Abstract: According to the perceptual model, our knowledge of others' minds is a form of perceptual knowledge. We know, for example, that Jones is angry because we can literally see that he is. In this essay, I argue that mental states do not have the kind of distinctive looks that could sufficiently justify perceptual knowledge of others’ mentality. I present a puzzle that can arise with respect to mental states that I claim does not arise for non-mental properties like being an apple and argue that this is explained by the fact that the looks of non-mental properties adhere to a certain explanatory principle that does not hold for mental states. This shows, I argue, that, even if we think mental states do have looks, these cannot offer sufficient grounds for perceptual knowledge of others' minds. In the final section of the essay, I suggest an alternative way of thinking about our knowledge of others' minds and about the sorts of looks or appearances that might be associated with mental states.

Many people have the sense that we simply cannot know about the mental states of other individuals in the same way we know about ordinary material objects. Nevertheless, a growing number of philosophers have come to endorse the surprising thesis that our knowledge of others' minds is a form of perceptual knowledge. They claim, for example, that we know that Jones is angry because we can literally see that she is angry (e.g., Cassam, 2007; Dretske, 1973; McDowell, 1982; McNeill, forthcoming; Smith, 2015). A common reaction to this view is incredulity. H. H. Price, for example, describes the idea that we can have perceptual knowledge of another’s mind as ‘extremely paradoxical’ (1931, pg. 56). This may be because it seems fairly clear that mental states are not observable in the same way that colors, shapes, or sizes are. How, one might wonder, could we possibly see that another person is angry when anger is simply not visible? Advocates of the perceptual model respond to this sort of concern by noting that we regularly take ourselves to see much more than the low-level properties of material objects such as their colors, shapes, sizes, and spatial relations. Thus, we not only see the redness of the object on the table in front of us, but we
can also see the object to be an apple. Although the property of being an apple is not among the basic observational properties of the apple, it still seems to be something that we can know about on the basis of perception.¹ So, given that we can see something to be an apple, why, the perceptual theorist asks, could we not also see Jones to be angry?

¹ In this essay, I shall understand an ‘observational property’ to be a property of an object that is presented in perceptual experience or is perceptually manifest (cf. Langsam, 2000). Basic observational properties are the low-level properties of objects which are widely agree to be manifest in perceptual experience, including color, shape, size, visible texture, spatial arrangement of parts, etc. There is considerable debate in contemporary philosophy of perception as to whether any other properties of objects are also perceptually manifest. Thus, several philosophers who defend what is often called the ‘rich-content view’ have argued that natural kind properties and other high-level properties are presented in the contents of perceptual experience (Masrour, 2011; Siegel, 2006; cf. Logue, 2013). If this right then, in my sense, they would be observational properties, but they would be non-basic. One crucial difference between these and basic observational properties is that, as M. G. F. Martin notes, only with the latter do we have ‘a necessary coincidence between having the look of that property and having that property.’ (2010a, pg. 208; cf. Price, 2009) It is partly because of this, that Martin would deny that natural kinds are observational properties. However, for the purposes of this essay, I wish to remain neutral as to whether there are non-basic observational properties. This is because, as we will see in what follows, advocates of a perceptual model of our knowledge of others’ minds do not need to take a stand on this issue.
For various reasons, we might be reluctant to accept the notion that a mental state can be directly presented in a perceptual experience. But the perceptual theorist does not really need this to be true in order to defend the thesis that we are able to acquire perceptual knowledge of another person’s mind.² This is because the defender of a perceptual model could argue for the thesis that we are able to see facts involving another person’s mental states, such as the fact that someone is angry, without having to see her mental states, without anything mental having to be directly presented in perceptual experience. This sort of thing is not uncommon. For example, it seems like we can see that the cup of tea is hot without seeing its heat (Dretske, 1969). We see it in virtue of seeing the steam rising from the cup. Strictly speaking, the heat itself does not enter into the content of our visual experience; but that does not prohibit our knowledge that the tea is hot from being genuinely perceptual.

