information available to the belief system fails to penetrate the relevant emotion system. This leads to the well-known phenomenon of *recalcitrance to reason* (D'Arms and Jacobson 2003), manifested when agents act out of emotions even though the emotions appear to contradict what they believe. For example, a snake phobic subject may believe that a picture of a snake is not dangerous, and yet be motivated to rapidly move away from the picture out of fear.

Second, in emotional actions the time spent collecting information about what to do is limited by one's preference for acting sooner rather than later. As a result, emoters often fail to properly consider the consequences of what they do. Slapping someone in a bar fight out of anger won't generally be preceded by careful consideration of the consequences—legal, moral, and practical—of engaging in an act of violence. As a result, we may pick a fight with a much stronger opponent.

Third, in emotional actions the way information is processed manifests various cognitive biases. The emoter's ability to focus attention, recruit memories, draw inferences, evaluate evidence, have accurate perceptions, assign probabilities, assess risks, and so on are all potentially impacted by being in the grip of an emotion (Loewenstein and Lerner 2003; Elster 2010). Limited information gathering and information processing biases are largely responsible for what some have referred to as the "shortsightedness" of emotional actions (Pacherie 2002, p. 76).

6.2 Bodily underpinnings

Emotional actions also tend to be characterized by concurrent bodily underpinnings. These include changes in facial, vocal and postural expressions, and changes in the autonomic nervous system. Whether we strike someone out of anger or kick a door out of anger, our angry actions tend to be accompanied by involuntary facial expressions such as a fixed stare, widened eyes, lowered eyebrows, and bared teeth, and by autonomic physiological changes such as an increased heartbeat and tremors (Ekman 1999). These bodily changes are not random, but rather geared towards facilitating a goal-oriented set of emotional actions like yelling, slapping, pushing, punching, spitting and so on. The bodily changes involved are often accompanied by distinctive bodily feelings, namely experiential states generated by perceiving one's body in turmoil (Goldie 2000a, b).

7 The real trouble with Humeanism

Now that we have clarified that the emotionality of emotional actions primarily consists of impulsivity and bodily underpinnings, we can offer a principled argument against Humeanism. As we have seen, Humeans believe that a constitutive explanation of emotional actions is adequate insofar as it sheds light on what makes a sequence of bodily changes an action rather than a happening. Accounting for the emotionality of actions is assumed to be a job better left to non-constitutive non-Humean explanations, which may cite emotions (as in Smith's

Author's personal copy

model), but always presuppose a constitutive Humean explanation in terms of beliefs and desires.

This division of labor suggests that Humean and non-Humean explanations can peacefully co-exist: the features that make an action emotional are supposed to be compatible with the features that make it an action in the Humean sense, namely being caused by a belief and desire pair. The problem is that whatever belief and desire pair may be said to motivate an emotional action, there is a non-emotional "twin" version of the action caused by the very same belief and desire pair.

If so, saying that an emotional action is explained by a belief and desire pair which is in turn explained by being in the grip of an emotion tells us nothing whatsoever about what makes the resulting action emotional: the very same belief and desire pair that allegedly caused the emotional action can cause a perfectly non-emotional version of the same action.

Consider Smith's proposal that John's rolling around in his dead wife's clothes can be explained by the desire to roll around in the clothes and by the belief that he can do so by doing just what he is doing. This belief and desire pair can motivate one to roll around in one's dead wife's clothes in the absence of any grief.

Suppose that John has a fetish for his dead wife's clothing. As a result, John regularly and spontaneously rolls around in his dead wife's clothes without experiencing any grief nor, by assumption, any other emotion. What explains John's unemotional action according to the Humean account? Evidently, the same belief–desire pair that explains his grieving action: John's desire to roll around in his dead wife's clothes and his belief that he can satisfy his desire by doing just what he is doing.

This shows that the belief and desire pairs that allegedly cause emotional actions are insufficient for causing actions that are truly emotional: something extra accounting for emotionality would need to be added on, and it cannot simply pertain to the causal origin of the belief and desire pair. Whatever causal origin John's belief and desire may have, that belief and desire pair alone cannot account for the emotional aspect of the action we need to explain.

But things get worse for the Humean. Besides shedding no light on the emotionality of an action, the Humean explanation leads us to actively expect the action not to manifest markers of emotionality. Consider a *postfunctional emotional action* like continuing to shoot a dog long after the dog has died (DeLancey 2002). If this action were caused by the desire to kill the dog and by the belief that one can do so by shooting the dog, as Humeans would have it, we would expect the action to *stop* once the shooter realizes that the dog is dead. The fact that the action continues, thereby revealing its distinctive emotional impulsivity, suggests that what caused it is something other than the pair of beliefs and desires Humeans would posit.

This problem does not only affect "extreme" emotional actions like postfunctional ones, but garden-variety emotional actions as well. Suppose someone believes that the roof is in danger of caving in under heavy snow and desires to bring his family to safety. We would expect him to assess the likelihood of an imminent roof collapse, and depending on this assessed likelihood gather one's family, collect all vital items that need to be brought out of the house, locate the best exit, and finally get out of the house. We would expect him to do all these things in a hurry

Author's personal copy

commensurate to the perceived likelihood of an imminent collapse and without distinctive bodily changes.

