

Psychopathy, autism, and basic moral emotions: Evidence for sentimentalist constructivism

please cite published version of this article in *The Bloomsbury Companion to the Philosophy of Psychiatry*

Abstract:

Philosophers and psychologists often claim that moral agency is connected with the ability to feel, understand, and deploy moral emotions. In this chapter, I investigate the nature of these emotions and their connection with moral agency. First, I examine the degree to which these emotional capacities are innate and/or ‘basic’ in a philosophically important sense. I examine three senses in which an emotion might be basic: developmental, compositional, and phylogenetic. After considering the evidence for basic emotion, I conclude that emotions are not basic in a philosophically important sense. Emotions, I argue, are best understood as socially constructed concepts. I then investigate whether these emotions are necessary for moral agency. In order to do this I examine the philosophical and psychological literature on psychopathy and autism (two conditions defined in terms of empathic and emotional deficits). Persons with psychopathy appear incapable of distinguishing moral from non-moral norms. Additionally, while persons with autism often struggle to develop their empathic capacities, they are capable of understanding and deploying moral emotions like guilt and shame. I conclude that, in line with the conceptual act theories of emotion, that only contagion-based empathy is necessary for the acquisition of moral concepts.

*Erick Jose Ramirez
Santa Clara University*

Few human experiences are as universal as emotional experience. Or so it is thought. There are, to be sure, cross-cultural differences in emotional *concepts*. For example, the emotional life of the Ifaluk people, who all live on a few small islands within the Federated States of Micronesia, is dominated by four emotions: song, fago, ker, and metagu/rus. Although these emotional concepts are often translated as justified indignation, compassion-love-sadness, happiness-ecstasy-joy, and fear-panic-surprise-anxiety respectively, many wonder whether these translations do justice to the Ifaluk emotional life (Cardoso 2015). Although few scholars dispute the existence of cultural variation across emotional concepts, many believe that this variation masks an underlying unity of emotional *experience*.

In this chapter I provide a critical examination of what are called “basic” emotions.

Philosophers and psychologists have long believed that human beings, with few exceptions, are born with the capability of feeling and expressing a small but important set of emotions. Although happiness, sadness, disgust, surprise, fear, and anger are often the prime suspects on a list of basic emotions, the exact names and number of these emotions varies widely from as few as two to as many as a eighteen (Ortony & Turner 1990).

I begin by explaining the allure that basic emotions have held for psychologists and, especially, for philosophers. Sentimentalism, the view that moral values are at least partially constituted by emotional content, lies at the heart of this allure. I will survey which, among several, senses of “basic” philosophers have traditionally been interested in when they search for basic emotions before I shift my attention to the empirical evidence for and against the existence of these emotions. I will conclude that the best available evidence actually points us away from basic emotions and toward a conceptual act model of emotional concepts. If our interest in emotion is grounded in our interest to understand moral psychology, I argue that we can learn a lot about the role of emotion in moral psychology by examining the moral concepts of psychopaths and high functioning persons with autism. Both psychopathy and autism are typically characterized in terms of empathic deficits. However, the empathic deficits associated with each condition are quite distinct. While individuals diagnosed with psychopathy routinely struggle to acquire moral concepts, high functioning persons with autism do not. I conclude by suggesting the psychopath’s lack of moral concepts is best understood as a result of their lack of empathic contagion and is exactly what a conceptual act model of emotion would predict.

¹ Sentimentalist moral psychology, I claim, is enriched by incorporating these findings into their

¹ Throughout this essay, I make reference to “psychopaths.” Human beings are neurodiverse, meaning that human beings experience and cognize the world and human social concepts in diverse ways. In using the term psychopathy, I intend to describe a class of persons diagnosed with a very specific set of character traits (elaborated in section V of this chapter). I do not intend to license any normative judgements about the status, standing, or value of such persons.

normative frameworks.

I. The Allure of Basic Emotions

Aside from those with purely theoretical interests, why might we care about whether our emotions are basic (in some sense) or not? One answer comes to us courtesy of early modern philosopher David Hume. Hume argued that our emotions, not our reasons, are responsible for our moral and aesthetic values (Hume 1740/2002). When we judge that something is beautiful or ugly, right or wrong, just or unjust, what we are doing, Hume claimed, is projecting our emotional states onto the external world. To better understand the evaluative world, Hume argued, we must better understand our emotional lives. Hume is not alone in sensing this connection between how we *feel* about something and our value *judgments* about it. Sentimentalist philosophers believe that this observation represents a deep truth about the nature of morality.

Basic emotions are especially interesting to philosophers trying to make sense of the link between emotion and value. Sentimentalist philosophers turn to basic emotions to explain two seemingly essential features of morality. First, it seems true for most of us that moral judgments are not passive experiences. To judge that something is wrong is to feel some kind of motivation to *do something* about our judgments. If moral judgments were purely about what we *think*, this feature of our moral judgments would be harder to explain. Picture someone who believes that humans are causally responsible for global climate change. Assume further that this person does not think that that it's *wrong* for humans to cause global climate change. Would we be surprised if such a person fails to do anything about climate change? Probably not. Now picture someone who believes that anthropogenic climate change represents an injustice both to future generations and to the globe's

poorest people. It would seem odd if someone who sincerely judged it wrong to contribute to climate change did nothing or, worse, increased her carbon footprint. Sincere moral judgments, in other words, seem to be at least weakly motivating. We can make sense of this fact, assuming it is a fact, if our sincere moral judgments were emotionally laden. Emotions are intrinsically motivating in the sense that feeling an emotion brings with it an impulse to act.

