# Genuinely collective emotions*

Bryce Huebner
Department of Philosophy
Georgetown University

**Abstract**: *It is received wisdom in philosophy and the cognitive sciences that individuals can be in emotional states but groups cannot. But why should we accept this view? In this paper, I argue that there is substantial philosophical and empirical support for the existence of collective emotions. Thus, while there is good reason to be skeptical about many ascriptions of collective emotion, I argue that some groups exhibit the computational complexity and informational integration required for being in genuinely emotional states.*

Does the United States *regret* its decision not to intervene in the Rwandan genocide? Were the Teamsters *angry* about the recent decision to open American borders to Mexican trucking companies? Did the Republican Party *feel upset* about its lack of success in the 2008 election? Most people would offer the same reply to all three questions: of course not! Commonsense, as well as the received view in philosophy and cognitive science, holds that individuals can have regrets, but the United States cannot; individuals can be angry, but the Teamsters cannot; and, although the members of the Republican Party were probably upset by these results, the Republican Party itself was not. The dominant view in philosophy, as well as in commonsense psychology, is that the collective term in such claims should be read as a plural term, and the accompanying statements treated as generic statements that call for a collective or distributed reading that adverts to the mental states of the individuals in these groups. But why should we accept the view that individual people can be in emotional states while groups of people cannot? My goal in this paper is to show that we should not be so willing to adopt this view. Indeed, I argue that there is substantial philosophical and empirical support for the existence of collective emotions.

## 1. Commonsense resistance to collective emotions

We often speak and write in ways that appear to ascribe emotions to a various human and non-human entities, as well as objects. "Susanne *regrets* her decision to live in this neighborhood"; "Germany *regrets* its genocidal past"; "My cat is *unhappy* when she finds her food bowl empty"; and, "My car *was angry* when I finally started it after over a year".

* This paper has had a long history, and has gone through many revisions on the basis of the compelling criticisms offered by numerous friends, acquaintances, and anonymous referees. Without the helpful criticisms that were voiced by these people, the central arguments in this paper would be far less compelling. Special thanks are due to Adam Arico, Bill Blattner, Dan Dennett, Aaron Garrett, Marcus Hedahl, Justin Junge, Joshua Knobe, Mark Lance, James Mattingly, Ram Neta, Mark Lange, Bill Lycan, Jesse Prinz, Andrea Scarantino, Mikko Salmela, Susanne Sreedhar, Justin Sytsma, and audiences at Georgetown University, Georgia State University, and The University of North Carolina – Chapel Hill.

But while such sentences occur in ordinary language, this does not establish that Susanne, Germany, my cat, or my car can be in emotional states. While some sentences that include mental states terms are intended as claims about *psychological states*, many others—despite their similarities in surface grammar—are not intended to convey anything that is literally true. As philosophers have long noticed, the ordinary usage of mental state terms ranges over a heterogeneous hodge-podge of genuine mental state ascriptions, dubious attempts at such ascriptions, instrumentally useful metaphors, and generic claims of various sorts (cf., Dennett, 1989). But, this leaves philosophers with a complex task of distinguishing those cases in which mental state ascriptions ought to be regarded as literally true from those cases in which these ascriptions are mere rhetorical flourish.

Of course, philosophical arguments have been marshaled to suggest that our practices of holding groups accountable for their actions, and expecting that groups should apologize for their morally problematic behavior, provide a plausible basis for positing collective guilt and remorse (cf., Gilbert 2001, 2002; Tollefsen 2006; and the essays in May and Hoffman 1991; see Kutz 2001 for a counterargument).[1] However, any appeal to collective emotions is likely to seem at least *prima facie* implausible. Faced with the suggestion that "Microsoft *believes* the time is right for world domination", it is reasonable to take the claim at face value, and doing so will likely yield plausible predictions about Microsoft's future behavior. However, faced with the suggestion that "Microsoft *feels melancholy* when reflecting on the loss of innocence that has accompanied its rise to power", it seems more reasonable to feel apprehensive and to demand an explanation. Of course, there are numerous differences between these claims, but the type of state that is being ascribed is clearly salient. Thus, although collective beliefs, desires, and intentions have seemed plausible enough to warrant philosophical attention, collective melancholia has seemed absurd enough to be rejected without reflection.

As described, we would expect Microsoft to feel a particular way, but it is not the sort of entity that can feel anything at all. With such worries in mind, Margaret Gilbert (2002, 119) argues that we must distinguish emotions from feelings, that feelings are inessential, though common concomitants of emotional states. This being the case, she argues that collective emotions require no specific phenomenology; they are unconscious emotions. On first blush, Gilbert seems to have the force of commonsense on her side. In a recent study, Joshua Knobe & Jesse Prinz (2007) found that people tend to judge that sentences like 'ACME Corporation *regrets* its recent decisions' sound OK, but sentences like 'ACME Corporation *feels upset*' sound weird.[2] But, while such data provide some support for

---

[1] A brief note regarding terminology: I use the term 'collective' in a more inclusive sense than it is typically used in the literature on collective intentions, plural subjects, and shared commitments. While discussions of collective mental states in philosophy tend to focus on states that are shared by groups of people, the position that I develop—building on theories of distributed cognition in the cognitive sciences—focuses on states and processes that include people as well as the technological structures in which they are embedded. As there is no obvious term that can be applied to groups *plus* technologies, I retain the term collective mentality to refer broadly to the states of groups as well as groups *plus* their technological environments.

[2] As has often been noted, the comparison between 'regrets' and 'feels upset' is not clearly the right comparison—and comparisons that differ only in the inclusion of the 'feels' locution must be examined to determine how robust this difference really is. Taking into account the various subtleties of mental state ascriptions, subsequent experiments have revealed that the phenomena are more complex than these initial data suggests. Justin Sytsma & Edouard Machery (2009) found that commonsense psychology draws a distinction between individual-appropriate and group-appropriate behaviors; Adam Arico found that contextual information has a significant impact on commonsense judgments about the sentences ascribing phenomenal states to groups; and my colleagues and I (Huebner et al 2010) found a significant difference between the ascriptions of mental states to groups in East Asian and Western settings using stimuli that used 'minimal pairs' of sentences that differed

Gilbert's claim, no one should be persuaded of the existence of collective emotions by statistical regularities in commonsense judgments. After all, even if some ascriptions of collective emotions 'sound plausible', and even if they are 'read literally',[3] there are likely to be plausible non-mentalistic and literal interpretations of such sentences. Such alternative interpretations must be ruled out before we attempt to draw ontological conclusions from these ordinary language data.

The most striking problem with ascriptions of collective emotion is that they fail to distinguish holistic and collectivist views of these mental states (cf., Pettit 1996). According to the *holist,* an individual's mental states depend on her social associations; according to the *collectivist* the groups *themselves* can be in genuinely mental states. As Robert Wilson (2001) argues, ascriptions of mental states often function as claims about certain psychological states of individuals that tend to be manifested only within the context of particular group relations. If individual mental states depend on social relations in this way, then they must be understood relationally, as opposed to as intrinsic properties of individuals. But this fact is far from sufficient for licensing the ontologically robust claim that some groups *as such* can be in mental states. Unfortunately, many arguments for collective emotion can only establish the weaker holistic claim.

Karl Jaspers (1947) argued that there was a sense in which every German should feel co-responsible for the atrocities perpetrated by the Third Reich. Building on this claim, Larry May argues that a 'metaphysical guilt' for collective actions arises where group membership yields a 'shared identity' and where a person "did not but could have (and should have) responded differently when faced with the harms committed by his or her fellow group members" (May 1991, 240). Similarly, Tollefsen (2006) defines collective guilt as "the guilt one feels in response to the harms committed by one's group". But none of these claims demonstrate the truth of the stronger collectivist claim. While collective guilt is clearly a social phenomenon, these descriptions of collective guilt suggest nothing more than individual emotions that are manifested "as social abilities, as ways of negotiating aspects of the social world" (Wilson 2004, 418). While there is little doubt that members of various groups commonly experience such feelings, there is no reason to suppose that these are emotional states of a collectivity.

To demonstrate the existence of such genuinely collective emotions, it is necessary to show that there are emotional states that are not merely states of individuals in aggregation. In attempting to develop a non-aggregative account of collective emotion, Margaret Gilbert (2001, 2002) distinguishes three types of 'guilt feelings' that are present in the context of a various groups. 'Personal feelings' are the result of actions that an individual *herself* carries out. There are many cases in which we must advert to social-relational properties in individuating these states; but individuals experience personal feelings because of *their own contributions* to some collective end. 'Membership feelings'

---

only in the agent to which a mental state was applied, though we also replicated Knobe & Prinz's original effect. Yet, as I have argued elsewhere (Huebner, 2010), commonsense judgments about mental states and processes are likely to be highly malleable and, as such, are more likely to tell us far more about ideology than ontology. That is, while such experiments might tell us something important about the default strategies that people employ in evaluating the plausibility of various sorts of mental state ascriptions, they are far less likely to be informative when we turn to questions about the furniture of the world (Thanks are due to an anonymous referee who pushed me to clarify my views on this point).

[3] Faced with such data, philosophers are likely to worry that the people in these experiments are interpreting these sentences in some metaphorical or figurative sense. After all, there are many ways of using emotion terms in meaningful ways without ascribing emotional states. However, when Arico and his colleagues (unpublished data) asked participants to categorize sentences as figurative (e.g., "Einstein was an egghead") or literal (e.g., "Carpenters build houses"), sentences that attributed mental states to individuals (e.g., "Some millionaires want tax cuts") and groups (e.g., "Some corporations want tax cuts"), tended to be categorized as 'literally true'.

are the result of collective actions even where the individual that feels them did not contribute to the specified action. Jasper's claim—and May's claim following him—is that a sort of moral taint affects the member of a group who did not respond as she should have to the morally problematic actions of her group. The fact that a person can say of her group that it carried out some action, while she was a member of that group, provides grounds for a feeling of guilt even in those cases where she did not *herself* behave in a morally problematic way. Finally, the non-aggregative alternative suggested by Gilbert (2001, 139) is termed 'collective feelings': For us collectively to feel guilt over our action A is for us to be jointly committed to feeling guilt as a body over our action *A*. This brings us closer to a genuinely collective emotion. However, even here it is unclear how such a state is supposed to be understood *as a collective emotion* as opposed to a distribution of individual states that are directly sustained by the causal and conceptual connections between the individuals in a collectivity.

On Gilbert's view, the emotional states of a 'plural subject' are understood as resulting from the commitments that each of the individuals has made to feel guilty or to feel remorse. On the basis of this model, non-aggregative collective emotions result from the ways in which the members of these groups coordinate their actions to produce some intentionally specified behavior. However, as Robert Rupert (2005) has argued, from the standpoint of psychological explanation, we gain no additional explanatory power by appealing to collective mental states in these cases.