What we would *not* expect is what we observe when people are motivated by fear. If one is in the grip of fear that the roof will collapse and obliterate one's entire family, one may start sweating and trembling, forget to pick up the coats necessary for protecting one's children in the outside cold, mistake a common crack in the roof for a clear sign that the roof is about to cave in, trip over objects, scream unclear and inconsistent directives, and more generally display the shortsightedness typical of emotional actions.

One last stand may be available to Humeans, namely giving up on the idea that the sort of belief and desire pairs that cause emotional actions are *ordinary* beliefs and desires. For example, Humeans may propose that emotional actions are caused by belief and desire pairs in which the desire is *impulsive* and *bodily* as a result of its *strength* or *intensity* (e.g. Marks 1982; Solomon 2003). Such strength could in turn be explained as resulting from the strength of the desires that constitute the emotions causing the relevant belief and desire pairs. In addition, beliefs may be described as a special kind of *affective beliefs* that are partially informationally encapsulated, and consequently able to co-exist with contradictory *cognitive beliefs* as in the case of emotional recalcitrance (e.g. Nussbaum 2001).

We find these moves to be patently ad hoc since the various markers of emotional actions are embedded by fiat into the belief and desire pairs that supposedly cause them, without being exportable to any other species of the belief genus or the desire genus. The trouble with the claim that emotional actions are caused by *sui generis* beliefs and desires, elastically stretched to fit whatever properties emotions may be found to have, is that the claim becomes at best trivial and at worst unfalsifiable (Scarantino 2010).

Moreover, it is far from clear that the specific proposals offered to account for what makes emotional beliefs and desires distinct from ordinary ones are persuasive. For example, we reject the view that the desires motivating an emotional action are always stronger than the desires motivating a non-emotional twin version of the same action, unless we make this definitionally true of emotional desires and consequently not especially enlightening.[5]

To go back to our earlier example, we should not assume that John's desire to roll around in his dead wife's clothes out of grief is necessarily stronger than his desire to roll around in his dead wife's clothes out of a fetish. In either case, the desire at work is stronger than any competing desire in the motivational sense because each John ends up rolling around in his dead wife's clothing rather than doing something else. And we do not know of any viable way to compare John's fetishistic and grief-driven desires so as to demonstrate that the former is always weaker than the latter.

Just to mention some possible suggestions, it does not seem true that the fetishistic desire is weaker than the grieving desire in terms of its "felt violence" or in terms of the vividness of its associated "imagery" or in terms of its ability to

[5] See Gendler (2008) and Benbaji (2012), among others, for reasons to resist the view that recalcitrant emotions involve beliefs.

Author's personal copy

"grab attention" or in terms of its ability to "produce pleasure" (Mele 1998; Sinhababu 2009). It is also far from clear how the two desires would fare in a counterfactual scenario in which they competed against each other (e.g. a scenario in which John has both desires, but the fetishistic desire has one piece of clothing as its object, the grieving desire has a different clothing item as its object, and only one piece of clothing can be retrieved). So we conclude that the idea that the desires involved in emotional actions are necessarily stronger than the desires involved in non-emotional twin actions is ad hoc and far from convincing.

On our view, Humeans should simply accept that their favorite explanation of action does not apply universally to all intentional actions. Emotions are not causally responsible for emotional actions by courtesy of mediating beliefs and desires but much more directly, in ways we need to understand. In the next section, we develop an alternative Emotionist account of emotional action explanation, grounded in the Motivational Theory of Emotions that one of us has developed (Scarantino 2014).

8 The Motivational Theory of Emotions

The core thesis of the Motivational Theory of Emotions (MTE) is that emotions are action control systems designed to prioritize the pursuit of certain goals over others. Here, we summarize two components of MTE: (1) prioritized action and inaction tendencies, (2) the two-level control structure of emotional actions.[6]

Before getting into the details of MTE, we wish to emphasize that the solution to the problem of arational actions we offer in this paper can survive, *mutatis mutandis*, even if one does not accept MTE in its entirety.[7] What is required for our solution to go through is essentially the acceptance that an emotion involves an action tendency with strong motivational power, either constitutively or by virtue of a reliable causal connection, and that the same action tendency can give rise to different actions.

This is a significant advantage of our solution, because it makes it in principle compatible not only with theories that are similar in spirit to ours (e.g. Deonna and Teroni 2012) but also with theories whose shortcomings have motivated the emergence of MTE, i.e. perceptual theories (e.g. Prinz 2004) and cognitivist theories (e.g. Marks 1982; Nussbaum 2001; Solomon 2003). This being said, MTE offers the best platform for the solution we propose because it does not require problematic explanations of how emotions come to involve an action tendency with strong motivational power (see our discussion in the preceding section, and Scarantino 2014 for a broader discussion of the advantages of MTE over competing theories in explaining emotional motivation).