Second, emotions appear to lie at the bottom of many moral and aesthetic disagreements. In a series of studies, psychologist Jonathan Haidt asked subjects whether certain actions were wrong. These included instances of harmless but clearly aberrant behavior. For example, Haidt included cases where subjects used an American flag to clean a toilet or in which a family decides to eat their recently deceased dog instead of burying it (Haidt, Helena Koller, & Dias 1993). In another study, Haidt asked his subject to determine whether or not it would be wrong for two siblings to engage in a consensual incestual relationship (Haidt 2001). As a result of a phenomenon now referred to as “moral dumbfounding,” Haidt concluded that his subjects grounded their judgments about the moral permissibility of an act based not on good reasons but instead on how those acts made them *feel*. For example, Haidt found that his subjects continued to say that consensual incest was wrong even when he concocted cases in which all of the reasons they gave to justify their judgments no longer applied.² “Moral dumbfounding” remains a controversial thesis about evaluative judgment (Jacobson 2013). However, sentimentalist philosophers see it as powerful evidence in favor of their position. To these philosophers, emotions appear to play a fundamental role in our moral judgments and moral practices.

If our moral judgments are, in fact, grounded on emotion, then the allure of basic emotions is

² The most common reasons given against incest included the risk of birth defects and a concern over possible abuse. In one version of Haidt’s incest case, for example, the siblings met only as adults, did not have a history of abusive relationships, highly enjoyed the experience, and agreed to be sterilized before sexual contact. In such cases, subjects continued to judge that the act was wrong. When pressed for a reason, many claimed that incest was simply disgusting.

easier to understand. Basic emotions offer a solution to a problem that many philosophers believe must be explained by any adequate moral theory: the problem of disagreement. It is a truism to claim that people disagree about values. It is harder to explain what these disagreements say about the nature of morality.

Disagreements, for example, occur both within a culture and between cultures. Whether abortion ought to be permissible, for example, continues to be the subject of intense disagreement within and between cultures. Although many disagreements might be explained by the fact that one, or both, parties to the disagreement have imperfect or incorrect information, it seems possible that equally well-informed persons can disagree about what goodness consists in or whether a particular action is morally justifiable. Although many philosophers confront moral disagreement by embracing moral relativism (the view that moral claims are limited in their scope or application), sentimentalists hoping for a more objective response to the problem of moral disagreement have turned to basic emotions for help.

The idea is that a shared response, or sentiment, somehow moors us in a common subject matter, making it possible for us to disagree substantively about what a thing has to be like in order to be such that we should feel this sentiment toward it. Thus, if the sentimentalist is right, it is because our evaluative concepts have a special tie to shared human sentiments that we are able to engage meaningfully in debates over their application. (D'Arms 2005, 13)

If evaluative judgments are anchored by our emotional responses and at least some of those emotional responses are basic then moral disagreements can be understood as disagreements about the conditions under which a particular emotion should be felt. If these emotions are *basic* then there also may be a mind-independent answer available to help settle such disputes. In other words, basic emotions offer the prospect of securing a universal domain of cross-cultural human values. In order to see why such

objective standards exist, I need to say more about what senses of “basic” philosophers and psychologists have in mind when they speak of “basic emotions.”

II. What’s basic about basic emotions?

What are we saying when we say that an emotion is basic? In this section I briefly survey three possible answers to this question. Emotions might be basic in a developmental, compositional, or phylogenetic sense (Ortony & Turner 1990). Importantly, an emotion may be basic in all three senses. They are not mutually exclusive. Although it is a live question whether emotions are basic in any sense, in this section I argue that phylogenetically basic emotions hold the greatest promise for sentimental moral philosophers.

To say that an emotion is developmentally basic is to say that the emotion appears early in the course of human development. That some emotions may be developmentally basic is not particularly debated among philosophers or psychologists. Prime candidates for developmentally basic “emotions” include positive and negative feeling states. These states define the valence of all experience. Although positive and negative feelings are an important part of lived experience, such states are too “thin” to serve the philosophically complex roles that moral emotions do.³ They cannot ground moral disagreement and lack the content to explain the differences between different kinds of moral assessments.

To say that an emotion is compositionally basic is to say that these emotions work like atoms with respect to molecules. Compositionally basic emotions can be used to build other emotions and do not contain any other emotions as parts.

³ To call an emotion “thin” is to say that it lacks evaluative content. To say that a state feels good (or bad) is not to say all that much about it. Contrast that with saying that one is disgusted by something. Disgust is “thicker” in that it not only tells us something about the *valence* of the feeling, it *condemns* its object.

The main criterion used by those who advocate basic emotions in the [compositional] sense focuses on the interrelationship of the emotions, rather than directly on the nature of the eliciting conditions. Here, an emotion is regarded as basic if it contains no other emotion as a component. (Ortony & Turner 1990, 324)

Schadenfreude, for example, is the name given to a type of pleasure derived from the suffering of others. Schadenfreude is unlikely to be compositionally basic. It decomposes into several more primitive states: empathic pain derived from the suffering of others and pleasure caused by that experience. Many emotions might be thought to be compositionally basic though this form of basicness also will not be sufficient to serve the philosophically useful functions that philosophers appeal to basic emotions for. Compositionally basic emotions may vary widely across cultures and hence fail to account for disagreements.

