

Jesse Prinz on Emotional Conditioning (20170508)

Alonzo Fyfe

To believe that something is morally wrong (right) is to have a sentiment of disapprobation (approbation) towards it. (EBMJ 33)¹

Jesse Prinz presents this as the "first approximation" of the thesis he seeks to defend in "The Emotional Basis of Moral Judgment."

Prinz argues that empirical research in psychology can help us to answer moral questions. Specifically, he argues that empirical research in psychology tells us that moral judgments are based on emotion. Among the set of empirical observations that he draws upon is the fact that parental instruction of children substantially involves, not the teaching of facts, but emotional conditioning. This moral conditioning aims to link, in the child, negative emotions to misdeeds.

I wish to argue that moral instruction – even if Prinz's description of it is accurate - does not provide evidence in defense of his thesis. Quite the contrary. If Prinz's account of moral instruction is correct, it provides an argument against his thesis. Rather than identify a misdeed as that to which one has a negative emotion, it says that a misdeed is "something else" and that the purpose of moral instruction is to link negative emotions to this "something else".

However, as Prinz said, the version of his thesis that I quoted above was his "first approximation" of that thesis. He then introduced a set of refinements. In what follows I wish to examine those refinements to try to determine Prinz's refined thesis. Then I wish to show that the objection above still applies – the phenomenon of moral instruction is still inconsistent with that refined thesis. Then I want to make one small adjustment to Prinz's set of premises and show that there is an alternative thesis, similar to Prinz's refined thesis but also recognizably distinct, that can handle the phenomenon of moral instruction.

Prinz's Refinements

After presenting us with a "first approximation" of his thesis, Prinz mentions that some refinements are needed. He does not present us with a final, refined thesis. Consequently, I will try to collect his suggested refinements into a final refined thesis.

¹ Jesse Prinz, "The Emotional Basis of Moral Judgments", *Philosophical Explanations*, Vol. 9, No. 1, March 2006, 29-43. References to this work will be abbreviated (EBMJ p.##)

First, Prinz points out that it is possible to express a moral judgment without having much sentiment behind it. Moral philosophers are prone to this phenomenon as they often do not experience a strong sentiment about the wrongs they use to illustrate their moral arguments. A scholar might become accustomed to speaking and writing about racism to the degree that she can mention its wrongness without having a strong sense of disapprobation at the time she is saying that it is wrong.

Prinz handles these types of cases by claiming that a sentiment is a dispositional state, not always an actual occurrent state.

A sentiment is an emotional disposition, and I can have a sentiment without manifesting it. I love John Coltrane, and I can truly self-ascribe this sentiment, but I am not always experiencing the hedonic rush that I get when I listen to A Love Supreme. Likewise, I can testify that I think gender discrimination is wrong without experiencing any outrage. (EBMJ p. 38)

We can make this explicit in his thesis: To say that something is wrong is to be disposed to have an attitude of disapproval towards it. Having, at a particular moment, no attitude – because one’s mind is on other matters, for example, or one has dealt with the issue constantly – does not refute this thesis. Though, according to Prinz, it would be odd to think that a person morally objects to some action while, at the same time, he never really cares whether people perform that action.

A second and third refinement is that the sentiment only identifies something as wrong if a sentiment held, “under conditions of full factual knowledge and reflection, and freedom from emotional biases that I myself would deem as unrelated to the matter at hand,” (p. 35).

One person, being pushed by another, may have an immediate attitude of disapprobation at the act of being pushed. Then she learns that the other person had pushed her out of the path of a runaway trolley. Here, the judgment made with the awareness of the reason one had been pushed is the judgment that counts – not the judgment made in haste and in ignorance.

Prinz also allows a person to dismiss sentiments that can be traced to a bias that the person would not endorse. This allows for the possibility, for example, that an agent may judge an action performed by a woman to be wrong. He is then made aware of the fact that if a man were to perform that same action, he would not judge it wrong. He has a further sentiment that it is wrong to have one standard for men and a different standard for women. This allows him to discard the judgment based on his sexist sentiments.