² Several advocates of the perceptual view nevertheless do defend the claim that mental properties are manifest in perceptual experience. Most prominently, the idea that we are perceptually acquainted with mental states is often proposed by theorists those who adopt some sort of ‘embodied perception model’ of our knowledge of others’ minds (e.g., Gallagher, 2008; Krueger and Overgaard, 2012; Overgaard, 2014; Smith, 2015; and Stout, 2010). Indeed, some of these embodied perception theorists seem to think that this is the only way one could possibly acquire perceptual knowledge of another person’s mind (cf. McNeill 2010). Others argue that the notion of a mental state being perceptually manifest is the only accurate characterization of the phenomenology of our experiences of other people (Gallagher, 2008; Krueger, 2012; Smith, 2015). It is not obvious to me that this last claim is true; however, the issue is outside the scope of this essay. I shall return to the embodied perception model in section 6.
Perceptual knowledge that $x$ is $F$ does not require $F$ itself to be perceptually manifest or figure directly in the content of one's experience. So even if mental states cannot be perceptually presented to a subject, this would not show that our knowledge of others' minds is not a form of perceptual knowledge.

Several philosophers have appealed to a conception of perceptual knowledge along these lines in order to explain our knowledge of others’ minds. Prominently among them, Quassim Cassam draws on Dretske’s theory of epistemic-seeing in order to specify the following four conditions for acquiring perceptual knowledge of another person’s mental states:

"The sense in which I see that the Bursar is angry is that

(i) I see him,
(ii) he is angry,
(iii) the conditions are such that he wouldn't look the way he looks now unless he was angry, and
(iv) believing that the conditions are like this I take him to be angry." (2007, pg. 163)

---

3 In order to apply these conditions to other cases, we need to generalize them as follows (cf. Dretske, 1969):

"The sense in which I see that the $x$ is $F$ is that

(i) I see $x$,
(ii) $x$ is $F$,
(iii) the conditions are such that $x$ wouldn't look the way it looks now unless $x$ was $F$, and
(iv) believing that the conditions are like this I take $x$ to be $F$. 
When these four conditions hold, Cassam claims, we are able to see facts concerning another person's mentality, such as the fact that the Bursar is angry, without having to directly perceive the person’s mental states, without anything mental having to be perceptually present.

On a view like Cassam’s, perceptual knowledge is possible because the relevant properties (e.g., being an apple, being hot, or being angry) are associated with a distinctive appearance or look. If there were no distinctive way that apples looked, we would not be able to visually discriminate or identify apples, which is to say that there would be no way for us to see an object to be an apple, even if the object were an apple and we were staring directly at it. Thus, to know on the basis of perception that a particular object is an apple, that object must manifest a look that is actually distinctive of apples and we must also have the capacity to visually recognize apples on the basis of this look (cf. Millar, 2011). It is worth noting that this is true even when it comes to basic observational properties like color, shape, and size. We are able to see that the apple is red only because being red has a distinctive look, a look that permits us to visually discriminate red from the other colors.

To say that some property has a distinctive look or appearance is to say that most of the objects in our actual environment which have that look in fact instantiate the property (Millar, 2014; cf. Martin, 2010a). So, being angry would have a distinctive look $L$ only if most of the people in our actual environment that look $L$ are in fact angry. This is important for Cassam’s condition (iii) because for it to be to be true of the Bursar that he would not look the way he does unless he were angry, the way the Bursar looks must be distinctive of anger, in the sense that most of the people who look that way are in fact angry. If the Bursar’s look were not distinctive of anger, if it was for instance a look typical of someone who is sad, then
it would not be true that the Bursar would not look the way he does unless he were angry, for it seems he could easily look the same way when sad.

So regardless of whether or not mental states can be directly perceived or can be the sorts of things that figure in the contents of perceptual experience, if we are to have any perceptual knowledge of another person's mental states, these too must have distinctive appearances or looks. If there were no distinctive way that it looked for a person to be angry or sad, there would be no way to visually discriminate angry people from others, which is just to say there would be no way to see that someone is angry.