Lastly we can call an emotion basic in a phylogenetic sense. To call an emotion phylogenetically basic is to say something about the *causal* history of that emotion. A phylogenetically basic emotion arises as a product of natural selection. Psychologist Paul Ekman writes that

[t]here are two key issues, which I use the adjective *basic* to convey about the position I have adopted... (1) There are a number of separate emotions which differ one from another in important ways. (2) Evolution played an important role in shaping both the unique and the common features which these emotions display as well as their current function. (Ekman 1994, 170)

Prime candidates for phylogenetically basic emotions include anger, disgust, fear, surprise, happiness, and sadness. A phylogenetically basic emotion is also likely compositionally basic in the sense that human cultures can expand upon their evolutionarily gifted repertoire of emotions to form more complex blends (transforming guilt into survivor's-guilt, for example). That an emotion is phylogenetically basic therefore makes it a prime candidate for solving the problems of moral

disagreement and for making sense of Haidt's cross-cultural moral dumbfounding research.

This is because if phylogenetically basic emotions exist then all neurotypical humans will share these emotions. If all neurotypical humans share these emotions then they can serve as common points of contact for moral conversations about the function and purpose of these emotions. Additionally, because they are evolutionary products, there will be an answer to questions about what the appropriate objects are for these emotions. This would allow empirically-inclined philosophers to make progress on addressing difficult moral questions that are grounded on those emotions. For example, if shame evolved to serve a specific purpose then there may be a way to answer questions about whether it is appropriate to be ashamed of our appearance, class standing, or gender identities. It is little wonder that sentimentalist philosophers have tended to focus their attention on phylogenetically basic emotions (D'Arms & Jacobson 2000). In what follows, I use the term "basic emotion" to refer to phylogenetically basic emotions.

III. Evidence for basic emotions

What evidence do we have that emotions are basic? Charles Darwin theorized that animals (including human animals) behaved in relatively stereotyped ways when placed in specific sorts of circumstances (Darwin 1873/2009). Dogs will often adopt threatening or submissive body postures, for example, in order to establish a pack hierarchy. Humans will often flush and hide their faces when performing in socially costly ways. Emotions, he reasoned may have evolved as strategic solutions to universal life problems and his observations of stereotyped, seemingly automatic, behavior was evidence of this. In the wake of Darwin's arguments, behavioral evidence dominated in the search for basic emotions.

In this section, I survey this evidence. I focus predominantly on facial expression research that

has been emerging out of Paul Ekman's research group for over four decades. I argue that, while Ekman's research suggests that emotion and facial expression are connected, facial expressions do not give us evidence for basic emotion. Some have also looked to the body and the brain as a source of evidence for basic emotions. Although many have claimed to find evidence for emotionally distinct physiological or neurological profiles, I argue that meta-analyses of these studies provide, at best, evidence for a much weaker thesis than the existence of basic emotion. Evidence about a person's face, body, and brain can give us useful information about the valence of their experience and about its intensity but no combination of these three pieces of information will tell us which particular emotion someone is experiencing. We will need to turn to other models of emotion to explain this result.

In 1971, Paul Ekman and Wallace Friesen published a study on facial expressions that would create a new paradigm in emotion research (Ekman & Friesen 1971). Ekman and Friesen believed that facial expressions represented specific emotional states. In order to test this claim, Ekman and Friesen needed to determine whether people the world over agreed about what facial expressions meant.

Ekman and Friesen chose the Fore people of Papua New Guinea as subjects. The Fore were chosen because they lived in small groups that were relatively isolated from Western cultural influences. Ekman and Friesen showed pictures of American subjects to members of the Fore. Given their isolation, the Fore were seen as excellent subjects to test Darwin's supposition about the link between behavior, especially facial expression, and emotion. Each picture was a closely cropped image of a person's face which Ekman and Friesen had taken great pains to pose. Each Fore was presented with one of three emotion words (translated from English into the Fore's native language) and then asked which emotion the face in the image represented. If facial expressions gained their meaning a result of socio-historical facts then the Fore, who lacked knowledge of Western culture, should have

responded somewhat randomly when asked to match facial expressions to emotion words. This is not what Ekman and Friesen observed. The Fore converged on the same sorts of answers that their Western counterparts converged on. Smiling faces were connected to happiness, frowning faces with anger, and so on.

Other researchers would quickly adopt Ekman and Friesen's methodological approach and apply it across the globe (Ekman 1972b; Matsumoto 1992). The seemingly stunning convergence of the world's peoples to questions about the meaning of facial expressions was interpreted as a powerful vindication of Darwin's suggestion that emotions were basic. Ekman argued that basic emotions were what he called "affect programs." Affect programs are mental modules. As such, they are non-conscious evolved neural circuits which produce stereotyped responses. Affect programs are triggered by specific emotionally relevant input and produce emotionally distinct outputs (Ekman 1972a). Each emotion, on this view, is distinguishable from others in virtue of the inputs and outputs associated with the it. Fear, for example, can be understood as a response to the perception of something as an imminent threat. When triggered, the fear module activates and prepares a creature to fight or flee the source of the threat. These preparations include emotion specific changes to a creature's physiology (increased heart rate, blood pressure, and so on) and, in humans and other apes, includes the production of an emotion-specific facial expression. Although the facial expression research paradigm has been successfully replicated by researchers for decades, we should be wary of the conclusion that Ekman and others have drawn from it. In particular, I suggest that we jump too hastily from the convergence of facial expression judgment studies to claims about affect programs and basic emotion (Russell 1995).