---

[6] For a description of other features of MTE, most importantly of how it accounts for the intentionality of emotions and for the type identification of different emotions, see Scarantino (2014).

[7] We thank an anonymous referee for emphasizing this point to us.

Author's personal copy

MTE extends Frijda's (1986) theory of emotions according to which emotions are types of *action tendencies*, understood as "states of readiness to execute a given kind of action", where the kind of action "is defined by…[the] end result aimed at" (p. 70). We refer to the end result as the *relational goal* of the emotion. For example, fear has the relational goal of avoiding a certain target appraised as dangerous, anger has the relational goal of removing a certain obstacle appraised as blameworthy, disgust has the relational goal of removing an object appraised as noxious, and so on.

On Frijda's view, action tendencies become emotions when they acquire *control precedence*:

> Action tendencies – and action readiness changes generally – clamor for attention and for execution. They lie in waiting for signs that they can or may be executed; they, and their execution, tend to persist in the face of interruptions; they tend to interrupt other ongoing programs and actions; and they tend to preempt the information-processing facilities…Evidently, then, action tendencies are programs that have a place of precedence in the control of action and of information processing. We therefore say: Action tendencies – action readiness changes generally – have the feature of *control precedence* (Frijda 1986, p. 78).

On this view, fear involves an avoidance tendency with control precedence, anger involves an attack tendency with control precedence, disgust involves an expelling tendency with control precedence, and so on. Now, not every emotion involves a prioritized *action* tendency. Grief, for instance, seems to *reduce*, rather than increase, one's willingness to pursue specific relational goals. MTE understands grief as an *inaction tendency*, namely a tendency associated with the relational goal of *not relating as such*. As a result, the grieving person is in a state of generalized disengagement from the world.[8] This inaction tendency is also prioritized, in the sense that it acquires control over the organism, tends to persist in the face of interruptions, and pre-empts information processing. In what follows, we will think of emotions as special kinds of prioritizing action control systems, which can be expressed either by action or inaction tendencies with control precedence.[9]

The second component of MTE that we want to highlight is that emotional actions are generated by a two-level control structure. The first level is constituted by the prioritized (in)action tendencies we have just discussed. The second level is constituted by a set of general-purpose capacities that determine whether and how the prioritized (in)action tendencies are manifested. We call this set of capacities

[8] We are here focusing on the lack of physical actions characteristic of grief. On the other hand, grief may involve significant mental actions. For a discussion of the complex combination of mental activity and physical passivity characteristic of some emotions involving inaction tendencies, see for instance Jacobs et al.'s (2014) paper on depression.

[9] The full-fledged version of MTE also includes reflexive actions, a point we will disregard in what follows (see Scarantino 2014).

Author's personal copy

rational control.[10] We consider the interplay between prioritized (in)action tendencies and rational control to be the crucial design feature of emotions, which (generally) allow for input in the course of execution, making it possible to flexibly adapt to the circumstances at hand.

Rational control involves a *deliberative phase* and an *executive phase*, both constrained by the fact that a prioritized (in)action tendency is up and running while the emoter is trying to exercise control. In the deliberative phase, the emoter must determine *whether* the relational goal of the emotion should be pursued and, if so, *how* it should be pursued, translating the abstract goal of the emotion (e.g. attacking an opponent) into a set of situated sub-goals that achieve the abstract goal in a concrete context (e.g. picking up a bottle from a nearby table and hitting the opponent on the head with it).

An important objective here is to monitor that the emotion's relational goal and its situated sub-goals are achievable and compatible with the emoter's other goals and values. For example, an emoter may wonder whether getting back at her boss in some way after being insulted during a meeting is compatible with her career goals. Furthermore, the emoter may wonder whether a particular way of getting back at her boss—e.g. cursing at him in the middle of the meeting, slapping him after the meeting, etc.—is compatible with her moral values.

In the *executive phase*, the emoter must guide and control the implementation of the situated sub-goals that achieve the relational goal of the emotion, making sure the situated sub-goals are transformed into a proper sequence of motor sub-goals, that unforeseen contingencies are adapted to, and that side effects are taken into account in the course of execution. For example, the situated sub-goals of picking up a bottle from a nearby table and hitting the opponent on the head with it must be turned into a precise sequence of executable motor commands, the movements of the opponent must be adapted to, and possible side-effects (e.g. the appearance of a gun in the hands of the opponent) considered.

This process is different from standard practical reasoning in two respects. First, the relational goal is *given* to the emoter by the automatic activation of a prioritized action tendency, rather than *chosen* through a process of deliberation in which different courses of action are considered in light of one's beliefs and desires. Second, both the deliberative and the executive phase take place while an action tendency with control precedence is up and running. As the degree of control precedence of the prioritized (in)action tendency increases, the scope and effectiveness of rational control decreases and in some cases is crowded out all together, as demonstrated by emotional reflexes such as recoiling in fear from a suddenly looming object.

With these components of MTE in place, we can finally turn to the explanation of arational actions.