It also follows from this formulation that if the agent deems gender to be relevant to the matter at hand, then it would follow that if he would condemn in a woman what he would not condemn in a man, that this would be a legitimate moral claim. Some may consider this an objection to Prinz's thesis – as I would. However, my objections will lie along a different track, so I will set this problem aside for now. I will say a little more about it near the end.

We can now formulate Prinz' thesis as saying:

To believe that something is morally wrong (right) is to have a disposition to have an attitude of disapprobation (approbation) towards it under conditions of full factual knowledge and reflection and freedom from emotional biases that I myself would deem unrelated to the matter at hand.

However, this is only a "second approximation" of his thesis. To get to the final refined version, we have another question to answer.

Belief, Truth, and Meaning

In attempting to determine Prinz' refined formulation, I encountered an important area of confusion.

"I believe that X is P", "X is P" and "X means, at least in part, P" are three different claims. Prinz's "first approximation" and the "second approximation" that I drew from his further comments both talk about what it takes to believe X is P – what it takes to believe that something is morally wrong. However, in other parts of his essay, Prinz seemed to be talking about what it takes for something to actually be morally wrong. At other times Prinz wrote about the meaning of moral terms.

To see the difference between these three types of claims, let us consider the claim, "there is a pink elephant in the room." We may be able to establish that for Pete to believe that there is a pink elephant in the room he must have a blood alcohol content of at least 0.2%. Yet, even if Pete is suitably drunk and being drunk is necessary and sufficient for him to have the belief, this says nothing about whether "there is a pink elephant in the room" is true. That is a separate question. Both of these claims are different, in turn, from the claim that the phrase, "there is a pink elephant in the room," spoken by Pete, has as a part of its meaning that Pete has a blood alcohol content of 0.2%. It is possible that one may be able to make this inference, but it may be a causal (material) inference, not a logical inference.

Similarly, it may be the case that for Patty to believe that infanticide is wrong she must have an attitude of disapprobation towards infanticide. Yet, even if Patty has this attitude and has the corresponding belief, it tells us nothing about what it takes for infanticide is wrong (if anything). And though we may be

able to imply that Patty has an attitude of disapprobation towards infanticide from her belief that infanticide is wrong, this may be a causal and not a logical inference.

Prinz's "first approximation" is stated as a thesis about moral judgments – about what it takes to believe or judge that X is wrong. It is not necessarily a thesis about X being wrong or about what it means to say that X is wrong.

Yet, elsewhere, Prinz's clearly talks about the meaning of terms. In other passages, he seems to be concerned with what it takes for something to actually be moral or immoral – that it is something capable of causing a particular sentiment in the speaker.

On the matter of what makes a moral claim true – for something to be moral or immoral – Prinz writes the following:

The judgment that something is funny is justified by our amusement, because causing amusement is constitutive of being funny. If moral judgments are sentimental, and they refer to response-dependent properties, then the judgment that killing is wrong is self-justifying because killing elicits the negative sentiment expressed by that judgment and having the power to elicit such negative sentiments is constitutive of being wrong. (EBMJ p. 37)

The claim here is that the judgment that something is wrong simply is the judgment that it is such as to cause a particular sentiment within us. Thus, he will argue, a moral claim has a truth-value – and that some moral claims can be true. This is because the claim that something can cause a particular sentiment within us can have a truth value, and can sometimes be true.

Prinz distinguishes his thesis from emotivist non-cognitivism by asserting that moral claims, on his thesis, are 'truth-apt'.

Sentimentalism so-defined has a major advantage over expressivism. Moral judgments are truth-apt, if they refer to response-dependent properties, just as their surface form would suggest. (EBMJ p. 35)

Moral propositions are capable of being true or false. However, Prinz takes no steps that suggests that he is heading towards an error theory regarding moral statements. That some of these truth-apt statements can be true seems to go without saying.

Indeed, Prinz offers sentimentalism as a way of explaining self-justifying intuitions.