The affect program approach to emotion is grounded on several assumptions. First, the claim

that emotions evolved to produce specific facial expressions. Second, that humans evolved to detect these signals in order to better understand the emotional lives of those around us. Third, that these signals are difficult to convincingly fake. All three assumptions have been challenged in the last twenty five years.

Facial expression judgment studies have been subject to many lines of criticism. One line of criticism takes issue with Ekman's methodology, which most facial expression judgment studies adopted. Ekman's paradigm utilized what is known as a within-subjects forced-choice design. Although, by itself, the use of any particular study design is neither helpful nor hurtful, methodological decisions can have an impact on we analyze the results of each study (Russell 2016). A "forced-choice" experiment is one in which subjects must select a response from those provided by the research. In the case of facial expression studies, critics point out that the forced-choice response format artificially raises the degree of convergence. The problem with using forced choice within subject experiments was perhaps best demonstrated in a series of studies in which, depending on the options available, subjects could be made to converge on different answers, even non-sense answers, when asked what a specific face expresses (DiGirolamo & Russell 2016).

To say that an experiment is "within-subject" is to say that each subject of the experiment is exposed to the entire stimulus set (Russell 1995). In the context of facial expression studies, it means each subject is shown all of the facial expression photographs used in the experiment and asked to make judgments about them. Why might this pose a problem for judgment studies? Repeated exposure to several faces cues subjects to pay closer attention to the differences between faces and therefore impacts their judgments about the meaning of each face. This feature serves to artificially raise rates of convergence. Facial expression judgment studies that adopt "free-response" formats

(where subjects are asked to write in what they believe a face expresses) or “between-subject” designs (where subjects are only shown a single facial expression) show significantly lower rates of convergence than Ekman’s more traditional protocol. This is especially true for studies using members of non-Western cultures as subjects (Gendron, Roberson, van der Vyver, & Feldman Barrett 2014).

Critics argue that these critical experiments show us that it is the study *design* utilized by most facial expression researchers that produces such high rates of convergence on answers. While these methodological criticisms do not rule out the existence of basic emotions, they weaken the kind of evidence that facial expression studies provide for them.

A second set of criticisms levied against the facial expression paradigm targeted the assumption that basic emotions evolved to produce specific facial expressions as signals. In a study, conducted in the 1920s, psychologist Carney Landis tried to trigger strong emotional states in his subjects. Landis did this in various ways including asking subjects to place their hands into a bucket of frogs, smell ammonia, and view at pornographic images; at the end of the experiment, Landis asked his subjects to decapitate a living rat (Landis 1924). Throughout the experiment, Landis took photographs of his subject’s facial expressions. To his surprise, he found that subjects rarely produced the stereotyped facial expressions commonly associated with basic emotions. Although Landis’ approach has methodological (and ethical) issues of its own, Landis’ data would prompt other researchers to test the assumption that basic emotions evolved to produce specific facial expressions.

Jose Fernández-Dols and his colleagues would attempt to test these assumptions in the late 1990s. They were interested in the kinds of facial expressions people spontaneously made while undergoing strong emotional episodes. For example, they observed gold medal winners when they won their competitions, as they waited during the award ceremony, and during the moment they had

their medals placed around their necks (Fernández-Dols & Ruiz-Belda 1997). They also studied facial expressions generated spontaneously in a number of different public spaces including those of soccer fans during a match (Fernández-Dols & Crivelli 2013). If facial expressions evolved as non-cognitive (unconscious) signals of basic emotions, then soccer fans and gold medal winners seem like the perfect test subjects.

Like Landis, Fernández-Dols' results did not conform to the expectations of the facial expression paradigm. He discovered that, while subjects reported strong feelings throughout their gold medal experience, for example, that they only produced the expected facial expression when they were under public scrutiny. As the medal was placed on their necks, gold medal winners smiled. Smiles were missing, however, during other elements of their experience. Fernández-Dols and others argued that facial expressions were more likely to be the product of social signalling than of automatic affect programs. Attempts to discover the strength of the link between felt emotion and facial expression has consistently demonstrated that link to be weaker than predicted by affect program theorists (Reisenzein, Studtmann, & Horstmann 2013). Coupled with the methodological criticisms of the affect program paradigm, emotion researchers began to look for other avenues in support of basic emotions.

Evidence from facial expression judgment studies, while initially seen as strongly supporting basic emotions, have received a critical second-look in the late 20th and early 21st centuries. Basis emotion theorists have, in the meantime, tried to look for other types of evidence to support their views. In particular, these researchers have held out hope that our bodies and brains will provide them with tell tale clues of basic emotion. If discrete emotions evolved to solve specific life conflicts, these emotions might be identifiable by the discrete patterns of physiological or neurological activation they

trigger. An explosion of studies emerged in the late 1990s and 2000s all claiming to have discovered emotion-specific markers for anxiety, anger, sadness, disgust, happiness, and the like. However, when subjected to meta-analysis, the strength and uniqueness of these emotion-specific markers falls dramatically (Phan, Wager, Taylor, & Liberzona 2002; Wager, Kang, Johnson, Nichols, Satpute, & Feldman Barrett 2015). At the time of writing, intense disagreement exists about the robustness of emotion-specific physiological or neurological markers.