> After all, every step in the construction of such representations, as well as every step in the causal sequence alleged to involve the effects of those representations, proceeds either by brute physical causation (e.g., photons emitted from the surface of the page stimulate the reader's retinal cells) or by causal processes involving the mental states of individuals (Rupert 2005, 5).

Regardless of how one chooses to spell out such a non-aggregative theory in terms of commitments to mutually realized plans (e.g., Bratman 1993; Gilbert 1989; Pettit 2003; Searle 1995), the semantic properties of these collective states will diverge radically from the psychosemantic properties of familiar genuinely mental representations. In the case of the 'collective feelings' for which Gilbert argues, there is a straightforward sense in which collective emotions are reducible to individual commitments and rules for their aggregation. However, in the case of an individual's emotion, no similar reduction is possible. Even though the subpersonal mechanisms responsible for the production of an emotion are likely to traffic in representations of some sort, it is highly unlikely that neural structures or computational subroutines are trafficking in commitments of any sort at all (cf. Velleman 1997). Following Rupert (2005), I argue that attempts to classify public language structures as *mental* representations are unlikely to provide any explanatory advantage beyond what is gained by a more thorough understanding of the psychological states of the individuals that compose a group, and a more thorough understanding of the decision procedures used in the formation of collective decisions. So, the key question is whether there are any purported collective emotions that could not be exhaustively characterized in terms of intentional states of individuals in aggregation.

## 2. A new approach to distributing cognition

While previous attempts to defend the possibility of collective emotions have appealed to facts about the structure of ordinary language and the presuppositions that we find in commonsense psychology, I propose a more revisionary approach to establishing the

existence of genuinely collective emotions. In what follows, I argue that some collective representations are produced in a way that parallels the integration of representations in an individual mind.[4] I begin from a philosophically and scientifically plausible view of individual mental states, and I show how this view can be extended to some collectivities. The view that I adopt is a version of the computational theory of mind (cf., Dennett, 1978; Fodor, 1980); although this theory has its opponents, its status as the received view in cognitive science—and in most corners of the philosophy of mind—is nearly as secure as the status of belief-desire psychology, which also has its opponents. On my favored approach, psychological explanations in terms of individual mental states require: 1) specifying a task that can be solved by an individual; 2) explaining how an individual solves this task by appealing to the operation of a set of computational systems; 3) explaining the operation of these computational systems in terms of the operation of simpler components; and, 4) repeating this process until a purely mechanical level of explanation is achieved (Dennett, 1978). So, for example, an individual's capacity to distinguish various visual stimuli begins by appeal to the operation of large-scale computational systems in the visual system, the operation of the visual system is then explained by decomposing it into the component systems by which it is implemented (e.g., edge and color detectors), and eventually the operation of these component systems will be explained in terms of the neuronal mechanisms and biochemical reactions that implement these computations.

This model of psychological explanation is old hat in the cognitive sciences; and, while it is typically assumed that the proprietary range of such homuncular decompositions begins at the level of individual actions, it is also standard fare to acknowledge that this model of reverse engineering allows for a high degree of implementational plasticity (Wilson, 1995). That is, philosophers tend to acknowledge the possibility of implementing the computational systems that are required for having a mind in a variety of media so long as the relevant range of counterfactual invariance is maintained to provide for psychological explanations. Crucially, this sort of reverse engineering allows for the possibility that the various divisions of the mind could

---

[4] I defend the possibility of collective representations in Huebner (forthcoming); I also offer an extended discussion of collective representations in Huebner (in prep). The details of these arguments would take us well beyond the scope of the current paper. However, given that the possibility of collective representation plays a crucial role in the current project, a brief characterization of my view is in order. I begin by adopting a set of rough-and-ready criteria for representation articulated by Haugeland (1998) and Clark (1997). With these criteria in hand, I argue that there are some collectivities that possess 'internal' states and processes (often in the form of local 'trading languages', cf., Gallison 1997) that have the function of adjusting collective behavior, in ways that are not fully specified in the design of collectivity. These internal states and processes facilitate skilful coping with novel changes in the environment, and I argue that they can also be deployed for modeling various non-present features of the environment. Finally, I argue that there are "proper (and improper) ways of producing, maintaining, modifying, and/or using the various representations under various environmental and other conditions" (cf., Clark, 1997, p. 147), such that collective misrepresentation is possible. In response to the charge that collective representations are explanatorily superfluous, I argue (Huebner 2008) that person-level representational states possess a constituent structure that roughly parallels the constituent structure displayed by the integrated representational states of some collectivities. That is, I argue both person-level and collective representations are implemented by highly integrated, hierarchically organized, and massively parallel computational systems that operate on lower-level representations by way of competitive algorithms to yield representational outputs that can be consumed by the higher level systems that they compose. I have no doubt that this discussion is too compressed to be compelling (though some of the details will become more clear over the course of this paper), so skeptical readers are invited to take my arguments as establishing only the conditional conclusion that *if* there are collective representations, then there can also be collective emotional representations as well.

themselves be realized by autonomous cognitive systems.[5] With this in mind, some cognitive scientists and philosophers have moved away from an exclusive focus on the implementation of mental states in aggregations of neurons to focus instead on the ways in which mental states can be implemented by collections of people: perhaps the flow of information through networks of individuals embedded in various technological environments also allows for the emergence of genuinely collective cognition (cf., Barber, 2006; Brooks, 1986; Giere, 2002; Hutchins, 1995a, 1995b; Knorr Cetina, 1999; Sutton, 2006). The vast majority of this research has focused on the sorts of computational states that are easily individuated by the behaviors that they produce. This research, however, leaves open the theoretical possibility of extending this model in a way that can accommodate the existence of genuinely collective emotions. Even on the assumption that this model of distributed cognition can establish the existence of distributed cognitive states, the possibility of collective emotions does not immediately follow; perhaps the computational systems required to implement emotions preclude the extension of such states to collectivities. I believe that there is room to develop an account of such genuinely collective emotions, and the remainder of this paper is dedicated to this task.

I begin with a brief account of the representational and computational capacities that are necessary to implement a paradigmatic case of individual fear. In this discussion, I assume that emotional states are *at least* representational states with intentional contents; regardless of what other constraints must be placed on emotion, I assume that such states must *at least* have the function of carrying some information about the world (Dretske, 1981, 1988; Millikan, 1984; Prinz, 2002).[6] I hold that fear has the function of carrying information about dangerous things in our environment and I argue that collective fear must also be a representational state that has the function of carrying information about danger. Building on these considerations, I turn to a frequently discussed case of distributed cognition and I argue that the crew of a naval vessel can implement many of the representational and computational capacities that are required for being in an emotional state. However, after addressing a series of difficulties with this proposed case of collective emotion, I argue, in the final section of this paper, that there is an alternative and more plausible case of a collective emotion.

## 3. A raccoon, a trashcan, and an emotion

Let me begin with a situation that might evoke fear in an ordinary person. Suppose you are walking home from the office one night, completely engrossed in the details of an argument, when you hear a rattle in a nearby trashcan. You look up and a raccoon bounds

---

[5] If cognitive scientists discovered that human neurons were autonomous cognitive systems, this discovery would not rule out the existence of individual minds. In fact, Tecumseh Fitch (2008) has recently offered an argument attempting to establish this claim. Fitch argues that neurons should be understood as possessing a sort of nano-intentionality, and that this nano-intentionality is a necessary precondition for the capacity of human organisms to have full-blooded intentional thoughts. In conversation, Marc Lange has suggested that there is an analogous argument regarding the existence of superorganisms. Microbiologists have provided good reason for thinking that mitochondria are alive; but this does not preclude cells, or even whole organisms, from also being alive as well. In fact, it seems to be a necessary condition on life as we know it that living things can be constructed out of other living things. More importantly, the fact that an organism is alive does not itself rule out the possibility that superorganisms are alive in the same sense. To rule out the possibility of life at the level of the superorganism, it is necessary to develop another sort of argument.

[6] I do not agree with the details of Prinz's (2004b) proposal. However, his use of Lazarus' core-relational themes as the formal objects of emotional representation holds a great deal of promise. I return to this hypothesis about emotional representation below.

from the trashcan, looks at you, and begins growling and chirping in a characteristic display of raccoon aggression. You perceive the bounding raccoon as dangerous, and as this happens various cognitive routines are recruited in considering possible routes of escape as well as possible means of defense. Your eyes widen. Visual and auditory attention is focused on the raccoon. You watch and listen for any movement that might suggest a threat to your future wellbeing. Motor routines are initiated to prepare a quick response if the raccoon decides to lunge at you. Your heart rate and respiration accelerate. Your muscles become tense. You stop digesting your dinner.

In this case, the fear representation that is produced is representative of a number of the core features of individual emotional representations. First, consider the type of content that we find in an emotional representation. Emotional representations have both formal and particular objects (Kenny, 1963; Prinz, 2004a, 2004b). In this case, the emotion *formally* represents the presence of danger, but it also represents a *particular* bounding raccoon as the source of that danger. While this claim may seem to be grounded in a thoroughly antiquated metaphysical view of the world, a wide range of neurophysiological data suggests that the human brain contains evolutionarily ancient structures that are dedicated to *detecting* emotionally salient stimuli *as such*, and that this perception of emotional salience occurs prior to the deployment of cognitive systems that are responsible for high-level conceptualization. There are, for example, neural pathways projecting from the sensory thalamus directly to the amygdala, which facilitate the rapid categorization of emotionally salient fear stimuli *as dangerous* on the basis of very little information (Ledoux, 1996). More importantly, the perception of a thing as dangerous does not even require the perception of that thing as a particular object. As Paul Whalen and his colleagues (1998) have shown, when volunteers are presented with emotional faces for 33ms, followed immediately by neutral faces presented for longer intervals, they exhibit increased hemodynamic activity in the region of the amygdala associated with the detection of fear faces even though they have no consciously available experience of the emotionally salient stimulus.[7] More strikingly, Whalen and his colleagues (2004) have demonstrated that the amount of white in a (subconsciously-processed) masked eye is sufficient to yield this activity. As Whalen often puts the point: the amygdala is a relatively dumb system, but it is very good at looking for danger and ringing the alarm when it finds it. In short, there are neurological structures that are responsible for the construction of low-level danger representations on the basis of relatively sparse information. They "trigger an emotional bodily response without the mediation of any kind of judgment. The relevant perceptual centers don't even support categorical object recognition, much less sophisticated appraisal" (Prinz 2004a, 46).