If moral judgments are sentimental, and they refer to response-dependent properties, then the judgment that killing is wrong is self-justifying because killing elicits the

negative sentiment expressed by that judgment and having the power to elicit such negative sentiments is constitutive of being wrong. Sentimentalism explains the phenomenology driving intuitionism, and it shows how intuitionism might be true. (EBMJ p. 37)

Elsewhere, Prinz states that his thesis is about the meaning of moral terms. For example, he writes:

Here, I think something like Dreier's speaker-relativism is right. When I say that something is wrong, I refer (perhaps unwittingly) to the property of causing emotions of blame in me. (Or perhaps, saying that something is wrong means that it causes emotions of blame in us, where the 'us' refers to a group of people to whom I would morally defer. (EBMJ p. 35)

Or consider:

I conclude that ordinary moral concepts do not refer to properties that can be coherently characterized without reference to our sentiments. (EBMJ p.41).

This evidence suggests that, even though Prinz' first approximation concerned belief, the fully refined thesis included claims about the meaning of moral terms and what it takes for a moral claim to be true.

In summary, the final refined statement of Prinz's thesis should take it to be a claim not only about moral belief, but about the truth of moral claims and the meaning of moral terms. In Prinz's final analysis, for something to be morally wrong it has to be such as to create a particular emotional response on the person making the claim. This is not only a contingent truth, it is a logical truth.

This identifies what it is that Prinz wanted to argue for. With this in mind, I would like to look at one of his arguments in defense of this thesis – an argument from moral instruction. I wish to show that not only does Prinz's evidence not support his thesis, it actually provides an argument against his thesis.

The Argument from Moral Instruction

As a part of Prinz's argument in defense of this thesis, Prinz sought to argue that emotions are necessary for moral appraisal. In support of this point, he brought up the observation that emotions seem to be necessary for moral instruction. The way in which parents instruct their children is by using techniques that "recruit emotions".

Specifically, Prinz writes:

Unlike language, children need a lot of training to conform to moral rules, and parents spend a lot of time giving their children moral instruction. Interestingly, the three main techniques that parents use to convey moral rules all recruit emotions. One technique is power assertion (physical punishment or threat of punishment), which elicits fear. Another technique is called induction, which elicits distress by

orienting a child to some harm she has caused in another person ('Look, you made your little brother cry!'). The third technique is love withdrawal, which elicits sadness through social ostracism ('If you behave like that, I'm not going to play with you!'). Each technique conditions the child to experience negative emotions in conjunction with misdeeds. This does not prove that emotions are necessary for moral development, but it is suggestive. (EBMJ pp. 31-32)

In other words, what the parent is teaching the child is not a set of facts – and it is not being taught in the way that facts are taught. When it comes to moral training, the parents are looking at the emotions – using techniques of emotional conditioning – and seeking to link emotions to misdeeds. This suggests, according to Prinz, that, whatever morality is, it seems to concern the emotions.

If we look at these practices closely, we will discover that, even though it supports a link between morality and emotions, the link is inconsistent with Prinz's thesis.

The Child's Point of View

To see the problems that moral instructions raise for Prinz, I would like to look at it from the child's point of view. Let us imagine a child who has hit her younger brother, for which her parents have scolded her. They may have used any or all of the techniques that Prinz mentioned – pointing out that she made her little brother cry, taken something away from her child as punishment, and then sending her to her room.

Let us further assume that these techniques work the way Prinz says they do. The work to condition the emotions – to associate the misdeeds with negative emotions.

When the child's parents scold and punish her for hitting her brother, what does the child hear?

Or, more importantly, what does the child not hear?

When the parents tell her that hitting her brother is wrong, the child will not likely hear any invitation for her to look to her own sentiments to determine if she should attach the term "wrong" to hitting her brother. She hears her parents saying that the wrongness is a fact of the matter – something which she would be wise not to dispute, at least not at this point in her life.

I would like to illustrate this point by looking at a term that does seem to involve the type of agent-centered sentimentalism that Prinz ascribes to morality – the concept of "delicious". The child's parents may tell the child that tomato soup is delicious. The child tastes it and does not like it. She then learns, through interaction with her parents and others, that the fact that she does not like tomato soup means that it is not delicious-for-her. Tomato soup is only 'delicious' for those who like the taste.

This is quite different from what the child hears when her parents tell her that hitting her brother is wrong. If she were to look at her own sentiments and find no sentiment attached to hitting her brother she would not discover with her parents and others that hitting her brother is not wrong-for-her. She will discover them saying that hitting her brother is still wrong, and, furthermore, that she is a bad person if she does not share that sentiment.