Sentimentalist philosophers appealed to basic emotions in the hopes of addressing longstanding questions about the nature of moral judgments and moral disagreements. Basic emotions seemed poised to explain how moral disagreements, even radical ones, could be understood without appeal to moral relativism. However, in light of what has become a growing consensus away from basic emotions among psychologists, neuroscientists, and philosophers of mind, sentimentalists must look elsewhere for means of addressing these concerns. The picture of emotions emerging leans away from emotions as basic and toward a view of emotions as culturally informed perceptions.

IV. Beyond Basic Emotions

Although we have reason to doubt the existence of basic emotions, sentimentalist philosophers can salvage a great deal of useful material from the theories of emotion emerging in response to critiques of basic emotion. In addition, recent discoveries about the varying capacities all currently referred to as forms of empathy, especially as they appear in psychopaths and high functioning people with autism, not only give further evidence for the sentimentalist approach but can functionally replace basic emotions. That is, constructivist accounts of emotion can fulfill most of the roles that basic emotions were once thought uniquely suited perform. In this section I briefly survey contemporary

psychological constructivist and conceptual act theories of emotion. In the following, and final, section, I explain how psychopathic and autistic agency can point the way toward a more satisfactory moral psychology.

If emotions are not basic, then what are they? How do we explain the seeming universality of emotional experiences and emotional facial expressions? Psychological constructivist approaches to emotion address these questions by distinguishing between “affect” and “emotion.” Affect, on this view is a phylogenetically basic aspect of all experience including non-emotional experiences. Affect defines the “feel” or “phenomenology” of experience and is defined in terms of valence (positive/negative) and arousal (high/low) (Feldman Barrett 2006). Hunger, for example, might include a highly arousing but negative affective component. Although emotions include affect as a component, emotions are meaningful groupings of causes, affective changes, thoughts, and behaviors that have been given their meaning as a result of cultural construction.

Like the basic emotion approach, psychological constructionist models take evolution seriously and consider how biological and species-general factors have some role in the events that are conventionally named as “sadness”, “anger”, and “fear”. Like the appraisal approach, the psychological construction approach to emotion involves making meaning out of sensory cues. And, like social constructionist models of emotion, psychological construction approaches consider the [culturally] specific elements of emotion. As socially constituted artifacts of learning and culture, no emotion category is assumed to be biologically basic. (Barrett 2009, 1290)

Emotional concepts have top-down effects on affective experience. They structure how we understand the causes of affective change and the meaning of the change itself. Fago, song, and the like are a part of the emotional and social world for the Ifaluk. Emotional concepts structure and regulate how a person sees the world. In much the same way that we do not have to consciously make efforts to parse out our perceptions *into* cars, airplanes, computers, and the like, our emotional concepts structure our

affective experiences relatively automatically. Because emotional concepts must be learned (they are language and culture dependent), these approaches to emotion are comfortable abandoning the basic emotion paradigm.

According to the conceptual act model, emotion categories exist because groups of people agreed (for phenomenological and social reasons) that this is a functional way to parse the on-going mental activity that is realised in the brain. The model is consistent with the observation that some of the categories are cross-culturally stable (because they function to address certain universal human concerns that stem from living in large, complex groups), whereas other categories are culture specific. The conceptual act model hypothesises that the category instances named with emotion words are real, but they derive their reality from the human mind (in the context of other human minds). (Barrett 2009, 1293)

Conceptual act theories allow for a great deal of variation in terms of how cultures parse out affective space into emotional and non-emotional experiences. For example, “[m]any depressed Chinese people do not report feeling sad, but rather express boredom, discomfort, feelings of inner pressure, and symptoms of pain, dizziness, and fatigue. These culturally coded symptoms may confound diagnosis among Chinese immigrants in the United States, many of whom find the diagnosis of depression morally unacceptable and experientially meaningless” (Kleinman 2004). Such findings are consistent with the conceptual act theory.

I pause here to note that, on the conceptual act model of emotion, emotions are cultural artifacts. Sentimentalists interested in grounding moral values in emotion will therefore be driven to adopt a form of cultural relativism. Moralized emotional concepts will only make sense within those groups where such an emotion is available to structure affective experience. However, I argue that this acceptance has two upsides.

First, at least some emotional concepts are likely to be pancultural in virtue of being useful for structuring social relations and managing the distribution of scarce resources. Shame may be such an

emotion (Wong & Tsai 2007). Insofar as moral regulative emotions like shame are cross-culturally shared, their lack of basicness does not pose a problem for cross-cultural moral communication. Second, conceptual act models make predictions about how individuals acquire and use emotionally laden moral concepts. In the next section, I argue that investigations of the moral concepts of psychopathic and high functioning persons with autism largely support these predictions and provide further insight into the nature of our moral capacities.

V. Psychopathy, Autism, and Sentimentalist Moral Psychology

Conceptual act theories of emotion include a basic element, affect, and a constructed element, emotion. Although affect is a complex phenomenon, some elements may be more important toward the acquisition of moral concepts than others. In particular, although psychopathy and autism are often characterized in terms of empathic deficits, I argue that these deficits are radically different from one another. High functioning people with autism appear capable of acquiring moral concepts while psychopaths cannot. An investigation into their empathic capacities will help us make sense of our underlying moral capacities.