[10] See D'Arms and Jacobson (2006) for a philosophical analysis of emotional regulation, and Gross (1998) for a psychological analysis of emotional regulation.

Author's personal copy

9 Arational actions explained

Our solution to the problem of arational actions is *non-Humean*, *non-primitivist* and *pluralist*. It is *non-Humean* because, just like Hursthouse and Doring, we take emotions to be sufficient on their own to explain arational actions. It is *non-primitivist*, because unlike Hursthouse and Doring we do not simply say that arational actions are explained by being in the grip of an emotion or a feeling, with the latter understood as primitives. It is *pluralist*, finally, because we do not think arational actions constitute a unified category: different emotion-based explanations are required for different arational actions.

We introduce a general distinction between two ways in which emotions explain actions. In some cases, emotional actions are ways to achieve the *relational goal* of an emotion. We call these *instrumental emotional actions*. In other cases, emotional actions are ways to achieve a goal that is symbolically related to the relational goal of the emotion, or that diverts attention away from the relational goal of the emotion. We call these *displaced emotional actions*. Some of the arational actions in Hursthouse's original list are instrumental emotional actions, whereas others are displaced emotional actions. Garden-variety emotional actions, on the other hand, are all of the instrumental variety.

Let us now introduce some of the details of our proposal.

9.1 Instrumental emotional actions

Instrumental emotional actions (IEAs) occur when the emoter pursues the relational goal of an emotion, translates the goal into a sequence of situated sub-goals, and successfully guides and monitors the sub-goals' execution to completion. IEAs include two types of emotional actions. We call the first *directly instrumental emotional actions*, which are garden-variety emotional actions like running away out of fear or striking someone out of anger. These actions are directly instrumental because running away is a direct means of achieving the relational goal of fear, namely avoiding an object appraised as dangerous, and striking is a direct means of achieving the relational goal of anger, namely removing an obstacle appraised as blameworthy.

We call the second type of IEA *communicatively instrumental emotional actions*. These are emotional actions that achieve the relational goal of the emotion by means of an act of communication, namely an act that influences the behavior of a recipient by means of information transfer (Scarantino 2013). For instance, instead of striking someone out of anger, an angry person may engage in the communicative act of punching a hole into a wall in full view of the opponent. Punching a hole into a wall is not a direct means of achieving the relational goal of removing the obstacle constituted by the opponent, but it can become an indirect means of doing so if the opponent reacts to the reception of information about the emoter's strength and aggressiveness by retreating.

We think that several of Hursthouse's arational actions [1]–[6] are either directly or communicatively instrumental emotional actions. This marks an important discontinuity with other Emotionist models of arational actions, which have

Author's personal copy

explicitly denied that arational actions can be means to ends. Consider rumpling someone's hair out of love [1]. We take this action to satisfy a direct sub-goal of the relational goal of love, which we understand as an "urge toward proximity seeking in which proximity as such is the satisfaction" (Frijda 1986, p. 83). Rumpling someone's hair achieves this sort of proximity by making it physically embodied, and so it is directly instrumental.

Note that this is a different sense of instrumentality from the one contemplated by Hursthouse and Goldie, who have pointed out that we occasionally rumple someone's hair in order to express our love or achieve pleasure. Their view is that, in such cases, the emotional action is not arational and the Humean explanation in terms of belief and desire pairs is fully adequate. Our claim is different, namely that in standard cases we rumple someone's hair in order to achieve proximity with them, which may or may not express our love and lead to pleasure, but is not done *in order* to express love or achieve pleasure.

Another example of a directly instrumental emotional action may be covering one's face in the dark out of shame [6]. We follow Lewis (2008) and others in taking the relational goal of shame to be hiding a self appraised as flawed. Covering one's face is a way to hide from the world by removing the world from sight. Admittedly, this does not fully explain the action under consideration, because being in the dark *already* removes the world from sight. This is where other resources of the Motivational Theory of Emotions can help us.

Shame is not just a tendency to hide a flawed self: it is a *prioritized* tendency to do so, namely a tendency that clamors for attention and execution. This explains why emotional actions are impulsive, which is to say urgent and marked by partial informational access. So the action of hiding one's face out of shame while already in the dark may qualify as *postfunctional*: its characteristic urgency does not abate even after the relational goal of shame has been achieved, just like the urgency of killing a dog may not abate after the dog is clearly dead. Another factor likely at play here is that, due to the control precedence of the hiding tendency of shame, the shameful person displays cognitive biases in the way information is processed, which may also account for why he or she is not realizing that being in the dark is already a form of hiding.

Consider now the case of jumping up and down out of joy [4]. We understand joy as having the relational goal of undifferentiated engagement with the world. The joyful person is ready to engage in an open range of actions and actively prepares for this open engagement with a generalized state of arousal. By jumping up and down, the joyful person communicates an intent to interact, and this can lead to a shared moment (Scarantino 2014). As Frijda puts it, "joyful behavior entraps the other in an interaction" (p. 27). We take this to be another example of a communicatively instrumental emotional action.[11]

[11] Researchers of positive emotions have recently emphasized the functional aspects of an undifferentiated readiness to engage with the world, suggesting that it ultimately leads the emoter to "broaden" habitual ways of thinking and acting, and as a result "builds enduring personal resources" (Fredrickson and Cohn 2008, 782).