This argument speaks to one of the components of Prinz's thesis – the doctrine of internalism.

Externalism versus Internalism

Internalism is a doctrine that holds that when an agent believes that something is obligatory, then that agent will be motivated to perform that action. If an agent believes that something is wrong, then the agent will be inclined to avoid performing such an action. It is not possible for an agent to have a moral belief without having motivation.

Some moral internalists argue that moral judgments come with overwhelming motivational force. A person who believes that something is wrong would be motivated beyond all other concerns to refrain from performing that action. However, Prinz accepts weak internalism. Weak internalism allows one's motivation to avoid doing that which one judges to be wrong to be overridden by other concerns.

This project has important implications for our understanding of practical reasoning. It bears, most directly, on the debate between motivational internalism and externalism in ethics. Can we make moral judgments without being motivated to act? If emotions are linked to moral judgments in an intimate way, then the answer may be negative. Moral judgments are intrinsically action-guiding. I will defend a version of the internalist position. (EBMJ p. 30)

Prinz lists it as an advantage for his thesis that it accounts for a link between motivation and moral judgment. For a person to say, "that is wrong" (or just to believe it) and not to be averse to performing that action is considered as odd as saying, "this ice cream is delicious" and yet not enjoy the taste of it. If one says that something is wrong, and does not care to refrain from doing it, one does not understand the meaning of the term.

Internalism seems to have some appeal when we look at a moral claim from the point of view of the person who makes it. People do, in fact, seem to judge favorably what they are inclined to do, and to judge harshly that which they show a strong aversion. They do so with such consistency that one has reason to suspect that one cannot understand what it means to say something is good or bad if one lacked a motive to realize or avoid it accordingly.

The standard philosophical objection to internalism looks at the conceptual possibility of the amoralist. The moral externalist does not need to deny that people often link their moral judgments with their dispositions to act. The externalist needs only to argue that it is at least conceivable that a person can separate these two – can judge something to be wrong without having an inclination against performing it.

We can find support for this conceptual possibility in looking at a moral scolding from the point of view of the child. The child is told to apply “wrong” to acts such as hitting her brother without regard to her own sentiments. If she lacks the attitude of disapprobation towards hitting her brother, the conclusion is not, “hitting my brother is not wrong-for-me.” The conclusion is, “I am a bad person. A good person would have this attitude of disapprobation towards such acts.” She might actually care about whether she is a good person, and thus have some motivation not to hit her brother, but the concept of “wrong” her parents throw at her is still an externalist concept – one that tells her to assign this term to hitting her brother and does not to inquire into her sentiments before accepting their claim as true. In this way, it is quite different from a term like ‘delicious’.

The child can make the further inference that the term her parents use is the same term they learned as children when they were scolded for hitting their own siblings.

This message that there is a gap between what is wrong and the child’s attitude of disapprobation is reinforced by another fact that Prinz puts a lot of emphasis on – the fact that acts of emotional conditioning accompany the parents’ moral judgment

Emotional Conditioning

The claim that what the child hears when her parents tell her that it is wrong to hit her brother is an externalist concept of “wrong” that does not invite her to consider her own sentiments before accepting the claim as true. This is reinforced by the fact, as Prinz reports, that techniques of moral conditioning often accompany the claim.

Recall that Prinz included in these techniques actual punishment, ostracism, and ‘induction’ such as, “Look, you made your little brother cry.” All of these are designed to link negative emotion to behavior.

But if the goal is to link negative attitudes to misbehavior, then the child is being told that misbehavior has to be something other than whatever the child has negative attitudes towards. If the child understands misbehavior to be “behavior that I have an attitude of disapproval towards,” then what is this emotional conditioning for? “Misbehavior” – when accompanied by emotional conditioning, seems

to be saying, “misbehavior is this other thing – over there. It is, at best, what you would have negative reactions toward if we can get those negative emotions properly tuned.” Yet, the very idea of improperly calibrated sentiments implies a wrongness that is independent of those sentiments.