To be clear, these data do not speak to the person-level experience of fear; nor do they speak to the person-level representation of something as dangerous. Rather, they show that there are low-level representational systems that provide the computational resources from which such person-level representations can be constructed. These low-level representations stand in reliable causal relations to dangerous features of the world; and because of their evolutionary history, they have the function of conveying this information to computational systems that have the function of controlling and adjusting individual behavior. As the existence of horror movies and phobias clearly suggests, these systems can be 'tricked' into believing that there is something dangerous in the vicinity. However, under normal conditions these processes facilitate skilful coping with novel

[7] Numerous experiments confirm that this methodology successfully masks the initial stimulus and leaves only the second stimulus available for conscious processing; moreover, in post-experiment debriefings, these volunteers reported having seen only the neutral faces.

changes in the environment because of the way that they are harnessed to systems that are dedicated to the control of action. In understanding the operation of these systems, the computational story can be told exclusively at the level of subpersonal systems whose representations are either broadcast exogenously through the production of immediate coping behavior, or broadcast endogenously to conceptual systems that can provide the representational materials that are required by introspective systems. Thus, although these representations are never likely to be accessed *as such* from the first-person perspective, they are deeply integrated with computational systems that trigger further computations that can be readily monitored and evaluated in a way that can be subjectively recognized as fear. Such considerations provide a rough account of the implementation of the formal content of a fear representation. But, what of the particular objects of fear?

It is important to keep in mind that being in an emotional state can yield new beliefs about the world, as well as new desires about what ought to be done. In the case of a fear representation, subpersonal processes must focus person-level attention on dangerous objects, while at the same time diverting attention away from other cognitive tasks. When this happens, information about the danger that is posed by a particular threat also tends to be evaluated for its overall significance to the organism. But, this often requires recruiting higher-level conceptual systems that can engage in planning as well as evaluating the range of possible responses to a particular threat.[8] In deploying these higher-order systems, the threat is categorized as belonging to a particular kind of object or situation so that forward-models can be constructed and evaluated in establishing a plausible plan for dealing with the threat. At this point, the emotional representation can be said to have a particular object. By categorizing the threat and engaged higher-order representational systems to evaluate the sort of danger that is posed by an object or situation, conceptual systems are brought online to mobilize richer and more robust coping strategies (at the same time, producing further, non-conscious, evaluations of the probability of harm, the capacity to cope with the threat, and the urgency of coping with the threat).

While emotional representations have indicative content, they also have imperative force. This being the case, they tend to be implemented by reactive processes that lie beyond the realm of endogenous person-level control.[9] As Paul Griffiths and Andrea Scarantino (2005) put the point, emotions are action-oriented representations that reflexively reorient behavior in ways that facilitate coping with emotionally salient stimuli. Fear representations, for example, typically evoke behavior that causes one to avoid the threat posed by a dangerous stimulus. This occurs when low-level systems responsible for modifying cognitive routines and producing preparatory physiological responses are brought on-line to modify behavior in light of the perception of an emotionally salient stimulus. Such representations are typically deployed where engaging in quick-and-dirty computations can immediately facilitate action as this makes more ecological sense for an organism in a rapidly changing environment than does engaging

---

[8] Thanks are due to Marc Hauser for pushing this point. An emotional representation that merely tracked danger or proximity to indigestible items (disgust) would be unable to account for all of the ways in which emotions play a role in our overall cognitive economy.

[9] This claim has a rich philosophical pedigree. This is especially clear in the early modern philosophers (e.g., Descartes, Spinoza, and Hume), who used the Latin term '*patior*' to connote the idea that such states were something to be endured and submitted to. Even Stoic philosophers who notoriously treated emotions as types of judgments recognized that emotional states were something that one undergoes, as indicated by the use of the term 'pathê'.

in cognitively taxing reflection; and they typically activate motivational structures to produce reactions that are consistent with being threatened by a dangerous stimulus.

Different emotional representations have distinct action-tendencies that are produced immediately and unreflectively in light of the perception of an emotionally salient stimulus.[10] These action tendencies result, at least in part, from the physiological changes induced when an emotionally salient stimulus is detected. Such changes typically include an increase in proprioceptive monitoring, the tensing or relaxing of muscles, the increase or decrease of heart rate and respiration, and the secretion of various sorts of neurotransmitters and hormones in order to prepare the organism to act in response to the danger. In the case of fear, these bodily changes are initiated to insure that the organism will be prepared to fight, flee, or freeze in response to the threat. Emotions must, thus, be understood as reactive or irruptive states that are capable of radically reorienting various sorts of behavioral and cognitive processes. No one makes this claim more forcefully than Charles Darwin (1872/1965, 38):

> I put my face close to the thick glass-plate in front of a puff-adder in the Zoological Gardens, with the firm determination of not starting back if the snake struck at me; but, as soon as the blow was struck, my resolution went for nothing, and I jumped a yard or two backwards with astonishing rapidity. My will and reason were powerless against the imagination of a danger which had never been experienced.

The reason why this is the case, and this point is of the utmost importance in making sense of the possibility of collective emotions, is that many of the computational processes that are required for emotional representation are subpersonal mechanisms that lie beyond our capacities for person-level control. Because emotional representations are generated when quick-and-dirty computations are carried out to reorient the cognitive activity of an organism, emotional representations are not bound by the standards of deliberative rationality—whatever those might be. Emotional representations often diverge from the representations produced by more deliberative and reflective systems. This is why it does not matter how safe a person with a phobic fear of flying believes her airplane to be, her beliefs will typically be unable to override the fear representation produced at the thought of flying.

This brings me to a final observation about the underlying architecture of an individual fear representation. To yield a first-personal emotional representation, various subpersonal processes that are dedicated to the production of changes in cognitive processing and physiological states must be *coordinated and integrated* as a single representational state that is available to the introspective systems that produce person-level access to subpersonal representations. However, it is possible to explain, at the subpersonal level, how various subsystems access one another's representations. Conceptualized thoughts trigger the production of linguistic representations that are "either sent to exogenous broadcast systems (where they become the raw material for personal speech acts), or are endogenously broadcast to language comprehension systems which feed directly to the mindreading system" (Huebner & Dennett 2009, 149). Unless

[10] It is quite plausible to claim that fear representations have both intentional content and imperative force. However, things become less clear when you turn to a state like wistful melancholy for the world of one's childhood. However, even in this case, emotional representations play an integral role in producing the action tendencies required for interpersonal interactions. Wistful melancholy leads us to share memories with close friends and yields subtle behavioral-patterns that express important social cues. As Griffiths and Scarantino (2005) note, the orientation of social action is an integral aspect of most emotional representations.

this occurs, there is no way to generate the first-personal understanding of the emotional state as one of the simple and unified states that organisms like us often experience. Because the subpersonal processes that are responsible for responding to dangerous stimuli are typically integrated representations that can be treated as unified emotional states that are available for introspective monitoring, they tend to yield the sorts of meta-representations that allow us to experience ourselves as being in an emotional state. That is, these processes typically yield a conscious recognition that 'I am afraid', though they need not, and do not always do so.[11] In the heat of the moment, fear captures introspective attention in the service of constructing person-level coping strategies; this capture of attention leads us to be conscious of many of our emotional states (Clore, 1994; Nash, 1989).

Keeping these claims in mind, I suggest that a homuncular analysis of emotional representation is plausible. Successfully representing a raccoon as dangerous, while simultaneously preparing to respond to the threat of this animal, requires the reorientation and integration of numerous sub-personal computational routines.[12] Emotions are *fast*, *mandatory*, and to a significant degree, *cognitively impenetrable*,[13] suggesting the presence of a set of integrated, layered, and insular systems (Griffiths & Scarantino, 2005). This structure is implemented by a rich computational architecture that requires at least the following. First, a capacity to perceive things as dangerous; this in turn requires *a sensory system* that is capable of providing inputs to the relevant emotional circuits, *systems dedicated to recognizing emotionally salient stimuli* as such, and more robustly *conceptual systems* dedicated to interpreting the precise nature of the threat posed by a particular stimulus. Second, *attentional mechanisms* that can be deployed to focus sensory systems on the threatening stimulus as well as to focus *cognitive systems* on the computations required to develop strategies for overcoming the threat. Third, systems that are dedicated to producing and monitoring system-wide changes must be recruited in order to prepare the individual to cope with the emotionally salient stimuli (e.g., preparing the system to fight or flee). Finally, the output of these multiple systems must be *coordinated as a single state of fear* and systems dedicated to the introspective monitoring of this coordinated representation of fear must be mobilized in order to facilitate system-level coping strategies.

Although the internal structure of emotional representations is often obscured from the standpoint of first-person cognition, emotional representation requires the integration of a variety of computational mechanisms to yield practical activity. But, this provides a foundation for constructing an argument for the possibility of genuinely

---

[11] Richard Lane (2000) offers neurological evidence for the claim that distinct systems underwrite the production of a unified emotional representation that can be experienced and the actual experience of that emotion: dorsal anterior cingulate cortex (ACC) produces a unified emotional representations, but rostral ACC and medial pre-frontal cortex yield the awareness of the emotional state.

[12] This account of emotion has numerous predecessors and is likely to seem quite familiar. George Pitcher (1965), Nico Frijda (1986), Jon Elster (1999), and Paul Ekman (2003) all argue that a number of different representational capacities are required to account for paradigmatic emotional experiences. Klaus Scherer (1987) notes that the computational structure of emotional representation requires integrating various cognitive structures. Unlike these views, however, I do not claim that *what it is for something to be an emotion* is for it to be the result of these processes; I merely argue that representing the world as our emotions do requires that the outputs of a number of component mechanisms be integrated to generate a unified emotional representation.

[13] Of course, it does not follow that emotions are completely cognitively impenetrable. I suggest only that the operation of emotional mechanisms is, to a significant degree, outside of endogenous control. The degree to which the construction of emotional representation is malleable is an important question, but it is beyond the scope of this paper.

collective emotions. In the next section, I ask whether there are cases in which analogous computations are implemented in a distributed cognitive system that consists of multiple agents. My specific target is the well-known analysis of distributed cognition in the navigation crew of the *USS Palau* (Hutchins 1995a).

## 4. Collective fear on a naval vessel?

As the *USS Palau* was returning to port it began to lose pressure in its main steam drum, forcing the ship's engineer to shut the throttles. The ship was moving fast, and the only way to slow such a large vessel is by reversing its propellers. Without steam power, there was a significant threat of leaving deep waters, running aground, and damaging the ship (or destroying the port). Although Hutchins does not discuss the possibility, this seems like it could be the sort of dangerous situation in which collective fear might be produced (assuming that the *USS Palau* is itself a genuinely cognitive system). To see how a fear response might be produced in the crew of a naval vessel, it will help to briefly trace the response to the loss of steam power that occurred aboard the *USS Palau*.

The response began when the ship's engineer detected a loss of pressure in the main steam drum with no apparent cause. Recognizing that this was a potential danger, he shut the throttles and notified the bridge conning officer that he was doing so. He thereby produced a low-level danger representation. The bridge conning officer, recognizing that there were further implications of the loss of pressure for the ship's steering mechanisms, prepared the navigation crew to deal with the potential threat by ordering the helmsman to center the ship's rudders, an action that played an analogous role to physiological changes in the individual that are often classified as action tendencies (e.g., the modulation of muscle tension, heart rate and respiration in preparing to cope with a dangerous raccoon). At this point, additional computational processes were engaged to evaluate 1) the probability that the ship actually was in danger; 2) the capacity of the crew to cope with the threat; and 3) the urgency of adopting a coping strategy. The engineer noted that he was securing the backup boiler, and the ship's captain, noting that the ship was currently moving too fast to drop anchor, called the ship's bosun and told him to assemble a crew forward, ready to drop anchor if necessary.