Author's personal copy

9.2 Displaced emotional actions

The main focus of the debate on emotional actions has been on the most puzzling items on Hursthouse's list: gouging holes in someone's picture [2], kicking a door [3] and rolling around in one's dead wife's clothes [5]. In these cases, the action performed does not satisfy the relational goal of the emotion that caused it, and so the instrumental explanation we have outlined so far does not apply. To account for these emotional actions, we now introduce the notion of *displaced emotional actions* (DEAs).

Our account of DEAs emerges from adding to MTE's framework insights from Goldie and Kovach and DeLancey. Goldie points out that there are often *symbolic relations* between the objects of emotions and the objects of emotional actions. In [2], Joan's photo (the object of Jane's emotional action) straightforwardly *represents* Joan (the object of Joan's hatred), so that when Jane scratches out the eyes in the picture, this action symbolizes her scratching out Joan's eyes. Other cases of straightforward representation include Voodoo dolls resembling the person to be cursed, statuettes and wax effigies, all of which can be the objects of emotional actions (e.g. putting needles in a Voodoo doll out of hatred, sleeping with a statuette of one's lover in the bedroom, etc.).

Goldie (2000a, p. 132) also considers other possible symbolic relations, such as the one of part-to-whole or whole-to-part (synecdoche), or the one of property or relatum of a thing standing for the thing itself (metonymy). For example, when the pope washes a poor man's feet out of compassion during Christmas celebrations, the poor man's feet may stand for the whole of human suffering. Or when someone kisses a lover's ring out of love, the ring may stand for the lover herself.

We think it is important to emphasize that for a symbolization process to occur there need not be either a relation of iconic similarity (as in the picture's case), or a relation of part-to-whole (as in the case of a poor man's feet), or a relation of being a property or relatum of the thing being symbolized (as in the ring's case). We often invest objects with symbolic value in a completely arbitrary way through an act of imagining, as when we imagine that a random stone is our enemy and kick it out of anger, or when we transform a random piece of soap into the hand of a lover and caress it out of love.[12] In order for a symbolic process to take place one must simply take a certain object to stand for another, perhaps only for the purposes of a fleeting emotional action.

Also, we do not follow Goldie in emphasizing "civilizing constraints" as the primary engine for the emergence of the symbolic processes involved in some arational actions. A symbolization process may have many other causes that have nothing to do with civilizing constraints. Examples include the absence of a loved one (which may lead to kissing her ring), the physical impossibility of hurting someone who is already dead (which may lead to shooting his car), prudential considerations (which may lead to putting needles in a Voodoo doll instead of actually hurting someone that may retaliate), and many other reasons.

[12] We thank Achim Stephan for helpful discussions of this point.

Author's personal copy

We pointed out earlier that Goldie's analysis does not apply to arational actions in which a symbolic relation between the object of the action and the object of the emotion is missing. Goldie (2000b) suggested that in such cases "the better explanation…will be the cruder one of just "venting" your emotion on the nearest thing to hand" (p. 31). We think we can improve on this cruder explanation and unify symbolic and non-symbolic instances of non-instrumental emotional actions by incorporating the ethological concept of displacement.

Kovach and De Lancey (2005) were the first to mention displacement in this context, comparing Jane's gouging the eyes in Joan's photo to an angry lioness attacking a tree trunk after being bitten by her cub. Since "an aggressive response to the cub has been redirected at another target" (p. 119), this is a case of displaced (or redirected) action. We find this suggestion helpful, but more needs to be said about *why* displaced actions come about.

Tinbergen (1939) and Kortlandt (1940) introduced the notion of displacement to make sense of animal behaviors that appear to be "out-of-context" (Zeigler 1964). Examples of such behaviors include the following: domestic fighting cocks exhibit displacement feeding, suddenly pecking at the ground; some bird species exhibit displacement preening when copulation is prevented or interrupted; male sticklebacks exhibit displacement digging when they encounter another male at the border of their two territories (Tinbergen 1951).

Displacement actions emerge in two types of situations (Zeigler 1964). In the first, the animal experiences two or more competing and incompatible motivations. Call this a *conflict situation*. An example is the male stickleback's displacement digging. Because he is at the border of two territories, he is motivated to both attack and flee. These competing motivations give rise to the displaced action of digging. In the second situation, the animal is prevented from completing whatever it is motivated to do. Call this a *thwarting situation*. For instance, when a male stickleback is prevented from copulating due to insufficient female response, he "invariably shows nest-ventilating movements" (Tinbergen 1951, p. 116).