For example, if we were to understand time as "whatever time is shown on a clock," then there would never be a need to adjust the clock – to set it to a different time or to adjust the speed at which it progresses. The clock could never be wrong. We may, to draw the analogy closer to Prinz's view of morality, say that we must be in a state so as to read the clock correctly – under good lighting, for example. This would correspond to Prinz's claim that morality is whatever disposes us to have a certain reaction if we were fully informed and free of biases. Yet, the view that the correct time is whatever the clock says viewed under these conditions is inconsistent with the claim that the clock might need to be adjusted. The very idea of adjusting the clock suggests that the clock can be wrong, that we need to look someplace else to determine the correct time, and then make an adjustment to the clock.

Our scolded child, then, can only make sense of this emotional conditioning if she assumes that the wrongness of hitting her brother is something other than that which she has a sentiment of disapprobation towards – otherwise she would not need to have her sentiments adjusted.

The child will find further evidence for this interpretation in claims that her parents make such as, “You should be ashamed of yourself.” Note that this phrase does not mean, “You would be ashamed of yourself if you were fully informed and free of those biases that you do not endorse.” It means, literally, “Your sentiments, if they were properly calibrated to this situation, would be sentiments of shame, and, if it is not, then your sentiments are mistaken.”

Prinz himself agrees that moral instruction works with molding sentiments, not with teaching the child to be aware of relationships to sentiments that the child already has – or even would have under full information – but with altering the child’s sentiments:

Emotional conditioning and osmosis are not merely convenient tools for acquiring values: they are essential. Parents sometimes try to reason with their children, but moral reasoning only works by drawing attention to values that the child has already internalized through emotional conditioning. (MCCR)²

² Prinz, Jesse. "Morality is a culturally conditioned response." *Philosophy Now* 82 (2011): 6-9., https://philosophynow.org/issues/82/Morality_is_a_Culturally_Conditioned_Response, last referenced 04/23/2017. References to this work will be abbreviated (MCCR)

The purpose of moral conditioning is to change sentiments, not to draw attention to the relationships between states of affairs and sentiments that already exist. This implies that the sentiments themselves can be mistaken, and that wrongdoing can be something other than that to which the sentiments actually point.

All of these features tell us that, even if emotional conditioning works exactly as Prinz tells us it does, this is not consistent with the child learning that rightness and wrongness are expressions of her own sentiments of approbation and disapprobation. Emotional conditioning tells us that there is such a thing as a miscalibrated emotion – an emotion that does not correctly identify what is wrong. This possibility suggests that a person’s attitude may be in need of adjustment.

And how many times has a child heard this? “You had better change your attitude young man/woman.”

Moral Osmosis

Prinz also asserts that children acquire their moral training, in part, through osmosis.

Children also learn by emotional osmosis. They see their parents’ reactions to news broadcasts and storybooks. They hear hours of judgmental gossip about inconsiderate neighbors, unethical coworkers, disloyal friends, and the black sheep in the family. Consummate imitators, children internalize the feelings expressed by their parents, and, when they are a bit older, their peers. (MCCR)

In hearing the adults around her talk about the co-worker’s disloyalty, the rudeness of the shopper at the grocery store, and the recklessness of the driver on the road, according to Prinz, the child will come to share these sentiments. She will mimic those around her and come to acquire the same attitude of disapprobation towards disloyalty, rudeness, and recklessness. Because she adopts these same sentiments, she makes the same moral judgments.

Yet, here, too, we need to pay attention to what the child actually hears. When she hears that Uncle Jeb is a slob and does not care how filthy his house gets, that her cousin Jenny is spoiled and self-centered, and that Jimmy’s friend Bob in college is a bad influence because all he wants to do is party, she notices that here, too, the people who are speaking are not inviting any attention to the attitudes of the people they are speaking about.

In fact, she hears people condemning others in the same terms they used to condemn her. In the same way that she learns that she cannot change the judgment by appealing to her own sentiment, the other people they complain about cannot change the judgment according to their sentiment. Continuing the analogy to what is “delicious”, the people who her parents and others complain about cannot claim that

the actions are not ‘wrong-for-them’ in virtue of lacking a sentiment of disapprobation in the way they can dispute the claim that something is not “delicious-for-them” in virtue of the fact they do not like it.