In preparing to deal with this threat, the crew of the *USS Palau* faced a series of further computational difficulties. A number of the devices that are typically used for navigating the ship are powered by the steam turbines. So, the loss of steam power affected the capacity of the crew to produce an authoritative representation of the ship's location. When the navigational tools went off-line, the crew's attention had to be shifted to the chart that was used to keep track of the ship's location, adopting new cognitive strategies for tracking the location of the ship, and attention was diverted from the familiar strategies for keeping track of the ship's location. As the crew examined the previous course of the ship and the changes in rudder orientation, they were forced to adopt a series of overcompensations to ensure that the ship did not run aground. This yielded a slightly erratic path, similar in many respects to the behavior that we find when cognitive mechanisms are redeployed to deal with a dangerous raccoon. In such cases, a person often fails to focus on the relevant facts about her environment, and has to overcompensate for failing to see some obstacle that is in her way, until it is nearly too late. She might find herself tripping and falling to the ground unless she overcompensates for the loss of balance.

While the reorientation of computations and the modulation of the physical processes aboard the ship played an analogous role to the production of emotional action tendencies in an individual, this leaves open the possibility that the only representations

involved were the representational states of the individual crewmembers. Hutchins (1995, 117), however, argues that the representation of a ship's location can only be produced through the integration and coordination of a number of different lower-level representational systems. The navigation system of a ship consists of a number of sub-routines, each of which is sensitive to some one-dimensional constraint (Hutchins 1995, 118). To establish a complete representation of the ship's trajectory, speed, and location, these various representations must be coordinated and represented as a location on the map that can then be acted on by the captain. The captain of a modern naval vessel, however, only acts upon the representations that have been produced by the members of the navigation crew.[14]

The representational states and processes that we find among the navigation crew of the *USS Palau* operate do not require a ship-level experience of fear; nor do they require a ship-level representation of the dangerous situation. However, here we find representational and computational systems that have the function of conveying information about dangerous situations to other representational and computational systems that have the function of controlling and adjusting the ship's behavior. The integration of these representations facilitate skilful coping with novel changes in the environment because of they way that they are harnessed to systems dedicated to controlling the behavior of the ship; but this story can be told exclusively in terms of systems whose representations are either broadcast exogenously through the production of immediate coping behavior, or broadcast endogenously to the map to provide the representational materials that are required for evaluating the current state of the ship. Thus, although the representations that are distributed across the crew are never accessed by the ship *as such*, they are deeply integrated with computational systems that trigger further computations that can be readily monitored and evaluated in a way that parallels the production of an individual's fear representation. Moreover, these component representations yield new, highly distributed, beliefs about the environment and new global strategies for engaging with that environment. These representational systems also focus attention on relevant dangers and divert computational resources that would otherwise be focused on carrying out other tasks. This allows for the evaluation of the overall threat that is posed to the ship by recruiting person-level conceptual systems to both plan for and evaluate a range of possible responses. At this point, forward-models can be constructed to establish a plausible plan for dealing with the threat; having categorized the threat, and having engaged the computational systems that are employed in the evaluation of the sort of danger that is posed by an object or situation, further computations can then be initialized to mobilize robust coping strategies (while at the same time, further non-conscious evaluations of the probability of harm, the capacity to cope with the threat, and the urgency of coping with the threat continue to be carried out).

In this case, we find a computational architecture that shares a great deal in common with the computational architecture required for the production of a fear representation in an individual. The redeployment of attention, the reorientation of cognitive processes, and the production of action tendencies all play an important role generating the behavior of the system that consists of the ship and its crew. However, this representation

---

[14] Although the captain plays an important role in acting on the representations produced by other members of the crew, most of the computations that are necessary for navigating a modern naval vessel occur in the absence of the captain's orders. But I suggest that this too is analogous to individual cognition. First-person reflective consciousness is likely to come in only after most of the work has been done, and it is likely to play a less significant role than might typically be supposed. Unfortunately, I do not have the space here to further argue for this claim.

cannot be localized as the representational state of any member of the crew, nor can it be seen as the state of any apparatus aboard the ship. It is the state of the navigation crew as it is situated in the ship.[15] Strikingly, if the anthropological report of this incident is to be believed, none of the members of the crew exhibited fear, and it seems that there was something akin to a tone of boredom in the captain's voice, making his responses seem entirely routine (Hutchins, 1995, 2). Various component systems of the *USS Palau* were coordinated to yield a single collective representation that directed the immediate activity of the vessel even though no individual was ever afraid. But something does not seem right about ascribing a genuinely emotional state to the crew of the *USS Palau*.

## 5. Consciousness and collective emotions:

My claim that the crew of the *USS Palau* can be in an emotional state is a consequence of my argument that emotions can be exhaustively characterized in terms of the representational and computational processes that we find in a hierarchically organized and highly distributed system. But, as we all know, one philosopher's modus ponens is another's modus tollens; and indeed, it seems reasonable to suggest that since my claims about the computational structure of emotions yield the conclusion that a naval vessel can be afraid, this account of emotion is thereby implausible. I concede that such a response has a great deal of intuitive plausibility. After all, such a collectivity is unlikely to have a conscious experience of fear. So, if emotional states and processes are conscious states, then the implausibility of collective consciousness should impugn the possibility of collective emotions.

It does seem plausible to claim that emotional states are conscious states. If a person claimed that she was angry about racial profiling, but was left 'cold' by every report of racial profiling, it would seem more reasonable to deny her claim that she was angry than concede her anger and claim that it lacked the phenomenology that is typically associated with anger. Jaak Panksepp (2000, 138) has shown, "most people consider feelings to be among the most important aspects of emotion";[16] Robert Kraut (1986) argues that the commonsense understanding of emotion is, for this reason, to be identified with the commonsense understanding of a feeling; and, Max Bennett & Peter Hacker (2003, 214) argue that there is "no significant difference between *having an emotion* and *feeling an emotion*". From the perspective of emotion research, William James (1884) all but explicitly rules out emotions that cannot be felt by arguing that emotions are nothing more than bodily changes and the experience thereof; Sigmund Freud (1915/1950, 109) claims that it is "the essence of an emotion that we should be aware of it, i.e. that it should become known to consciousness"; and, Gerald Clore (1994, 285) argues that every emotion is f*elt and "*emotions that are felt cannot be unconscious by definition".

---

[15] I do not intend to make the familiar and dubious claim that if we wish to attribute a cognitive state to a collectivity, and no member of the collectivity is in that state, then the state *must be* a state of the collectivity. Rather, I only wish to suggest that the coordination and integration of the various representational states that are operative in this case share enough in common with the subpersonal computational architecture that we find in an individual fear representation to license the claim that the state of the ship is an emotional representation. Thanks to an anonymous referee for pushing me to clarify this point.

[16] There are also data to suggest that cognitive changes and autonomic changes are *also* among the most important aspects of emotion. But, more interestingly, this result is fairly plastic. Panksepp found that music majors were likely to treat feeling as the most important aspect of an emotional representation while philosophy majors were likely to treat cognitive changes as the most important aspect of an emotional representation. This, of course, points to one of the serious inadequacies of appealing to commonsense-psychological intuitions as evidence for metaphysically robust claims.

Summing up this dominant view, Richard Shweder (1991, 183) claims that even "three-year olds, Ifaluk islanders, and psychoanalysts (in other words, almost everyone, except perhaps the staunchest of positivists) recognizes that *emotions are feelings*".

There is no doubt that the claim that emotions require consciousness has a great deal of intuitive force and that it has garnered a great deal of theoretical support from both philosophers and scientists. Given that I have conceded that the navigation crew of the *USS Palau* is unlikely to be capable of being in conscious states (it is too spatially distributed, too insufficiently integrated, and its members communicate over low-bandwidth channels that are unlikely to instantiate consciousness).[17] So, defending the possibility of collective emotions requires a substantial argument for the claim that some emotional states are completely non-conscious; otherwise the fact that the *USS Palau* cannot be conscious immediately rules this out as a case of collective emotion.

To begin with, consider a case where commonsense psychology seems to allow for the possibility of unconscious emotions. Suppose that Natalja is discussing the trajectory of her current romantic relationship with her therapist, Milena. Milena listens carefully and then exclaims, "Natalja, don't you realize that you're afraid of commitment!" Not being one to be told that she is afraid when she doesn't feel afraid, Natalja responds: "Of course I'm not. I'd be happy to commit to Marko if it weren't for these exciting new opportunities at work. It's not that I'm afraid, it's just not the right time." Yet, Milena knows better. She knows that every time a relationship starts to get serious, Natalja begins to focus on routes of escape, she attends to every flaw in her partner's behavior (noting that they *never* remember what brand of tea she drinks and they *never* remember her favorite flavor of ice cream); she begins to look for changes in the relationship that might constrain her; she gets tense when talking about her relationship; and, she even forgets to eat. Although Natalja is keenly aware of the status of her current relationship, she is completely unaware of her fear of commitment.

Many people are likely to be familiar with such cases as people spend hundreds of thousands of dollars each year trying to navigate similar unconscious fears. Bookstore shelves are filled with self-help books that are intended to diagnose unconscious fears of commitment and failure, and they teach readers behavioral strategies for overcoming these fears. More importantly, while such fears are not consciously experienced, they are likely to be treated as emotions because 1) they have the right sorts of associated action tendencies, 2) they are largely automatic and lie largely outside of the realm of endogenous control, 3) they produce the right sorts of physiological changes, and 4) they modulate cognitive processing in the way that a conscious emotion would. In short, given the similarities in behavior, person-level cognition, and computational states and processes, it seems unreasonable to deny unconscious emotions of this sort (cf., Gilbert 2002). But if commonsense psychology allows for these sorts of unconscious emotions, then perhaps the intuitive resistance to collective emotions suggests little more than a failure of imagination. Since we are typically conscious of our emotions, we assume—and we assume far too quickly—that all emotions must be conscious; we try to imagine the crew of the *USS Palau* as *feeling* afraid, and we fail miserably.[18]

---

[17] Some theories of consciousness (Dennett, 1991; Lycan, 1987) need not rule out the possibility of collective consciousness. However, the mere possibility of collective consciousness is insufficient as a reply to the objection that there are no collective emotional representations in this naval vessel.

[18] Joshua Knobe informs me that the idea of collective happiness conjures up an image of the members of the Boston Red Sox arranging themselves in the shape of a smile—my guess is that the rest of us do no better in imagining a group, as such, exhibiting an emotional state.