Psychopathy is a controversial name for a controversial diagnosis. In referring to psychopathy, I refer to the condition first identified by Hervey Cleckley in the 1940s and whose diagnostic criteria have been refined by Robert Hare and his psychopathy checklist (PCL-R). I do not simultaneously intend to refer to the diagnosis of Antisocial Personality Disorder (ASPD) in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association 2013; Hare, Hart, & Harpur 1991). Psychopathy, on this understanding, names a complex spectrum disorder diagnosed primarily in terms of character traits and emotional capacities. Because it is a spectrum

disorder, I limit my attention to a subclass of psychopaths referred to as successful psychopaths. Successful psychopaths score highly on the PCL-R but also lack other rationality-undermining comorbid illnesses (Babiak, Neumann, & Hare 2010).

Psychopaths are said to lack an ability to feel empathic distress. Empathic distress occurs when we find ourselves immediately pained when we witness another person suffering. It is the name given to a particular form of empathy that is basic in all three senses of the term. It is developmentally basic in the sense that, for nearly all humans, empathic distress develops by around twenty four months. It is compositionally basic in the sense that empathic distress cannot be broken down into still more primitive experience. Lastly, because this capacity is found in many mammalian species, there is evidence that it is phylogenetically basic as well. Empathic distress is thought to be mediated by mirror neurons in somatosensory cortex that activate in response to our perception of the suffering of others (Fecteau, Pascual-Leone, & Theoret 2008). In addition to lacking empathic distress, psychopaths are also characterized by personality traits including glibness, superficial charm, deceitfulness, and so on (Ramirez 2016). Although psychopaths are characterized in terms of an inability to feel empathic distress, successful psychopaths are often masters of another form of empathy: mindreading. This form of “[e]mpathy is the process by which one attains a cognitive grasp of, belief about, or knowledge of another’s mental states” (Battaly 2011, 286). Mindreading empathy is an especially useful trait for the psychopath as it allows them to deceive and manipulate others by predicting their mental states.

Autism is also a spectrum disorder. Autism, like psychopathy, is often characterized in terms of emotional deficits, empathic difficulty, and impairments in social function (McIntosh, Reichmann-Decker, Winkielman, & Wilbarger 2006).⁴ Although both psychopathy and autism

⁴ Proponents of neurodiversity would, rightly, reject characterizations of autism that cast it as a mental illness or disorder (Ramirez 2016). For the purposes of my argument, I focus only on population-level differences between persons with autism, psychopaths, and neurotypical persons. I leave aside questions about whether autism is best

include empathic deficits, these deficits refer to different capacities. Where the psychopath lacks a capacity for empathic distress and excels at mindreading, the high functioning agent with autism has an intact capacity for empathic distress but struggles with basic mindreading abilities.⁵

The empathic differences between high functioning persons with autism and psychopaths translate into differences in moral understanding and moral capacity. Psychopaths have difficulty distinguishing between what psychologists call moral norm violations and conventional norm violations. Although the distinction is controversial, norms of convention are thought to be authority-dependent, not based on harm, and limited in their context. For example, the norm for raising one's hand before speaking is authority-dependent. If a speaker announces that her audience should feel free to ask questions without raising their hands, the rule no longer applies. When studying these distinctions, psychologists construe moral norms as authority-independent, serious, harm based, and universal.⁶ Psychopaths tend to see all norms as norms of convention (Dolan and Fullam 2010). Additionally, psychopaths lack the sort of moral motivation I referred to earlier when discussing the special nature of sincere moral judgments. Because they see all norms as mere norms of convention, moral norms lack special force.

Most of us can easily note, for example, that there is a difference between how norms like 'it is wrong to wear white after Labor Day' and 'it is wrong to steal a person's identity.' Whereas the powers

viewed as a mental disorder or as an example of variation in human modes of cognizing the world.

⁵ As with any spectrum disorder, generalities like these paint very broad strokes about the capacities that individuals have. Although some researchers have argued that persons with autism do in fact have a diminished capacity for empathic distress, in comparison with the deficits seen with psychopaths, the differences between high functioning persons with autism and neurotypical persons are less stark. For example, psychopaths have profoundly diminished myelination in areas thought to contain mirror neurons which might explain their pronounced lack of empathic distress (Ly, Motzkin, Philippi, Kirk, Newman, Kiehl, & Koenigs 2012).

⁶ Although the moral/conventional distinction is controversial, I use it here to show that psychopaths inherit a kind of moral disability as a result of their lack of empathic distress (Greenspan 2003). I need not commit myself to the claim that the sense of "moral" defined by the moral/conventional task is essentially correct.

that be may change fashion norms, prohibitions against harming others non-consensually would continue to apply even if an authority figure claimed that they did not. The tendency to see moral norms as independent from authority runs so deep that many, including religiously orthodox communities, see moral norms as independent of god's authority (Nucci & Turiel 1993).

Children, including children with autism, begin to mark the moral/conventional distinction at a little over two years of age (Turiel 1977). Psychopaths, on the other hand, do not consistently or clearly note the difference. Although the moral concepts of individuals with autism can be structured differently from neurotypical moral concepts, their ability to mark the moral/conventional distinction and, especially, their desire to universalize the application of moral rules, makes clear that the moral concepts of persons with autism are full-blown moral concepts (McGeer 2008; Kennett 2002). The conceptual act theory of emotion can help to explain why persons with autism retain their capacity to form moral concepts and why psychopaths cannot.