From time to time – though this would be very rare – she may hear those around her talk about what some distant group of people view as ‘wrong-for-them’. These people could be in the distant past, or a distant part of the world, or a somewhat isolated culture nearby. They may identify some things as ‘wrong-for-them’ because of their religion, or cultural tradition, or local custom. From this, she may get an idea of a type of ‘wrong’ that is merely an expression of the attitudes of an individual or group.

However, she should recognize a difference between this use of the term ‘wrong’ to mean ‘wrong-for-them’ and the ‘wrong’ she uses when talking about Uncle Jeb and her cousin Jenny. When parents and other adults speak of the morality of other cultures in an anthropological sense – in a sense merely of reporting what they believe – those claims are not associated with emotional conditioning. The child learns that she is not really expected to adopt those sentiments – that her sentiments can point to something other than what is wrong-for-them (whomever her parents are talking about) without the implication that she is a bad person – that she needs to adjust her attitude.

‘Wrong’ in the more common sense has a different meaning – one that implies that the child herself not only does, but should, adopt the attitudes that those around her endorse. The implication that comes from not having the proper attitudes towards that which she hears people condemn in others is not, “therefore, these things are not wrong for me.” The implication is, “therefore, I, like Uncle Jeb or my cousin Jenny, am a bad person.”

Calibrating Moral Sentiment

So far, I have argued that Prinz’s claims about emotional conditioning teaches an externalist conception of morality where the child finds no invitation to base moral judgments on her own system, and emotional conditioning is used to calibrate sentiments to external moral facts.

But what are these external moral facts that these sentiments are calibrated for?

Rather than canvas the various forms of moral realism and objectivism, I wish to present an option that seems to come from Prinz’s own claims and research regarding moral judgments. I wish to start with a set of propositions that I think Prinz would have to accept. Imagine that you have crash-landed on a habitable planet and that the following describes your situation:

(1) Imagine that you find yourself in a community of individuals who are substantially sentimental creatures. That is to say, their sentiments provide almost all (if not all) of their motivation to perform intentional actions.

(2) These sentiments are not fixed. Instead, you discover that interactions between these creatures and their environment can alter these sentiments.

(3) You are a part of these creatures' environment.

(4) You can determine, at least approximately, how interactions with the environment can change their sentiments.

In addition to these four items, which I do not think Prinz would dispute, Prinz's discussion of emotional conditioning suggests a fifth item that – though it can be made consistent with the other claims mentioned above, does not fit Prinz's refined thesis.

(5) It would be prudent for you to promote in other creatures those sentiments that are consistent with your interests – desires that are compatible with you obtaining that which suits your interests or that, at least, will give the creatures a reluctance to act in ways that would thwart your interests.

To illustrate this proposition, it would be prudent of you, assuming that the first four criteria are met, to emotionally condition predators to stay out of the territory where you live. It would be prudent to have large animals accept the burden of pulling a plow or performing other work.

With respect to more intelligent intentional agents, it is at least reasonable to believe that you would be well served to create in them aversions to lying, breaking promises, theft, vandalism, and assault. To take these tools and to use them only to express emotions, when you can do something with them that would be constructive in improving the quality of your life in the context of interacting with other intentional agents, would appear to be a waste.

However, Prinz seems to see this emotional conditioning as merely the expressions of emotions – something that children then pick up, but which does not serve any other purpose. In an earlier part of his discussion, he reports on the experiments of Jonathan Haidt and others that suggest that we can alter a person's moral judgments by altering their emotional states. One often-cited experiment involves giving agents descriptions of various types of actions and asking them to evaluate the degree of wrongness of the action in a setting where used tissues, greasy pizza boxes, and the like prompt feelings of disgust. Those who evaluate the actions in this disgusting environment tend to make harsher

judgments than those who evaluate the actions in a clean environment. This suggests that judgments are the expressions of sentiment.³

Prinz appears to be assuming the legitimacy of these moral judgments. However, Karen Jones points out that we would further expect these research subjects, upon being debriefed on the results of this research, to revise their previous moral judgments – to dismiss the disgust as morally irrelevant factors. In fact, Jones argues, a subject who would hold on to her judgment in light of this information would be charged with failure to understand what moral judgments are about. This second-order moral judgment, or judgment about the judgments one makes when disgust is triggered, goes against the idea that moral judgment is merely an expression of emotion.