Unfortunately, *pace* Gilbert (2002), this argument cannot succeed in undercutting the force of the worry about the implausibility of collective consciousness. To see why this is the case, consider two senses in which such psychoanalysis cases fail to establish the possibility of non-conscious emotions.[19] First, although Natalja's fear is *cognitively inaccessible*, there is an important sense in which there is something that it is like for her to be in a state of unconscious fear. The computational processes that are responsible for the production of her emotional representations are physiologically implemented. This being the case, they might yield phenomenal consciousness in the absence of awareness. As Prinz (2004b) puts the point, psychoanalysis examples can only establish that emotions can be inaccessible to consciousness, they cannot establish the existence of phenomenally unconscious emotions. However, no matter how rich the representational states and processes are among the crew of the *USS Palau*, there is never going to be anything that it is like to be a ship. If emotions must be phenomenally conscious, and psychoanalysis examples do not establish otherwise, then collective emotions will be impossible. Building on this worry, it might be argued that every case of an unconscious emotion can be brought to conscious awareness. However, no matter how rich the representational capacities of a collectivity are, they will never be brought to conscious awareness for *the ship*.

Prinz (2004b) is surely right that the experience of fear is implemented, at least in entities like us, by physiological changes that typically have an accompanying phenomenology. This being the case, at every time where a person is unconsciously afraid of commitment, there will be something that it is like to be that person. Moreover, this will be true even when she is unaware of her fear. However, this leaves open two distinct possibilities: either 1) there is something that it is like to be in an emotional state that is inaccessible to introspective and reporting mechanisms, or 2) there is something that it is like to be a person who is in a non-conscious emotional state, but there is nothing that it is like to be in that non-conscious emotional state. Psychoanalysis cases are consistent with both (1) and (2); so, establishing the existence of genuinely non-conscious emotions—of type (2)—requires a stronger argumentative strategy.

The way forward is suggested by an argument for the existence of non-conscious moods offered by Laura Sizer (2006). Sizer (2006) offers a case where a friend claims that you have been irritable all day. While there are many cases where we initially object to such characterizations, there are also cases where we agree after further reflection. The crucial thing to notice in cases where we agree is that the relevant sort of reflection does not involve examining conscious experiences over the course of the day to see whether you have, in fact, been experiencing irritable sensations.

> If you did, you would not find one that is uniquely indicative of irritability. Irritability and anxiety, for example, tend to feel very similar. But they are certainly very different moods. The relevant facts here are the pattern of thoughts and behaviors you have exhibited throughout the day. You realize that you have snapped at anyone who came near you, broken a few pencils and even now resent having your irritability pointed out to you. In other words, you recognize your mood by examining the patterns of thoughts and behaviors, by reflecting on how you have thought about and approached the world today, not by considering how you feel. (Sizer 2006, 132)

[19] Thanks to an anonymous referee for pressing me to clarify this point.

On the basis of this example, Sizer argues that moods are dissociable from the experiences of moods. Of course, such cases do nothing more than increase the intuitive plausibility of the claim that that there can be something that it is like to be a person who is the subject of an unconscious mood, even if the mood is not phenomenally conscious in any sense. I believe that a parallel argument can be constructed *mutatis mutandis* to demonstrate the intuitive plausibility of the claim that emotional states can sometimes be dissociated from the experiences of emotion. However, it is necessary to move beyond mere appeal to intuitively plausible cases to demonstrate why there are unconscious emotional states of the sort required to establish (2).

I concede that the production of a fear representation will yield numerous physiological changes, and I concede that many of these physiological changes will have an accompanying phenomenology. However, I argue that the disjointed experience of each of these changes in isolation cannot merely be 'summed up' to generate an aggregate phenomenology of fear. That is, while there is something that it is like to feel your heart race, to feel tension in your muscles, and to feel short of breath, the experience of these physiological changes alone is not sufficient to yield an experience *of fear*. To *feel afraid*, these experiences must be experienced as a unified state that is introspectively available as such—but we have little reason to suppose that every fear representation should be consciously available in this way. As I argued above, emotional states are implemented by a variety of sub-personal mechanisms that together have the function of conveying information about danger. These sub-personal mechanisms facilitate the adjustment of person-level behavior in ways that are sensitive to salient sorts of dangers. In this way they produce action-oriented representations that reflexively reorient behavior in ways that facilitate coping with emotionally salient stimuli. A wide range of empirical data from the cognitive science of emotion (e.g., Damasio, 1999; Ledoux 1996; Prinz, 2004a; Tsuchiya & Adolphs, 2007) suggest that these computational processes do not always recruit introspection, and so do not always yield a conscious recognition that we are in an emotional state. However, even where sub-personal representations are not accessed by introspective or reporting mechanisms, they still facilitate engagement with perceived threats by motivating behavior, focusing attention, and diverting computational resources away from other sorts of tasks. The crucial thing to notice at this point is that although such processes often recruit introspective mechanisms in a way that allows for the evaluation of the overall threat by way of person-level conceptual mechanisms that facilitate planning and evaluating responses, they need not do so. Keeping these facts in mind, there are likely to be at least some cases where a person can be afraid, although she does not *feel afraid,* and where there will still be something that it is like for her to feel as she does.[20]

---

[20] I take this as the most plausible reading of the well-known experiment by Schacter and Singer (1962) in which volunteers thought they were participating in a study on the effects of vitamin called Suproxin. Unbeknownst to the participants, the 'Suproxin' injections were really adrenaline injections. Some of the participants were informed that they might experience an increase in heart rate, blood flow, respiration, blood sugar, and lactic acid (the effects of adrenaline); some were told that they might experience headaches and numbness; the remainder were told nothing about potential side effects. The participants were then placed either in a 'happy condition' where they waited in another room where a stooge was flying paper airplanes and playing with hula-hoops, or they were placed in an angry condition where they waited in another room with a stooge who was becoming increasingly upset and where they were asked to fill out a survey containing probing personal questions. Schachter and Singer found no difference in the emotional states of those who anticipated the effects of the adrenaline injection. However, participants who were mislead or uninformed both felt and behaved in ways that associated with anger in the angry condition and in ways associated with happiness in the happy condition. Feeling the effects of the adrenaline, but having no explanation for the feeling, these subjects found themselves with the robust phenomenology of genuinely emotional states.

Put briefly, the various physiological changes that facilitate the production of coping strategies that are directed toward emotionally salient phenomena can operate in ways that will not yield a unified phenomenology of fear. That is, the sub-personal mechanisms that are responsible for detecting and engaging with dangers are not always accessed by introspective mechanisms. Of course, there is surely something that it is like for a person to feel as she does at every point where she is legitimately described as being unconsciously afraid. However, at least in some cases, the phenomenology of a person in a state of non-conscious fear will be exhaustively characterized in terms of what it is like for her to have her current eating habits, level of tension, and styles of cognitive processing. From the standpoint of first-person cognition, the phenomenology of various physiological changes might be present even though there is nothing going on that *feels* like an emotional state of fear.

Of course, someone who wanted to argue that these sorts of states are phenomenally conscious could still fall back to the claim that some phenomenal states are inaccessible from the first-person perspective *by any means whatsoever*. Such appeals are implausible.[21] However, if I were presented with the claim that phenomenal consciousness could be so radically dissociated from awareness as to be inaccessible to introspective monitoring and completely unreportable (cf., Block 2008), I would no longer see any reason to rule out the possibility of collective phenomenal consciousness. After all, once we have decided to entertain the possibility of phenomenal consciousness that is *in principle* inaccessible and completely unreportable, all empirical bets are off. Perhaps there are empirically undetectable high-bandwidth connections between the crew members of the *USS Palau* (or of some doppelganger *USS Palau*) that yield phenomenal consciousness in the absence of any awareness. Perhaps we are just so ill-informed about the physical substrates of phenomenal consciousness that even the crew of the *USS Palau* is phenomenally conscious and there is no way for us to know this from the standpoint of current empirical models of consciousness. I, for one, do not find such claims plausible. Of course, the members of the crew would not know that crew as a whole was phenomenally conscious, but this is utterly irrelevant. Neither the visual cortex, nor the amygdala, nor the fusiform gyrus knows that a person is phenomenally consciousness when she is presented with a stimulus; but this does not rule out the possibility of individual consciousness.[22]

I contend that if there is to be any force behind the claim that there are salient differences between the representations that are produced by the navigation crew of the *USS Palau* and the unconscious fear representations that are produced by the sub-personal mechanisms that we find in an individual, the argument must come from

---

[21] I do, however, address a version of this claim in (Huebner, in prep)

[22] If a representational or higher order theory of consciousness is correct, it will be consistent with our best psychological theory that collectivities can be conscious where they have the right functional organization. Someone might, however, argue that although higher order and representational theories of consciousness explain access to mental states, they do not explain phenomenal content, so they fail to account for what it is like to be in a phenomenal state. Where such claims rely on intuitions about the possibility of absent qualia or inverted qualia, however, they are likely to face difficulties. As David Chalmers (1996) convincingly argues, absent qualia and inverted qualia intuitions have the untoward consequence of allowing for fading qualia and dancing qualia. But, if fading qualia and dancing qualia are possible, then first-person introspective reports about phenomenology will have no evidentiary value because such reports will not be able to track changes in phenomenal content. More importantly, the scientific study of consciousness is possible only on the assumption that people are typically correct in their introspective claims about phenomenal states. So, unless one is willing to adopt a very strong dualism about the mind, a dualism strong enough to rule out the possibility of the scientific study of consciousness, arguments against collective emotion that rely on the potential consciousness of all emotion will be deeply problematic.

another tack. Perhaps the fact that individuals can be conscious of their fear representations has significant implications for the classification of their computational states as emotional states. This argument has a great deal of force, at least *prima facie*.

**6. Non-conscious emotions and homologies:**

On the basis of the argument that I have offered thus far, it would seem reasonable to infer that if an individual can be in a phenomenally unconscious emotional state, then a collectivity can as well. But, given that the states and processes of collectivities are completely incapable of being consciously experienced, it is still necessary to ask whether these states and processes belong to the same kinds as the phenomenally unconscious states and processes of an individual. Even at this point, there remain serious worries about whether similarities in computational systems are sufficient to underwrite the claim that a collectivity is capable of the same sorts of emotional states and processes as an individual. My guess is that people will be quite willing to allow that the *USS Palau* is capable of representing threats; but, there is an important respect in which its way of doing so diverges radically from the way in which individuals emotionally represent the world. This being the case, it is likely to seem illegitimate to classify these states as belonging to the same psychological kinds.