The conceptual act theory of emotion states that emotional concepts are acquired via cultural transmission. The primary vehicle for the cultural transmission of moral concepts is empathic contagion (Ramirez 2016). In particular, children with an intact capacity for empathic contagion will mimic the emotional expressions of those around them, learn to give specific patterns of emotional expressions names, and thereby develop emotional concepts that structure their adult affective experiences (Feldman Barrett 2006). Different cultures will structure their emotional concepts differently, according to their own needs, though some particular emotional concepts will prove cross-culturally useful.

If moral concepts gain their meaning from emotional concepts, as sentimentalists claim, then empathic contagion plays an essential role in the acquisition of moral concepts. Because high

functioning persons with autism are capable of empathic distress, they will still be able to form moral concepts. Their issues with mindreading can make the application of these concepts difficult but their ability to note the distinction between purely authority-driven norms of convention and harm-based moral norms make it clear that they understand the moral domain as a distinct, affectively backed, domain (McGeer 2008).

Psychopaths, on the other hand, lack the capacity to engage in empathic distress. They thus cannot acquire moral concepts using the traditional route. Their moralized concepts have the form of descriptive anthropological data. Their inability to feel pained by the harm they cause or to feel guilt as a result of violating moral wrongs can thus be traced back to their inability to form genuine moral concepts. This is just as the conceptual act theory of emotion predicts.⁷

Sentimentalists interested in acquiring an empirically adequate picture of the nature of moral concepts and the structure of moral education ought to pay attention to the moral concepts of high functioning persons with autism and psychopaths. From these groups we learn that empathic contagion is the primary vehicle for the acquisition of moral concepts. Given the constructivist nature of the conceptual act theory, a sentimentalist constructivism emerges. Although this might appear to cut off the possibility of moral univocity, it does no such thing. Agricultural practices are fairly common, for example, and fairly stereotyped. The importance of crop rotation, irrigation, and other practices have been independently discovered several times in different cultures. Agriculture is not a basic human practice, however. Similarly, emotions are not basic in the senses mentioned earlier in this chapter (though our tendency to form emotional concepts may be). Some emotions may be so useful, however, that many cultures may construct emotions that all bear a family-resemblance to one

⁷ Psychopaths are not entirely devoid of moral emotions. Psychopaths are capable of intense anger, for example and so-called “secondary” psychopaths appear capable of a degree of shame though this shame is often focused on failing to achieve the recognition that secondary psychopaths believe that they are owed (Morrison & Gilbert 2001).

another. Shame, for example, is an intensely social emotion and can track social relationships in highly complex ways. In cultures where such relationships are highly valued, an emotion very much like shame is likely to be found. Because at least some emotional concepts are useful to regulate human social interactions (shame and embarrassment being prime candidates), these emotions can still serve as a basis for explaining moral disagreements in a more limited fashion. To the degree that emotional concepts across cultures overlap with one another, then evaluative discourse about the values these emotions carve out is possible. Though less far-reaching than basic emotions, these emotions would permit cross-cultural moral engagement.

VI. Conclusion

Sentimentalist philosophers have turned to psychology and the neurosciences in the hopes of vindicating their intuition that moral concepts are grounded in our emotions. Though initially this empirical turn appeared to support the existence of basic emotion, these data have received significant criticism. In place of basic emotion, conceptual act theories have emerged to explain the seeming variation and universality of emotional experience. These views receive further support from the moral concepts of high functioning persons with autism and psychopaths. These groups provide strong evidence for the claim that empathic contagion is the primary vehicle for the acquisition of moral concepts. Because empathic contagion is basic, we should expect to see this capacity in all neurotypical human populations. Sentimentalists worried about the prospects of constructivism that emerges from conceptual act approaches can manage to explain many moral disagreements even without appeal to basic emotion.

Works Cited

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders 5th ed.* Washington, DC.
- Babiak, Paul, Craig Neumann, and Robert Hare. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences and the Law* 28, 174-193.
- Battaly, H. D. (2011). Is empathy a virtue? In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives*, Oxford: Oxford UP.
- Cardoso, F.M.S. (2015). Social regulation of emotion: A foray into the work of Catherine Lutz. *Revista E-Psi*, 5(2), 103-109.
- D'Arms, J. (2005). Two arguments for sentimentalism. *Philosophical Issues*, 15(1), 1-21.
- D'Arms, J. & Jacobson, D. (2000). 'Sentiment and value.' *Ethics* 110, 722–748.
- Darwin, C. (1873/2009). *The expression of the emotions in man and animals*. Oxford University Press
- DiGirolamo MA, & Russell JA. (2016). The emotion seen in a face can be a methodological Artifact: The process of elimination hypothesis. *Emotion*, Nov 28. [Epub ahead of print]
- Dolan, Mairead and Rachael Fullam. (2010). Moral/conventional transgression distinction and psychopathy in conduct disordered adolescent offenders. *Personality and Individual Differences* 49, 995–1000.
- Ekman, P. & Friesen, W. (1971). Constants across culture in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124-129.
- Ekman, P. (1972a). *Emotions in the human face*, New York: Pergamon Press.
- Ekman, P. (1972b). Universal and cultural differences in facial expressions of emotions. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971, pp. 207-283. Lincoln: University of Nebraska Press
- Ekman, P. (1994). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169-200.
- Fecteau S., Pascual-Leone A., Theoret H. (2008). Psychopathy and the mirror neuron system: preliminary findings from a non-psychiatric sample. *Psychiatry Research* 160, 137–144.
- Feldman Barrett, L. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10(1), 20-46.