Jones' observations also do nothing to support the thesis that moral judgments concern the usefulness of emotional conditioning. It is consistent with that thesis, in that we can dismiss the disgust triggered in these experiments as telling us nothing about the usefulness of promoting such a sentiment in others, and this is why it is morally irrelevant. However, we would need a lot of data points of this type to defend such a hypothesis. Here, I am merely offering it as an illustrative alternative to the thesis that Prinz appears to want to defend.

On this matter, there are two things that Prinz can say in his own defense.

First, please recall that the refined thesis says that a moral judgment on Prinz's account requires "freedom from emotional biases that I myself would deem unrelated to the matter at hand." If the agent judges this disgust to be irrelevant then, for that agent, it is irrelevant.

Though, on the other side of this, if an agent judges that his feelings of disgust are relevant to his moral judgment – if he does not see them as providing a bias - then they would, in fact, be morally relevant. The evaluator who accepted these external sources of disgust as being morally relevant would have them be morally relevant in fact. Against this, Jones argues that a person who appeals to these external sources of disgust as justifying a moral belief would not be using moral terms correctly.

We must also consider the fact that Prinz uses the fact that it is possible to alter a person's moral judgments by altering that person's emotional states – e.g., by creating a sense of disgust – as an

³ Schnall, Simone, Jonathan Haidt, Gerald L. Clore, and Alexander H. Jordan. "Disgust as embodied moral judgment." *Personality and social psychology bulletin* 34, no. 8 (2008): 1096-1109.

argument in defense of his thesis. It would be odd for him to go from this to arguing that these same sentiments are morally irrelevant.

Second, when Prinz attempts to defend his brand of moral relativism, he brings up the fact that some sentiments can be judged to be more useful than others. For example, in arguing that a moral relativist can still condemn Hitler, he wrote:

First of all, Hitler's actions were partially based on false beliefs, rather than values ('scientific' racism, moral absolutism, the likelihood of world domination). Second, the problem with Hitler was not that his values were false, but that they were pernicious. Relativism does not entail that we should tolerate murderous tyranny. When someone threatens us or our way of life, we are strongly motivated to protect ourselves. (MCCR)

And in defending his relativism from the claim that it does not allow for moral progress, he wrote:

In one sense this is correct; moral values do not become more true. But they can become better by other criteria. For example, some sets of values are more consistent and more conducive to social stability. If moral relativism is true, morality can be regarded as a tool, and we can think about what we'd like that tool to do for us and revise morality accordingly. (MCCR)

Here, Prinz acknowledges that we can evaluate a sentiment according to its usefulness, and that we have reasons to use emotional conditioning to promote useful sentiments.

However, his revised thesis grounds moral judgments on the sentiments that one has, not on the sentiments that one would have if one had those sentiments that are useful to others (that others have reason to cause him to have through emotional conditioning). In fact, he speaks against neo-sentimentalist theories that argue that morality is grounded on sentiments that are somehow, themselves, warranted in some way. This would appear to argue against basing morality on sentiments that are warranted in virtue of being useful – in virtue of being those sentiments that people generally have reasons to promote universally.

I would like to point out that there is still room here for morality to be grounded on sentiment.

However, it is not the case that my sentiment that distinguishes right from wrong. It is the sentiments that give people generally their reasons to praise or condemn in order to calibrate my sentiments and those of others. If properly calibrated, then I can use my sentiments as a guide to moral truth, but miscalibration remains a possibility. This is what the child learns when she is told that hitting her brother is wrong. She learns that emotional conditioning may be necessary to align sentiment with an independent moral fact.

Prinz's revised thesis appears to have some legitimacy when we look at moral judgment from the point of view of the person making the judgment. However, this changes significantly when we look at it from the point of view of the person being condemned. When the girl suffers her parents' condemnation for hitting her brother, she hears nothing that tells her that she can dismiss the claim that the actions are wrong so long as she has no sentiment of disapprobation towards hitting her brother. On the other hand, it does seem to make sense to hear her parents as saying, "In calling this 'wrong' we do not care one iota as to what your sentiments tell you on this matter, you are well advised to adopt a sentiment of disapproval towards acts like that of hitting your brother and to act accordingly."