As I have argued above, the most promising defense of the possibility of non-conscious emotions relies on the presence of shared physiological structures that can be used to explain why a non-conscious emotion ought to be treated as an emotion of the same kind as a consciously experienced emotion (Damasio, 1999, 2001; deGelder, Morris, & Dolan, 2005; Ledoux, 1996; Prinz, 2004a, 2004b; Tsuchiya & Adolphs, 2007). As Antonio Damasio argues, emotional states are best understood as bodily changes (broadly construed to range from visceral changes to low-level chemical changes that modulate cognitive processing) that are integrated with working-memory representations of particular emotional stimuli. Although such changes often capture first-personal attention and recruit introspective monitoring, conscious representation of emotional states are produced in cortical areas that need not be recruited to respond emotionally. On the basis of this argument, Damasio (2001, 781) draws a conceptual distinction between emotion and feeling, arguing that emotions "provide an immediate response to certain challenges and opportunities faced by an organism, the feeling of those emotions provides it with a mental alert". By his lights, conscious and non-conscious emotions belong to the same psychological kinds because they are implemented by largely similar underlying mechanisms. Joseph LeDoux (1996, 302) advances a similar view, suggesting that "brain states and bodily responses are the fundamental facts of an emotion, and the conscious feelings are the frills that have added icing to the emotional cake." Moreover, deGelder and his colleagues (2005, 18682) adduce neurological evidence in support of the claim that "fear is mandatory and independent of awareness"; and Tsuchiya and Adolphs (2007, 8) argue that because of the organization of the neural circuitry of emotional representations, "behavior can be motivated by the affective value of stimuli that are not consciously perceived and that do not induce any conscious feelings of the emotion." However, while these data can be used to establish the existence of non-conscious emotions, they also suggest a more philosophically robust argument against the claim that individual fear and the computational states of collectivities can belong to the same natural kind (cf., Griffiths, 1997).

It is often argued that whether a state or a process belongs to a particular psychological kind depends on the presence (or lack thereof) of a cluster of underlying physiological mechanisms. Psychological states are, at least in the most obvious cases,

implemented by biological mechanisms. Thus, structural homologies between mechanisms provide scientifically viable evidence for the existence of emotional states (and psychological states more broadly) in other mammals with which we share homologous mechanisms for implementing core emotional states. Moreover, the lack of structural homologies provides scientific evidence for where human emotions diverge from their evolutionary precursors. Finally, homologous mechanisms allow us to understand different sorts of organisms as having the capacity for emotional representation without relying on an implausibly liberal understanding of emotion (Rey, 1980). Given that the same physiological structures implement conscious and non-conscious emotional representations in an individual, both sorts of states are justifiably classified as emotional states; however, because collective representations and individual representations must be implemented by radically different sorts of mechanisms, these representations should not be understood as belonging to the same psychological kind. While individuals and collectivities both possess the capacity to track dangerous stimuli, the lack of homologous mechanisms suggests that the collective representations of danger are not emotions!

There is much to speak in favor of this objection. After all, neurological models of conscious and non-conscious emotion do provide compelling evidence for thinking that there are stringent requirements on the implementation of emotional representations *in organisms like us*.[23] Moreover, the homologous structures that we find in other mammals provide compelling reasons for thinking that they too have emotional representations that are strikingly similar to those that we find in humans. In examining these homologous structures, a rich cognitive science of emotional representations extending at least through most mammals can be developed (cf., Tsuchiya and Adolphs 2007). However, although the cognitive science of emotion has focused on organisms with which we share homologous structures for emotional representation, the fact that *these* emotional representations are so implemented cannot *entail* that these are the only structures that can implement emotional states. To establish the truth of the claim that such homologies are *necessary* for emotional representation it would have to be demonstrated that the level of biological implementation is the only level that is relevant for the homuncular decomposition that is necessary for psychological kind membership. I am not convinced that this is the case.

I contend that there are compelling reasons to think that many psychological generalizations are grounded at the level of the computational mechanisms, and that they are grounded only derivatively at the level of physiological implementation. Although emotions in humans, and perhaps nonhuman animals, are reliably implemented by particular sorts of physiological structures, the success of the intuition that emotional representations *require* these physiological mechanisms turns on a failure to understand the relationship between these physiological structures and the production of an emotional representation. The assumption that explanations in psychology must be articulated as accounts of the neurological mechanisms that implement those psychological states relies on a tacit acceptance of two-levelism about psychological explanation (Lycan, 1987). According to the two-levelist the software of the mind can only be implemented in the hardware of a brain like ours. However, there are three

[23] Even people who are unaware of these debates in the cognitive science of emotion recognize that emotional states are deeply tied to the production and reuptake of neurotransmitters such as dopamine, norepinephrine and serotonin. And this holds true even for the case of non-conscious emotions. After all, we all recognize that it is easy modulate emotional states chemically with SSRIs, anti-depressants, and other more illicit substances. The prominent use of psychiatric medications to modulate non-conscious emotional processes in order to prevent untoward conscious states suggests just this fact.

philosophically significant reasons for thinking that this claim about homology of structures places an unpalatably strong constraint on membership in psychological kinds.

First, such a view of emotion makes debates about the possibility of implementing emotional states in artificial agents both philosophically and empirically uninteresting. While there may never be an artificial agent that can experience emotional states, numerous cognitive scientists are perfectly willing to entertain the possibility of artificial emotional agents (Minsky, 2006; Picard, 1997; Scheutz, 2004; Scheutz et al., 2006; Simon, 1967). It seems far too cheap of a response to this research to claim that emotion is impossible in a robot because it lacks homologous structures to implement purportedly emotional states. The possibility of implementing emotions in artificial agents raises hard empirical and philosophical questions, and merely dismissing the possibility begs the question. Second, this argument rules out the possibility of alien life forms with emotions like ours. After all, even if there were an alien life form with a biological constitution *exactly like ours*, these organisms could have emerged only by parallel evolution; so, even in the face of robust neurophysiological similarities there would be no homologies of structure. Moreover, this argument also rules out the possibility of convergent evolution—in some distant future—that could yield the capacity for emotion in a non-human species that does not share a common emotional ancestor with humans.[24] Finally, this claim about homologous structures can only succeed at the expense of eliminating emotion as a natural kind even within the human species (cf., Griffiths, 1997). Our basic fear responses, are "short lived, highly automated, triggered in the early stages of processing perceptual information, and realized in anatomically ancient brain structures that we share with many other vertebrates" (Griffiths, 2004, 236). More 'complex' variants of fear, as exemplified by the fear of failure or the fear of public speaking, are functionally analogous to our basic fear responses in that they all detect danger in some loose sense. However, Griffiths argues that such functional similarities license generalizations that are far too shallow. If psychology is in the business of uncovering the biological mechanisms of thought, then mere functional analogies will be insufficient to underwrite induction to mechanisms across superficially similar behaviors. That is, no matter how similar complex fear representations are, behaviorally speaking, to basic fear responses, because these two sorts of representations are no doubt implemented in different neurological structures they will not belong to the same natural kind for the purposes of psychology.

There are good reasons to think that homologies are incredibly important for grounding the relevant range of counterfactual stabilities for biological psychology. But, it is not clear that they are important to ground the relevant range of counterfactual stabilities from the standpoint of psychology *simpliciter*. As Jerry Fodor (1968) puts this point, the fact that we have identified a certain mousetrap with its physical structure does not commit us to thinking that all mousetraps have to be built like that—otherwise it would be impossible to build a better mousetrap. Terms such as those that we use for emotional states are functional terms. So, even a complete account of the biological structures that realize particular emotional states fails to explain all of the counterfactual regularities that are important for psychology. To capture these counterfactual stabilities,

---

[24] An anonymous referee suggests that the objection from homologous implementation can be improved by weakening it to require similarities in implementational structure that are *either* homologous or analogous. While this would allow for the existence of alien emotions, artificial emotional agents, and convergent evolution, it would still seem to block the existence of collective emotions. I agree that this is one way to attempt to rehabilitate the argument. However, at this point, we are faced with concerns about the relevant *extent* of the analogies in structure. I contend that it will be difficult to offer a non-question-begging account of the relevant similarities, and my arguments in the remainder of the paper can be applied to this revision of the argument as well.

a biological theory must track all of the behaviors that constitute a particular sort of emotional state. But, *at the biological level* these states could form a heterogeneous class—yielding the conclusion that they should not be classified as belonging to a particular kind *except* by way of psychological explanation.

Let me spell this point out by way of an analogy. The claim that homologous physiological structures are necessary for emotional representations turns on the assumption that emotions, like bourbons, can only be realized by a narrow range of lower-level structures. What distinguishes bourbon from other sorts of whiskies is the grain from which the liquor is made. In order to make bourbon, the mash would have to be between 51% and 79% corn, with the remainder being wheat, rye, and malted barley. Thus, even if something tastes like bourbon but lacks this precise composition, it is not bourbon. I, on the other hand, am inclined to treat emotions like vodka. In order for a substance to count as vodka, it must be distilled from *something* that contains natural sugars that can be consumed by yeast. Provided that the resulting substance is filtered in order to make it relatively clear, odorless, and tasteless, it counts as vodka.[25] Put briefly, bourbon requires a particular occupant of the role thing-to-be-fermented, whereas in vodka anything with natural sugar can play that role.

Vodkas diverge at least on the level of chemical structure, which depends on the substance from which a particular vodka has been distilled. However, this divergence in structure is irrelevant for most of the inductions that we want to carry out regarding vodka. The substance from which vodka has been distilled is significant only to the extent that it allows for the production of an alcoholic substance that is within the normal range for alcoholic content, taste, and color. While vodkas are artificial kinds in the most straightforward sense possible, this analogy points to a version of the familiar claim that natural kinds in one science can remain invariant across a range of lower-level properties. Lower-level differences in structure are significant only to the extent that they are capable of underwriting the relevant range of counterfactual stabilities in a higher-level science. Regardless of how the computations happen to be carried out, and regardless of the computational structure in which they are implemented, it seems reasonable to hold that an entity is in an emotional state so long as the relevant range of computations are capable of sustaining the counterfactuals that are relevant for psychological explanations. Yet, unlike disputes over the extension of 'vodka and 'bourbon', the constraints on kind membership in psychology cannot be decided merely by an appeal to convention. Barring an incredibly strong anti-realism about psychological states, the nature and scope of the boundaries around psychological kinds must be grounded in their capacity to support *psychological* generalizations. But what sorts of neurological similarities are required for psychological explanations?

Neurological similarities concerning the types and densities of cells, and even the activity of a particular region of cortical tissue, do not always replicate similarities at the psychological level. So, just assuming that cortical similarities will be sufficient to underwrite the relevant range of similarities for explaining psychological phenomena might cause us to carve up the world in a way that diverges radically from the way we carve up things in terms of psychological phenomena. To put an even finer point on this claim, it is a well-known worry about functional magnetic resonance imaging (fMRI) that neuroanatomical localization is highly variable across subjects. Merely looking at the gross morphological structures that happen to be active at a particular time is never

[25] Mikko Salmela informs me that this claim sounds scandalous to a Finn, and he claims that it would sound false to many people from Northern Europe. However, the European Union has recently upheld my view, rejecting an initiative that would limit vodka to strong alcohol made from grain and potato.

sufficient to determine what the function of that region is. Only after we have decided on a functional task (hence the f in fMRI) is it possible to understand the function of that region of neural tissue. This suggests that neurophysiological states do not cluster into the right sorts of patterns to allow for psychological generalizations on their basis alone. But, this is precisely why Griffiths (2004) is compelled to adopt an eliminativist position about emotions given the assumption that emotional structures must be homologous in order to count as belonging to the same kind.