- Feldman Barrett, L. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, 23(7), 1284-1306.
- Fernández-Dols J. M. & Crivelli C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review*, 5, 24–29.
- Fernández-Dols, J. M. & Ruiz-Belda, M.A. (1997). ‘Spontaneous facial behavior during intense emotional episodes: Artistic truth and optical truth.’ In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression*, 255-274. New York: Cambridge University Press.
- Gendron M., Roberson D., van der Vyver J.M., Feldman Barrett L. (2014). ‘Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture.’ *Emotion*, 14(2), 251-62.
- Greenspan, P.S. (2003). ‘Responsible psychopaths.’ *Philosophical Psychology*, 16(3), 417-429.
- Haidt, J., Helena Koller, S. & Dias M.G. (1993). ‘Affect culture and morality, or, is it wrong to eat your dog?’ *Journal of Personality and Social Psychology*, 65(4), 613-28.
- Haidt, J. (2001). ‘The emotional dog and its rational tail: A social intuitionist approach to moral judgment.’ *Psychological Review*, 108, 814–834.
- Hare, R., Hart, S., & Harpur, T. (1991). ‘Psychopathy and the DSM-IV criteria for antisocial personality disorder.’ *Journal of Abnormal Psychology*, 100(3), 391-398.
- Hume, D. (1740/2002). II.3.3 415. In D.F. Norton & M.J. Norton (Eds.), *A treatise of human nature*. New York: Oxford University Press
- Jacobson, D. (2013). ‘Moral dumbfounding and moral stupefaction.’ M. Timmons (Ed.), *Oxford Studies in Normative Ethics*, Volume 2 (pp. 289–316). Oxford: Oxford University Press.
- Kennett, Jeanette (2002). ‘Autism, empathy and moral agency.’ *The Philosophical Quarterly*, 52, 340-357.
- Kleinman, A. (2004). *Culture and depression*. New England Journal of Medicine 351, 951-953
- Landis, C. (1924). ‘Studies of emotional reactions: General behavior and facial expression.’ *Journal of Comparative Psychology* 4(5), 447-509.
- Ly, M., Motzkin, J. C., Philippi, C. L., Kirk, G. R., Newman, J. P., Kiehl, K. A., & Koenigs, M. (2012). ‘Cortical thinning in psychopathy.’ *The American Journal of Psychiatry*, 169(7),

10.1176/appi.ajp.2012.11111627.

Matsumoto, D. (1992). 'More evidence for the universality of a contempt expression.' *Motivation and Emotion* 16: 363–368.

McGeer, V. (2008). 'The Makings of a moral sensibility: Replies to commentaries.' In W. Sinnott-Armstrong (Ed.) *The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, 281-296. MIT Press

McIntosh, D.N., Reichmann-Decker A., Winkielman, P., & Wilbarger, J.L. (2006). 'When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism.' *Developmental Science* 9(3), 295–302.

Morrison, D. & Gilbert, P. (2001). 'Social rank, shame and anger in primary and secondary psychopaths.' *The Journal of Forensic Psychiatry*, 12(2), 330-356.

Nucci, L. and Turiel, E. (1993). 'God's word, religious rules, and their relation to Christian and Jewish children's concepts of morality.' *Child Development* 64, 1475-1491.

Ortony, A. & Turner, T. (1990). 'Whats basic about basic emotions?' *Psychological Review*, 97(3), 315-331.

Phan K.L., Wager T., Taylor S.F., Liberzona, I. (2002). 'Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI.' *NeuroImage*, 16(2), 331–348.

Ramirez, E. (2016). 'Neurosurgery for psychopaths? The problems of empathy and neurodiversity.' *American Journal of Bioethics: Neuroscience*, 7(3), 166-168.

Reisenzein R., Studtmann M., & Horstmann G. (2013). 'Coherence between emotion and facial expression: Evidence from laboratory experiments.' *Emotion Review*, 5(1), 16 - 23

Russell, J.A. (1995). 'Facial expressions of emotion: What lies beyond minimal universality?' *Psychological Bulletin*, 118(3), 379-391.

Russell, J.A. (2016). 'A skeptical look at faces as emotion signals.' In C. Abell & J. Smith (Eds.) *The expression of emotion: Philosophical, psychological, and legal perspectives*, 157-172. New York: Cambridge University Press

Turiel, E. (1977). 'Distinct conceptual and developmental domains: Social convention and morality.' *Nebraska Symposium on Motivation*, 25, 77–116.

Wager T.D., Kang J., Johnson T.D., Nichols T.E., Satpute A.B., & Feldman Barrett L. (2015) 'A bayesian model of category-specific emotional brain responses.' *PLoS Computational Biology* 11(4):

e1004066. Doi: 10.1371/journal.pcbi.1004066

Wong, Y., & Tsai J. (2007). 'Cultural models of shame and guilt.' In J. Tracy, R. Robins & J. Tangney (Eds.) *Handbook of Self-Conscious Emotions*. New York, NY: Guilford Press: 210-223.