Conclusion

Prinz argues that a person's emotional state can heavily influence his moral judgments. However, the question remains as to whether one's emotional state should have the influence it does over those judgments – whether those judgments are little more than an expression of such a state.

Actually, it is likely that few people would dispute the claim that emotions do influence our moral judgments. The question is whether the sentiments provide the foundation for those judgments, or whether they cloud those judgments. We are constantly being warned of the dangers of allowing our emotions to cloud our judgments. Evidence that this type of influence does, in fact, occur makes a poor argument for the thesis that this is not the "clouding" of moral judgment but is the essence of moral judgment.

When the girl hit her brother, there is little reason to doubt that she acted on her sentiments – which may well not have included a sentiment of disapproval towards hitting her brother. When her parents condemn her for the action, they do not do so in a way that disputes that the girl will act on her sentiments. In fact, their condemnation takes this as given and, with this fact in mind, seeks to give her those sentiments to act on that are useful to others. They seek to include among those sentiments an aversion to such things as lying, breaking promises, theft, vandalism, and – particularly in this case – an aversion to acts like that of hitting her brother.

When we examine the usefulness of sentiments, as opposed to whether they are the direct objects of approbation or disapprobation – we get a sense of morality that is more in line with contemporary moral discussion.

Seeing morality in terms of useful sentiments can be expressed as a type of sentiment consequentialism: the right act is the act that expresses the best sentiments, where the best sentiments are those that

produce the best consequences. Those consequences, in turn, are what generally give other people reasons to promote the sentiment through the techniques of emotional conditioning. This would be like rule consequentialism, except that under the assumption overriding a sentiment (as opposed to violating a rule) may only be possible if more or stronger sentiments outweigh it. If other sentiments are needed, then evaluating the action would require evaluating the merits of the sentiments that would have been needed to cause that action.

The idea that morality aims at promoting useful sentiments can also be expressed as a type of Kantian ethics: act on those sentiments that one can will to be universal sentiments. However, we would have to recognize that, while there are sentiments people generally have reason to promote universally (such as aversions to lying, breaking promises, theft, vandalism, and assault), there are other sentiments that people generally have reason to promote – but not universally. These sentiments govern such things as what to wear, what to eat, where to live, whether to have children and how many children to have, and what profession to go into. The sentiments that people have reasons to promote universally govern the realm of moral obligation and prohibition. The sentiments that people have reason to promote – or at least no reason to disallow – in some people but not universally accounts for the category of non-obligatory permission.

Finally, this conception of morality could be presented as a type of virtue ethics, where a virtue is a sentiment people have reasons to promote universally and a vice is a sentiment people generally have reasons to discourage universally. Judith Thompson provides an account like this when she claims that, by a virtue, “What I have in mind is the idea that the fact of there being people who possess the virtues is good for us.”⁴ Though here, too, we have reason to mark off those sentiments that would be good for us if held universally and those sentiments that are good for us – or at least not bad for us - if they are held by some segment of the population, though not everybody. It is good for us that there are people who possess an interest in medicine, engineering, or teaching – but not necessarily good for us that everybody values one of these at the exclusion of the other two.

When we look at the usefulness of a sentiment – expressed in terms of the reasons that others have for promoting it universally using emotional conditioning – we also get a more externalist conception of morality. The sentiments that I have may be quite different from the sentiments that people generally

⁴ Judith Thompson, “The Right and the Good”, *The Journal of Philosophy*, Volume XCIV, No. 6 (June 1997), p. 282.

have reason to promote universally – which is the basis and the ground for their moral praise and condemnation.

We are accustomed to looking at morality through the eyes of the person who is making the moral judgment. However, there is another perspective to consider in understanding moral judgments – the perspective of the person being subject to moral praise or condemnation. From this point of view – from the point of view of the child being condemned for hitting her brother – the idea that one is being told that the wrongness of her actions depends crucially on her sentiments seems a bit problematic.