Of course, this claim is unlikely to dissuade a proponent of the homology objection. The assumption that only biological components can implement the computations required for an emotional representation could still lead one, following Rey (1980), to argue against taking functional-cum-computational states to be sufficient for emotion. Rey insists that in order to understand what is required for the functional architecture of an emotional state, we must incorporate physiological facts about neurotransmitters into our account, otherwise we run the risk of making our ascriptions of emotion too liberal and allowing the nation of China or the Bolivian economy to count as having emotional states. Note, however, that a retreat to this move makes the assumption that physiological structures of a particular sort are necessary for emotion sound question-begging. Adopting this conceptual argument to avoid an overly liberal account of emotion suggests a thoroughly chauvinistic view of the mind. As I noted above, this account of emotion immediately rules out the possibility of an intelligent species of silicon-based life form, whose brain used electrical rather than chemical transmissions between its neurons, as having the capacity to be in emotional states. To my mind, ruling out the possibility of collective emotion *a priori* on the basis of this sort of conceptual argument seems to be a misguided project for both the cognitive scientist and the philosopher of mind. And, if this is not enough to convince you, there is an even stronger argument against this conceptual resistance to collective emotions.

Neurotransmitters themselves are occupants of the functional role of transferring information between neurological structures, primarily between spatially separated neurons, in biological organisms like us. However, the fact that something is the occupant of a functional role does not rule out characterizing the occupant of that functional role functionally. In fact, there are good reasons for treating 'neurotransmitter' itself as a *functional role*. Once we recognize this, it becomes clear that what occupies this role will be determined, at least in part, by the system in which the neurons and neurotransmitters are embedded. In a silicon-based system, the role of neurotransmitter might be occupied by electric energy. In a system consisting of a group of humans, the role might be occupied by ordinary language structures like memos, emails, and vocalizations. There is no obvious reason to assume that this role could not be filled by something besides the chemical structures that we find in biological organisms like us. So, claiming that neurotransmitters of a particular sort are required for emotion merely begs the question against extending emotions to cognitive systems that are radically different from us.

In the absence of a more compelling explanation of the reason why emotion requires a biological implementation, I conclude that a more promising view treats emotions as tools for representing the world in particular ways. Emotions, like vodkas, require only the right sort of functional organization and there are numerous ways in which the relevant sorts of functional roles can be implemented. This understanding of emotion is capable of grounding the analogy between individual emotional representations and collective emotional representations to which I appealed in the opening sections of this paper. However, one final objection waits in the wings.

## 7. A final worry about genuinely collective emotions

Unfortunately, even if emotions must be individuated functionally rather than by physiological structures, the situation aboard the *USS Palau* still seems to fall short of being a genuinely emotional state. If Hutchins' anthropological report is to be believed, none of the members of the crew exhibited anything that looked at all like fear in this case. I initially took this to be a feature of the account, noting that we should expect something similar to be the case in the individual (recall that your amygdala is not likely to be afraid even if you are). However, I am now inclined to see this as a bug. After hearing an earlier version of this paper, Marcus Hedahl called my attention to a straightforward explanation for why we should not expect the members of this crew to be afraid in this case. Indeed, he convinced me that there is good reason to suppose that this explanation would be sufficient to block the possibility of a fear representation in this case. His reasoning went roughly as follows. The high degree of organization that allows for the possibility of collective representation in a collectivity such as this results from the fact that it is a military unit that must be organized to rapidly respond to military threats. However, as members of naval crews tend to be highly trained, they are going to be able to react in the face of danger without exhibiting fear—otherwise they would not be nearly as successful in carrying out their military duties. Indeed, had the crew members exhibited fear in this case, there would have been a much greater chance of the ship running aground because the sorts of rapid computations that needed to be carried out in this case would have been impeded by such feelings of personal fear. But the sort of training that allows an individual soldier to respond without fear is also the sort of training that would wipe out the possibility of fear in the distributed cognitive system that includes these crewmembers. After all, part of the training that is involved in getting individual crewmembers to inhibit fear in the face of danger is likely to involve a number of drills on how to respond if steam pressure is lost.

It seems reasonable to suppose that the reason why everything went off without a hitch in the case described above is that the capacity to detect danger has, in this case, been decoupled from the fear of running aground. Although the crew of the *USS Palau* saw that this was a case in which a threat was posed to the ship, and were thereby able to detect danger, nothing in this case looks like a genuine collective fear. I believe that this objection shows that the state of the *USS Palau* that I have been discussing should not be classified as a genuinely collective emotion, but it should instead be seen as a collective detection of danger in the absence of a fear response. What is lacking here is the sort of agitation that we find in the case of a fear representation; and for this reason, the case of a naval vessel is unlikely to offer the right sort of case for a genuine expression of fear by a collectivity. While this case clearly exhibits enough computational organization to yield genuinely collective representations, naval crews are likely to have the fear trained out of them. In the case described above, the crew of the *USS Palau* was able to skillfully navigate a dangerous situation, but there is little reason to think that this ought to be counted as collective fear. However, there are other collectivities that are likely to possess the sort of computational architecture that is required for collective emotional representation. Unfortunately, the cognitive anthropology required to establish that this is the case has yet to be carried out. So, I will close with an intuitively plausible case where a genuinely collective emotion might emerge, noting that until the relevant anthropological data is collected, this appeal will be insufficient to conclusively establish the existence of collective emotions.

Consider the collective representations that emerged during the closing days of the United States presidential campaigns in the fall of 2008. I suggest that there is good

reason to suppose that a fear representation emerged in the distributed architecture of the McCain-Palin campaign. As the election cycle was coming to a close, it rapidly became clear that there was nothing that this campaign could do to win the election. The Republicans were focusing on the wrong states; they were taking what were widely perceived as low-blows against the Democrats; and, rather than focusing on solidifying their positions (and, by this point, it was unclear whether they had positions), each new move turned out to be little more than a direct response to the perceived behavior of the Obama-Biden campaign. In this case, I suggest that it is quite likely that a number of changes took place in the strategies for processing information within the McCain-Palin campaign. Beyond changes in overt behavioral responses to a rapidly changing situation, the members of the campaign are likely to have exhibited similar changes in processing to those aboard the naval vessel described above. The various pollsters who had played an integral role in setting the agenda throughout the campaign shifted their attention away from what was going well for McCain-Palin and began to focus instead on what was going well for Obama-Biden. The various advisers within the camp concerned themselves only with the potential threats that were posed by the Obama-Biden campaign, and the strategic suggestions that they made began to focus on ways of neutralizing the threat. This, in turn, lead the strategists to attempt to evaluate these threats, examine their salience to the success of McCain-Palin, and develop strategies for negotiating them as rapidly as possible. In short, there was a radical redeployment of attention by the flailing Republican campaign to focus on whatever the Obama-Biden campaign was doing.

The representational states and processes that were operative within this campaign had the function of carrying information that could be made available for controlling and adjusting behavior in the political arena. The integration of these representations facilitated an unsuccessful, yet still skilful attempt to cope with novel changes in the political climate of the United States. This story can be told exclusively in terms of campaign-based systems whose representations were either broadcast exogenously through public relations systems, or broadcast endogenously in the form of memos that were sent to the members of the campaign who could evaluate current strategies. However, although these representations were never accessed by the campaign *as such*, they were deeply integrated with computational systems dedicated to the production of public behavior, and these behaviors could be monitored and evaluated in a way that parallels the production of an individual's fear representation. Moreover, these component representations yielded new, highly distributed, beliefs and new global strategies for engaging in the political arena. Attention was focused on relevant dangers, and computational resources that would otherwise be focused on carrying out other tasks were diverted. This allowed the campaign to evaluate the overall threat by recruiting person-level computational systems to plan for and evaluate a range of possible responses. Forward-models could then be constructed to establish a plausible plan for dealing with the Obama-Biden campaign; having categorized the threat, and having engaged computational systems to evaluate the sort of danger that the campaign was currently facing, further computations were likely to have been initialized to mobilize the coping strategies that were deployed by the McCain-Palin campaign (while at the same time, further evaluations of the probability of harm, the capacity to cope with the threat, and the urgency of coping with the threat continued to be carried out).

This distribution of computational resources across this collectivity suggests a computational architecture that is likely to have been sufficient for producing a fear representation. The redeployment of attention, the reorientation of cognitive processes, and the production of action tendencies all seem to have played an important role in generating the behavior of the campaign. However, this representation cannot be

localized as the representational state of any member of campaign, nor can it be seen as the state of any apparatus in the campaign. More importantly, unlike the case of the *USS Palau*, there was no training to prevent the production of a fear response, and the individual members of this collectivity were clearly in a state of agitation. Yet, it is highly unlikely that the state of fear within the campaign is best described by appealing to an aggregation of individual states of fear. While it may indeed be the case that some of these people who were genuinely afraid of losing the election, the individual person-level states are likely to have been far more variable than this. Some campaign workers expressed a hopeful attitude; some strategists were angry about the portrayal of the McCain-Palin camp as foundering; and many people probably felt saddened at having spent so much time on a campaign that was now headed for failure. This being the case, it seems likely that although we find a state of agitation in this collectivity that suggests a set of states and processes that were far more emotional, the emotional state of the McCain-Palin campaign can be predicted and explained only by appeal to the computational systems governing the behavior of the collectivity. Put in other terms, there is a real pattern in the behavior of this political campaign that would be missed by someone who focused exclusively on the behaviors of the individuals that compose this collectivity, and the reason for this is that a genuinely collective emotion emerged from the coordinated activity of the members of this campaign.

At this point, I hope to have shown that the most promising criticisms of collective emotions fail. If all has gone well, I have demonstrated that there are no principled philosophical reasons to deny that emotional states can be extended to some sorts of collectivities. Although I have only addressed the case of the fear representations, I have provided a model for extending emotional states to other collectivities provided they have the right sort of organization. Although there is much work to be done in specifying the particular mechanisms at play in the production of any particular collective emotional state, this is the empirical and philosophical work that can ground a robust research program directed at understanding genuinely collective emotions. As is often the case, commonsense psychology, as well as the received view in philosophy and cognitive science, is probably right *to some degree*. The United States probably does not regret its failure to intervene in the Rwandan genocide; and, the Teamsters were probably not angry about the recent decision to open American borders to Mexican trucking companies. However, the reason *why* these groups were not in emotional states is far more interesting than the brute intuition that individuals can be in emotional states but collectivities cannot. In each of these cases, the reason is that these groups are not organized in the right way to emotionally represent anything at all. However, rather than focusing on the enormous number of cases where collectivities do not represent, we will be better served by focusing on the cases where collectivities are organized in a way that allows for emotional representation. After all, it is typically true that focusing on patterns of failure is likely to yield less interesting results than focusing on patterns of success.