As Zeigler (1964) notes, in all cases of displacement behavior in animals "a specific response tendency has been activated but its overt manifestation has somehow been blocked" (p. 366). Drawing an analogy between emotional actions and displacement actions requires accounting for the *blocking* of emotional response tendencies. The Motivational Theory of Emotions is ideally suited for this task. MTE identifies emotions with prioritized action tendencies and makes room for impeding the satisfaction of the relational goal of the emotion or of one of its relational sub-goals. This can happen during the deliberative phase in which emoters monitor that relational goals and sub-goals are achievable and compatible with their other goals and values. Goldie's civilizing constraints play a role at this juncture because emoters must determine whether a specific emotional expression meets the standards of civilized life. The presence of civilizing constraints can lead to the emotional analogue of a *conflict situation*, which comes about when the relational goal or sub-goals of the emotion are deemed incompatible with the value system of the emoter.

Author's personal copy

As mentioned above, however, Goldie's civilizing constraints are not the only causes of conflict situations. Conflict situations can also emerge when, for example, two conflicting prioritized action tendencies are up and running at the same time. Imagine being mugged in an alley while on a date. This situation is likely to elicit both fear—a prioritized avoidance tendency—and anger—a prioritized attack tendency. Since the two tendencies cannot be manifested at the same time, a conflict ensues which leads to the blocking of at least one of the two tendencies.

Blocking can also emerge from emotional analogues of *thwarting situations*, which may or may not involve civilizing constraints. As mentioned before, in some cases the relational goal of the emotion won't be achievable at all. If one's wife dies, achieving proximity with her will become unachievable, thwarting the relational goal of love. If one's enemy dies, hurting him will no longer be possible, thwarting the relational goal of hatred.

Thwarting and conflict situations can lead to two types of DEAs. We call the first *symbolically displaced emotional action*. This is an emotional action whose object stands in a symbolic relation to the object of the emotion that caused it. Some of Hursthouse's arational actions are symbolically displaced. For example, gouging a rival's eyes is incompatible with most emoters' value systems. But gouging the eyes in someone's photo is compatible, and allows the symbolic satisfaction of the relational goal of hatred.

Rolling around in one's dead wife's clothes [5] is another example of a symbolically displaced emotional action. As Bonanno et al. (2008) have emphasized, grief is a "complex and enduring molar experience that generates various molecular components, including a range of specific emotions" (p. 798). These grief-induced emotions, sometimes described as belonging to different "phases" of grief, include both negative emotions such as sadness, guilt and anger, and positive emotions such as amusement, affection, happiness and pride. As a result, acting out of grief often amounts to acting out of one of the negative or positive emotions that grief generates. We think this is typically the case with respect to rolling around in one's dead wife's clothes, an action we explain as being caused by grief-induced affection or love. Since the emoter cannot achieve literal proximity to his deceased wife, the relational goal of his love is thwarted, and he resorts to the symbolically displaced action of rolling around in his wife's clothes.

Neither instrumental emotional actions nor symbolically displaced emotional actions can account for the cases Goldie labeled as "venting". To make sense of those, we need to introduce what we call *radically displaced emotional actions*. These are actions whose object does not stand either in an instrumental or in a symbolic relation to the object of the emotion that caused it. A prime example is that of kicking a door out of anger [3]. Other examples may include smoking a cigarette out of anxiety or adjusting one's tie out of fear prior to giving an important job talk.

Radically displaced emotional actions are the closest analogues of displacement behaviors in non-human animals, since the latter lack the cognitive resources required to engage in symbolically displaced emotional actions. Human beings do have such resources, but occasionally they do not deploy them, resorting instead to actions that are radically displaced in the sense that they are completely unrelated to

Author's personal copy

the relational goals of the emotions that cause them. So, we understand the action of kicking a door out of anger as being similar to the male stickleback's displaced nest-ventilating behavior.

One might object that grouping symbolic actions like [2] and [5] together with non-symbolic actions like [3] under the category of displaced emotional action is unmotivated.[13] The justification for this grouping, however, is the fact that all of [2], [3], and [5] arise from emotional analogues of conflict and thwarting situations, which are the two primary cases of displaced actions in the animal realm. It's hard to deny, after all, that Jane's scratching the photo of Joan is preceded by a conflict between her hatred of Joan and the social norms that she endorses, which proscribe scratching hated rivals; and kicking a door out of anger after getting an upsetting phone call from one's boss takes place, in part, because the object of one's anger (the boss) is not present, and hence not available as an object of emotional action. These uncontroversial observations, together with the well-confirmed ethological concepts described above, motivate our proposal that [2], [3], and [5] are all species of the same genus, namely displaced emotional actions.

Our response to this objection also helps to clarify how our account differs from and improves upon Goldie's. Although Goldie insightfully identified symbolic emotional actions, he failed to recognize that such actions share important features with certain non-symbolic emotional actions, like [3], which he lumped together under the catch-all term "venting". Our account corrects this shortcoming in Goldie's analysis by arguing that both "vented" emotional actions and symbolic emotional actions are types of displaced emotional actions. Finally, we have expanded on what counts as a symbolic emotional action (an act of imagining can make any arbitrary object stand for anything one wishes) while reducing the centrality of "civilizing constraints" in the symbolization process.