Thus far we might think the perceptual model advocated by Cassam and others is in pretty good shape because mental states do seem to have distinctive looks. Smiling in a particular manner, for example, might seem to be the distinctive way that happy people look and wincing or grimacing might seem to be the distinctive way that a person in pain looks. It might therefore also seem extremely plausible that the distinctive looks which we naturally associate with mental states allow us to have perceptual knowledge of others’ minds. In this essay, however, I shall argue that this last impression is mistaken. More specifically, I shall argue that even if we assume that mental states do have distinctive looks they do not have the kind of looks that could justify perceptual knowledge of another person’s mind. This is not for the obvious reason that different people manifest mental states in idiosyncratic ways, which may well be true, but rather because the looks associated with mental states are mediated by our agency and therefore can be too easily dissociated from those mental states. This indicates, as I shall argue, that there is a kind of ontological gap between a mental state and its distinctive look, which is absent when it comes to non-mental properties and their distinctive looks. For this reason, the distinctive look of a mental state is simply not the sort of thing that could enable a person to perceptually know anything.
My argument against the perceptual model of our knowledge of others’ minds centers around a puzzle that arises with respect to the distinctive looks of mental states that I shall present in the following section. In section two, I shall argue that this puzzle does not arise for the looks of non-mental properties like being an apple because the distinctive looks of those properties adhere to an explanatory principle that does not hold for the distinctive looks of mental states. As we will see by the end of section two, this has important consequences for how we should conceive of the conditions required for perceptual knowledge. In section three, it will be important to distinguish the problem that I am raising for the perceptual model from a rather different one that concerns the possibility of pretense or acting. In drawing this distinction, I shall argue that the mere possibility of pretense does not present a problem for the perceptual model, and the reason why also has an important consequence for how we should think about perceptual knowledge. In section four, I shall draw on the explanatory principle presented in section two in order to fully develop the problem that I want to raise for the perceptual theorist. In broad strokes, my argument will be that the overall look a person manifests at a particular time cannot be sufficiently grounded in her mental states for a spectator to acquire perceptual knowledge of her mind. In sections five and six, I respond to two objections to this argument that might be offered on behalf of the perceptual theorist. In my replies to these objections, I shall suggest that our knowledge of others’ minds is normally based on a person’s expressive behavior. This naturally raises a question about the relationship between expressions and mental states. So in the final section of the essay, I shall present a proposal which claims that expressive behaviors are necessarily connected to mental states, and I shall argue that this relation explains how we are able to acquire knowledge of another person’s mind on the basis of her expressive behavior.
1. Looking and Acting

For the sake of argument, let’s suppose that mental states do have distinctive looks or appearances, which means that there is a way it looks for a person to exemplify a particular mental state. This makes it more plausible to think that we are able to visually discriminate individuals who are angry or sad from those that are not on the basis of how they look and thereby acquire perceptual knowledge of others' minds. Yet, this does not entail that we are able to acquire perceptual knowledge of others' minds. Several things might stand in our way.

We know that it is possible for an object to misleadingly manifest the distinctive look of something. A schmapple looks just like a real apple but it isn't one. What schmapples have in common with genuine apples is that they share the same overall look (they are visual duplicates). So being perceptually acquainted with a distinctively applish look does not guarantee that we are seeing an apple. Less exotic examples of this phenomenon involve mentality. We know that a person can look angry or sad without actually being so. They can be pretending to be sad or acting. Thus, with the possible exception of basic observational properties like color or shape, it is generally true that it is possible for an object to have the look of F without actually being F. If this possibility were widespread, if, for example, we were in a world containing a rather significant number of schmapples, we would not be able know that something is F on the basis of its looking F, which is to say that we would not be able to see that something is F. This would be for the simple reason that the look of F would then not be distinctive of the things that are F - it would, for example, be the look distinctive of a class of objects consisting of both schmapples and apples. So in order to see that x is F, x must exemplify some look L that is distinctive of only the things that are F, which is to say that it must be the case that most of the things around here that manifest L are in fact F (Millar, 2000; 2014). This should not be thought to mean that some objects that are not-F
could not also manifest the look of F on occasion, only that when they do, the look they
manifest would not be distinctive of things that are not-F – the look in question would remain
the distinctive look of things that are F.