**8. Works Cited:**


Arico, Adam. 2007. Should corporate consciousness be regulated? (Poster), *Society for Philosophy and Psychology,* Toronto, ON.

Arico, Adam. Forthcoming. Folk-psychology, consciousness, and context effects. *European review of philosophy*.

Baber, Chris, Paul Smith, James Cross, John E. Hunter and Richard McMaster. 2006. Crime scene investigation as distributed cognition. *Pragmatics and communication, 14* (2), 357-385.

Bennett, Maxwell and Peter Hacker. 2003. *Philosophical foundations of neuroscience*. Oxford: Blackwell.

Block, Ned. 2007. Consciousness, accessibility and the mesh between psychology and neuroscience. *Behavioral and brain sciences, 30*, 481-548.

Bratman, Michael. 1993. Shared intention, Ethics, 104: 97-113.

Brooks, David. 1986. Group minds. *Australasian journal of philosophy, 64*, 456-470.

Chalmers, David. 1996. *The conscious mind* Oxford: Oxford University Press.

Clark, Andy. 1997. *Being there*. Cambridge: MIT Press

Clore, Gerald. 1994. Why emotions are never unconscious. In P. Ekman & R. Davidson (Eds.), *The nature of emotion: Fundamental questions*. Oxford: Oxford University Press.

Damasio, Antonio. 1999. *The feeling of what happens*. New York: Harvest Press.

Damasio, Antonio. 2001. Fundamental feelings. *Nature, 413*, 781.

Darwin, Charles. 1872/1965. *The expression of the emotions in man and animals*. Chicago: Chicago University Press.

deGelder, Beatrice, John Morris, and Raymond Dolan. 2005. Unconscious fear influences emotional awareness of faces and voices. *PNAS, 102*, 18682-18687.

Dennett, Daniel. 1978. *Brainstorms*. Montgomery, VT: Bradford Books.

Dennett, Daniel. 1989. *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, Daniel. 1991. *Consciousness explained*. New York: Penguin Press.

Dretske, Fred. 1981. *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Dretske, Fred. 1988. *Explaining behavior*. Cambridge, MA: MIT Press.

Ekman, Paul. 2003. *Emotion revealed*. New York: Times Books.

Elster, Jon. 1999. *Alchemies of the Mind: Rationality and the Emotions*. Cambridge: Cambridge University Press.

Fitch, W. Tecumseh. 2008. Nano-Intentionality, Biology and Philosophy, 23: 157-177.

Fodor, Jerry. 1968. *Psychological explanation*. New York: Random House.

Fodor, Jerry. 1980. Methodological solipsism considered as a research strategy in cognitive science. *Behavioral and brain sciences, 3*, 63-73.

Freud, Sigmund. 1915/1950. The unconscious (J. Riviere, Trans.). In *Collected Papers* (Vol. 4). London: Hogarth Press.

Frijda, Nico. 1986. *The emotions*. Cambridge: Cambridge University Press.

Galison, Peter. 1997. *Image and Logic*. Chicago: University of Chicago Press.

Giere, Ronald. 2002. Scientific Cognition as Distributed Cognition. In P. Caruthers, S. Stich & M. Siegal (Eds.), *The cognitive basis of science*. Cambridge: Cambridge University Press.

Gilbert, Margaret. 1989. On Social Facts. New York: Routledge.

Gilbert, Margaret. 2001. Collective remorse. In Jokic, A. (Ed.), War crimes and collective wrongdoing: A reader. Oxford: Blackwell Publishing.

Gilbert, Margaret. 2002. Collective guilt and collective guilt feelings. *The Journal of Ethics,* 6: 115–143.
Griffiths, Paul. 1997. *What emotions really are*. Chicago: University of Chicago Press.
Griffiths, Paul. 2004. Is emotion a natural kind. In Robert Solomon (Ed.), *Philosophers on emotion*. Oxford: Oxford University Press.
Griffiths, Paul and Andrea Scarantino. 2005. Emotions in the wild. In P. Robbins, and M. Aydede (Ed.), *The Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.
Haugeland, John. *Having thought*. Cambridge: Harvard University Press.
Huebner, B. (2008). Do you see what we see? An investigation of an argument against collective representation. *Philosophical psychology,* 21 (1): 91-112.
Huebner, B. (2010). Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and the cognitive sciences*.
Huebner, B. (forthcoming). Minimal Minds. *Oxford Handbook on Animals*. New York: Oxford University Press.
Huebner, B. (in prep). Distributed minds and collective mentality.
Huebner, B. & D.C. Dennett (2009). Banishing "I" and "we" from accounts of metacognition (Open peer commentary on P. Carruthers), *Behavioral and Brain Sciences,* 32: 148-149.
Huebner, B., H. Sarkissian, & M. Bruno (2010). What does the Nation of China think about phenomenal states? *Review of philosophy and psychology* 1:.
Hutchins, Edwin. 1995a. *Cognition in the wild*. Cambridge: MIT Press.
Hutchins, Edwin. 1995b. How a cockpit remembers its speeds. *Cognitive science, 19*, 265-288.
James, William. 1884. What is an emotion? *Mind, 9*, 188-205.
Jaspers, Karl. 1947. *The Question of German Guilt*, Ashton, E. (trans), New York: Capricorn Books.
Kenny, Anthony. 1963. *Action, Emotion, and Will*. London: Routledge and Kegan Paul.
Knobe, Joshua and Jesse Prinz. 2007. Intuitions about consciousness. *Phenomenology and cognitive science, 7*, 67-85.
Knorr Cetina, Karen. 1999. *Epistemic Cultures: How the sciences make knowledge*. Cambridge MA: Harvard University Press.
Kraut, Robert. 1986. Feelings in context. *The Journal of philosophy, 83* (11).
Kutz, Christopher. 2001. *Complicity*. Cambridge: Cambridge University Press.
Lane, Richard. 2000. Neural correlates of conscious emotional experience. In Richard Lane and Lynn Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 345-370). Oxford: Oxford University Press.
Ledoux, Joseph. 1996. *The emotional brain*. New York: Simon and Schuster.
Lycan, William. 1987. *Consciousness*. Cambridge: Bradford Books.
May, Larry. 1991. Metaphysical guilt and moral taint. In May, L. & Hoffman, S. (eds). *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics* (pp.239-354). Savage, MD: Rowman and Littlefield.
May, Larry and Stacey Hoffman, eds. 1991. *Collective Responsibility: Five Decades of Debate in Theoretical and Applied Ethics*. Savage, MD: Rowman and Littlefield.
Millikan, Ruth. 1984. *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.
Minsky, Marvin. 2006. *The emotion machine*. New York: Simon & Schuster.
Nash, Ronald. 1989. Cognitive theories of emotion. *Nous, 23*, 481-504.

Panksepp, Jaak. 2000. Emotions as natural kinds within the mammalian brain. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 137-156). New York: Guilford Press.
Pettit, Phillip. 1996. *The common mind*. Oxford: Oxford University Press
Pettit, Phillip. 2003. "Groups with Minds of Their Own," in Fred Schmitt (ed.), *Socializing Metaphysics*, Oxford: Rowman and Littlefield.
Picard, Rosalind. 1997. *Affective Computing*. Cambridge MA: MIT Press.
Pitcher, George. 1965. Emotion. *Mind, 295*, 326-246.
Prinz, Jesse. 2002. *Furnishing the mind*. Cambridge, Ma: MIT Press.
Prinz, Jesse. 2004a. Embodied emotions. In Robert Solomon (Ed.), *Thinking about feeling*. Oxford: Oxford University Press.
Prinz, Jesse. 2004b. *Gut reactions*. Oxford: Oxford University Press.
Rey, Georges. 1980. Functionalism and the emotions. In A. Rorty (Ed.), *Explaining emotions*. Berkeley: University of California Press.
Rupert, Robert. 2005. Minding one's cognitive systems. *Episteme: A Journal of Social Epistemology*, 1, 177–188.
Schachter, Stanley and Jerome Singer. (1962). Cognitive, social and physiological determinants of emotional states. *Psychological Review, 69* (379-399).
Scherer, Claus. 1987. Toward a dynamic theory of emotion: . *Geneva Studies in Emotion and Communication, 1*(1), 1-98.
Scheutz, Mathias. 2004. *Useful Roles of Emotions in Artificial Agents: A Case Study from Artificial Life* Paper presented at the American Association for Artificial Intelligence.
Matthias Scheutz, James Kramer, Chris Middendorff, Paul Schermerhorn, Michael Heilman, Dave Anderson, and Peter Bui 2006. *Toward Affective Cognitive Robots for Human-Robot Interaction* Paper presented at the American Association for Artificial Intelligence.
Searle, John. 1995. The Construction of Social Reality. New York: Free Press.
Shweder, Richard. 1991. *Thinking through cultures*. Cambridge MA: Harvard University Press.
Simon, Herbert. 1967. Motivational and emotional controls of cognition. *Psychological review, 74*(1), 29-39.
Sizer, Laura. 2006. What feelings can't do. *Mind and Language, 20*(1), 108-135.
Sutton, John. 2006. Distributed cognition: Domains and dimensions. *Pragmatics and cognition, 14*, 235-247.
Sytsma, Justin and Edouard Marchery. 2009. How to study folk intuitions about phenomenal consciousness. *Philosophical Psychology*, 22, 21-35.
Tollefsen, Deborah. 2006. The rationality of collective guilt. Midwest studies in philosophy, 30, 222-239.
Tsuchiya, Naotsu and Ralph Adolphs. (2007). Emotion and consciousness. *TICS, 11*(4), 158-167.
Velleman, David. 1997. How to share an intention. *Philosophy and phenomenological research*, 57: 29-50.
Whalen, Paul, Jerome Kagan, Robert Cook, F. Caroline Davis, Hackjin Kim, Sara Polis, Donald McLaren, Leah Somerville, Ashly McLean, Jeffrey Maxwell, and Tom Johnstone. 2004. Human amygdala responsivity to masked fearful eye whites. *Science, 306*(5704), 2061.
Whalen, Paul, Scott Rauch, Nancy Etcoff, Sean McInerney, Michael Lee, and Michael Jenike. 1998. Masked presentations of emotional facial expressions modulate

amygdala activity without explicit knowledge. *The Journal of neuroscience, 18*(1), 411-418.
Wilson, Robert. 1995. *Cartesian psychology and physical minds*. Cambridge: Cambridge University Press.
Wilson, Robert. 2001. Group-level cognition, Philosophy of Science, 68 (supp): S262-S273.
Wilson, Robert. 2004. Boundaries of the Mind: The individual in the fragile sciences: Cognition. Cambridge: Cambridge University Press.
Wilson, Robert. 2005. Collective memory, group minds, and the extended mind thesis, Cognitive Processing, 6: 227-236