Understanding the causal mechanisms underlying radically displaced emotional actions is going to be a significant challenge, in part because there is no consensus concerning the causal mechanisms underlying displaced actions in general (Anselme 2008). Tinbergen's original proposal was based on a hydraulic model rooted in drive theory: when an animal has a surplus of motivational energy that can't discharge itself via normal pathways, the motivational energy "sparks over". Andrew (1956) proposed instead a disinhibition model: when two (or more) motivations are in conflict, they inhibit each other and disinhibit a third motivation, which causes displacement behavior. A third proposal is that displacement behaviors are caused by the anxiety that results from being in conflict and thwarting situations (Maestripieri et al.1992). We won't enter this empirical debate here, but we suggest that whatever model ultimately emerges in ethology as best supported by the empirical data ought to be applied to radically displaced emotional actions as well.

[13] We thank an anonymous reviewer for raising this objection.

Author's personal copy

We conclude with a chart that recapitulates our account of the categories of arational actions:

Instrumental emotional actions		Displaced emotional actions	
Directly instrumental emotional actions	Rumpling someone's hair out of love [1], covering one's face in the dark out of shame [6]	Symbolically displaced emotional actions	Gouging holes in a photo out of hatred [2], rolling around in one's dead wife's clothes out of grief [5]
Communicatively instrumental emotional actions	Jumping up and down out of joy [4], punching hole in a wall out of anger	Radically displaced emotional actions	Kicking the door out of anger [3], fidgeting with one's tie out of fear

Although we have so far made the simplifying assumption that each token of an emotional action type has a unique explanation, things are bound to be more complicated. The same token of, say, punching a hole in a wall may be motivated both by the goals of communicating strength (communicatively instrumental) and venting (radically displaced). Similarly, kicking a door out of anger may be motivated both by the goals of symbolically hurting another (symbolically displaced) and venting (radically displaced). This is to say that real world emotional actions will often best be explained by a combination of the various models we have introduced.

Where does this all leave the term "arational"? We think the term is ultimately inadequate, and should be used only as a chapter or paper heading: it wrongly suggests that actions of the sort we have discussed are done for "no reason". On the contrary, we have unveiled various reasons why they are performed, which include pursuing the goal of an emotion or a goal symbolically related to it or simply venting. What remains importantly true is that none of such reasons can be construed in Humean terms, namely as combinations of beliefs and desires.

10 Conclusion

We have argued that Humean models of action explanation can't adequately explain emotional actions: the belief and desire pairs that Humeans cite to explain emotional actions fail to account for the emotionality of such actions. We have proposed an alternative Emotionist model, grounded in a well-developed theory of emotions, MTE, according to which emotions are prioritized (in)action tendencies. We have used this model, together with the ethological concept of displacement, to shed new light on Hursthouse's arational actions. Our analysis shows that such actions are either instrumental or displaced emotional actions. Both types of action are explained by the interplay between a prioritized (in)action tendency and rational control, rather than by a Humean belief and desire pair (as suggested by Humeans) or by the brute fact of being "in the grip" of an emotion or a feeling (as suggested by other Emotionist models).

Author's personal copy

Acknowledgments We thank Achim Stephan and the students of his "Research Seminar: Motivational Theory of Emotions" (Spring 2015) at the University of Osnabrueck, who have provided very helpful feedback on a previous draft of this paper. We also thank an anonymous referee of this journal and Nathan Dahlberg for critiques and suggestions that have helped us sharpen the paper. Finally, Andrea Scarantino thanks the Alexander von Humboldt Foundation for funding in the Spring 2015 semester.

References

Andrew, R. J. (1956). Normal and irrelevant toilet behaviour. *Emberiza Spp. The British Journal of Animal Behaviour, 4*(3), 85–91.

Anselme, P. (2008). Abnormal patterns of displacement activities: A review and reinterpretation. *Behavioural Processes, 79*(1), 48–58.

Benbaji, H. (2012). How is recalcitrant emotion possible? *Australasian Journal of Philosophy*. doi:10.1080/00048402.2012.699078

Berridge, K., & Winkielman, P. (2003). What is an unconscious emotion?(The case for unconscious 'liking"). *Cognition and Emotion, 17*(2), 181–211.

Bonanno, George A., Goorin, Laura, & Coifman, Karin G. (2008). Sadness and grief. *Handbook of Emotions, 3*, 797–806.

D'Arms, J., & Jacobson, D. (2003). The significance of recalcitrant emotions (or anti-quasijudgmentalism). In A. Hatzimoysis (Ed.), *Philosophy and the emotions*. Cambridge: Cambridge University Press.

D'Arms, J., & Jacobson, D. (2006). Anthropocentric constraints on human value. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 1, pp. 99–126). Oxford: Clarendon Press.

Damasio, A. R. (1994). *Descartes error: Emotion, reason and the human brain*. New York: Grosset/Putnam.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Harcourt Brace.