We also know that \( x \) can be F without having the distinctive look of F. Some of the
apples we come across are very poor indeed; they don't have that red, round look of an apple.
Indeed, it would be extremely difficult to recognize these fruits to be apples on the basis of
the way they look. So being F in no way guarantees that an object has the distinctive look F.
Nevertheless, the majority of apples we come across do have the distinctive look of apples
and, more importantly, they look like apples because they are apples. There is normally a
strong connection between the distinctive applish look they manifest to spectators and the
property of being an apple. Typically, when \( x \) both is F and looks F, \( x \) looks F because it is F.\(^4\)

\(^4\) Joel Smith (2015) claims that this kind of connection is necessary for a look to be distinctive:
‘\( L \) is the distinctive look of \( F \): if and only if most actual things that look \( L \) do so because they
are F.’ (pg. 286) But I don’t see why this would be true. Plausibly, \( L \) is distinctive of the F’s
insofar as it allows a spectator to discriminate and identify them on the basis of the way they
look. So it seems irrelevant whether or not objects have \( L \) ‘because’ they are F. If I paint all
of the Gala apples with a dark blue stripe down the centre, a spectator will be able to visually
discriminate those apples from the others, and so I think the apples would have a distinctive
look (Smith’s view denies that the blue line is a distinctive look). However, it would not
follow from this that the spectator could see that the apples were Galas (for one thing she
may not know about these unusual conditions). I therefore prefer a weaker notion of what it
takes for \( L \) to be a distinctive look. As will become clear in what follows, it is partly for this
But although this last claim is usually true, there are exceptions. Consider the following case:

**Angry Patrick Stewart**: The magnificent actor Patrick Stewart has been cast to play Hamlet at the local theater. During each performance, there is a time at which he looks angry. As it happens, before last Tuesday's performance, Stewart got some very bad news and actually is angry during the performance. Stewart both is F and looks F, but it is not the case that he looks F because he is F.

Because he is acting, Patrick Stewart can both be angry and look angry, but it is not true that he looks angry because he is. We see his behavior; we see him intentionally manifesting the distinctive look of anger. After all, part of what makes him such a splendid actor is that he excels at looking angry on cue. Thus, it is not an accident that Patrick Stewart looks angry. Nevertheless, since the look of anger is disconnected from his actual anger, we cannot see that Patrick Stewart is angry. The way Patrick Stewart looks is not a manifestation or expression of his anger and so it cannot be said to reveal or present his anger to us in any way. And if Stewart's look does not reveal his anger to us, it can hardly be the basis on which we come to know that he is angry.⁵

---

⁵ This case does not depend on any of the unique artistic or performative features of the acting context. If you are concerned that it does, simply imagine it happening outside of the theater. For example, suppose that Patrick Stewart intentionally looks angry whenever he goes
One straightforward way to realize that we cannot see that Patrick Stewart is angry is to notice that Angry Patrick Stewart violates Cassam's third condition (iii) on perceptual knowledge, which states that the circumstances must be such that Stewart would not look the way he does unless he was angry. In our example, it is not true that Stewart would not look the way he does unless he was angry. Although he is in fact angry, he looks the same way every other time he plays Hamlet, most of which are times when he is not angry. It is plausibly for this reason that audience members do not typically believe that conditions are such that he would not look the way he does unless he was angry (which means that Cassam’s condition (iv) would also be violated). This is notable because a quick glance at Cassam's conditions might lead one to conclude that whenever condition (ii) is satisfied and someone is actually angry, condition (iii) could be violated only if the way that person looks is not really distinctive of anger (as in the case of schm apples). Angry Patrick Stewart shows that this is wrong. Someone can be angry and non-accidentally manifest the distinctive look of anger but we may nevertheless be unable to see that they are angry because they do not look angry because they are angry. Being F and non-accidentally having the distinctive look of F are not jointly sufficient for securing perceptual knowledge.