DeLancey, C. (2002). *Passionate engines: What emotions reveal about mind and artificial intelligence*. Oxford,: Oxford University Press.

Deonna, J., & Teroni, F. (2012). *The emotions. A philosophical introduction*. London: Routledge.

Döring, S. A. (2003). Explaining action by emotion. *The Philosophical Quarterly, 53*(211), 214–230.

Döring, S. A. (2007). Seeing what to do: Affective perception and rational motivation. *Dialectica, 61*(3), 363–394.

Ekman, P. (1999). Basic emotions. *Handbook of Cognition and Emotion, 4*, 5–60.

Elster, J. (2010). Emotional choice and rational choice. In P. Goldie (Ed.), *Oxford handbook of the philosophy of emotion* (pp. 263–281). Oxford: Oxford University Press.

Fodor, J. A. (1983). *The modularity of mind: An essay in faculty psychology*. Cambridge, MA: Bradford Books/MIT Press.

Fredrickson, B. L., & Cohn, M. A. (2008). Positive emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 777–796). New York: Guilford Press.

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

Frijda, N. H. (2007). *The laws of emotion*. Mahwah, NJ: Erlbaum.

Frijda, N. H. (2010). Impulsive action and motivation. *Biological Psychology, 84*(3), 570–579.

Gendler, T. (2008). Alief and belief. *Journal of Philosophy, 105*(10), 634.

Goldie, P. (2000a). *The emotions: A philosophical exploration*. Oxford: Oxford University Press.

Goldie, P. (2000b). Explaining expressions of emotion. *Mind, 109*(433), 25–38.

Gross, J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271–299.

Hatzimoysis, A. (2007). The case against unconscious emotions. *Analysis, 67*(4), 292–299.

Hursthouse, R. (1991). Arational actions. *The Journal of Philosophy, 88*(2), 57–68.

Jacobs, K., Stephan, A., Paskaleva-Yankova, A., & Wilutzky, W. (2014). Existential and atmospheric feelings in depressive comportment. *Philosophy, Psychiatry, & Psychology*, *21*(2), 89–110.

Kortlandt, A. (1940). Wechselwirkung zwischen instinkten. *Archives Néerlandaises de Zoologie, 4*, 442–520.

Kovach, A., & De Lancey, C. (2005). On emotions and the explanation of behavior. *Nous, 39*(1), 106–122.

Author's personal copy

Lacewing, M. (2007). Do unconscious emotions involve unconscious feelings? *Philosophical Psychology, 20*(1), 81–104.

Lewis, M. (2008). Self-conscious emotions: embarrassment, pride, shame, and guilt. In M. Lewis, J. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 742–756). New York: Guilford Press.

Loewenstein, G., & Lerner, J. (2003). The role of affect in decision-making. In R. J. Davidson (Ed.), *Handbook of affective sciences* (pp. 619–642). New York: Oxford University Press.

Maestripieri, D., Schino, G., Aureli, F., & Troisi, A. (1992). A modest proposal: Displacement activities as an indicator of emotions in primates. *Animal Behaviour, 44*(5), 967–979.

Marks, J. (1982). A theory of emotions. *Philosophical Studies, 42*, 227–242.

Mele, AR. (1998). Motivational strength. *Nous, 32*(1), 23–36.

Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge: Cambridge University Press.

Pacherie, É. (2002). The role of emotions in the explanation of action. *European Review of Philosophy, 5*, 53–91.

Prinz, J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford: Oxford University Press.

Scarantino, A. (2010). Insights and blindspots of the cognitivist theory of emotions. *British Journal for the Philosophy of Science*, *60*(3), 729–768.

Scarantino, A. (2013). Animal communication as information mediated influence. In U. Stegmann (Ed.), *Animal Communication Theory: Information and Influence* (pp. 63–81). Cambridge: Cambridge University Press.

Scarantino, A. (2014). The motivational theory of emotions. In D. Jacobson, & J. D'Arms (Eds.), *Moral Psychology and Human Agency* (pp. 156–185). Cambridge: Cambridge University Press.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind* . Cambridge: Cambridge University Press.

Sinhababu, N. (2009). The Humean theory of motivation reformulated and defended. *Philosophical Review, 118*, 465–500.

Smith, M. (1994). *The moral problem*. Oxford: Blackwell.

Smith, M. (1998). The possibility of philosophy of action. *Human action, deliberation and causation* (pp. 17–41). Netherlands: Springer.

Solomon, R. C. (2003). *Not passion's slave*. Oxford: Oxford University Press.

Tinbergen, N. (1939). On the analysis of social organization among vertebrates, with special reference to birds. *American Midland Naturalist*, *21*(1), 210–234.

Tinbergen, N. (1951). *The study of instinct*. Oxford: Oxford University Press.

Winkielman, P., & Berridge, K. C. (2004). Unconscious emotion. *Current Directions in Psychological Science, 13*(3), 120–123.

Zeigler, H. P. (1964). Displacement activity and motivational theory: A case study in the history of ethology. *Psychological Bulletin, 61*(5), 362.