2. Looking the Way You Are

Is this sort of problem unique to others' minds? Defenders of the perceptual view may want to insist that puzzling cases like this are simply instances of more general difficulties facing perceptual knowledge. We must remember that the perceptual theorist is not claiming that to the market so that his adoring fans do not harass him, and suppose further that one day he also is actually angry, etc.
perceptual knowledge is not without problems, only that our knowledge of others' minds is perceptual. So even if we have some outstanding questions about perceptual knowledge, these do not necessarily undermine the perceptual model (cf. Dretske, 1973).

Along these lines, the following case has much in common with Angry Patrick Stewart:

**Village Fair**: I have agreed to enter a competition at my local village fair and have a bunch of wonky apples that do not look like apples. Since I have entered the fair however, and desperately want to win, I paint them in such a way that they look like apples. My wonky apples both are F and look F, but it is not the case that they look F because they are F.

My wonky apples will trick the judges only if I succeed in replicating the distinctive look of apples. So my mimicking the look of a genuine apple is motivated by my knowledge that my wonky fruits really are apples. Thus, there is one sense in which my wonky fruits look F because they are F - their actually being apples is causally relevant to the way they look after I paint them. Yet, even so, the look of my wonky apples does not depend solely on the property of being an apple. Rather, they look the way they do (like apples) only because something (me) has intervened to change their natural look. If we want to know why the wonky apples look like apples, there is a clear sense in which the answer 'because they are apples' is not a complete explanation. In this way, our impression that the wonky fruits do not look F because they are F seems to rest on a ‘non-causal’, or ‘not merely causal’, sense of ‘because’.

Can the judges in Village Fair see that my wonky fruits are apples? I think they cannot. As in Angry Patrick Stewart, the way my fruits look is not a direct manifestation of,
nor is it fixed by their being apples. Crucially, however, in this example Cassam's counterfactual condition (iii) is satisfied: the wonky fruits would not look the way they do unless they were apples. The competition is, after all, for best apples and it would be highly unlikely that I paint some other kind of fruit to look like apples in this context (if you are skeptical suppose that the judges eventually taste the fruit). This illustrates that Cassam's condition (iii) is too weak to secure perceptual knowledge.

In order to acquire perceptual knowledge that \( x \) is F, the relation between the distinctive look of F and an object's being F must be stronger than counterfactual dependence. If I am strongly disposed to causally intervene to change the overall look of my fruits only when they are apples, then it will be true that the wonky fruits at the fair would not look F, unless they were apples. Nevertheless, as with Angry Patrick Stewart, the fruits do not look F solely in virtue of being F. So it is not sufficient for the distinctive look of F to be reliably correlated with some property, which in this case happens to be F. Rather, in order to be able to see that \( x \) is F on the basis of how it looks, what must be true is that \( x \) looks F in virtue of its being F.

I therefore propose to revise Cassam’s condition (iii) for perceptual knowledge as follows:

(iii-r) the conditions are such that he looks the way he looks in virtue of his being angry.  

---

6 A comprehensive analysis of the 'in virtue of' relation is beyond the scope of this essay. But there are a couple of features that are worth highlighting. First, as I am using the term, the relation holds between facts: the fact that \( x \) looks \( L \) obtains in virtue of the fact that \( x \) is F. Second, this is a relation of determination in the sense that to say that one fact holds in virtue
Since this ‘in virtue of’ relation likely has some degree of modal strength, we can derive Cassam's version of condition (iii) from it (cf. Audi, 2012). The revised conditions for perceptual knowledge, now in generalized form, are as follows:

The sense in which I see that \( x \) is \( F \) is that

(i) I see \( x \),

(ii) \( x \) is \( F \),

(iii-r) the conditions are such that \( x \) looks the way it does in virtue of \( x \)'s being \( F \).

(iv) believing that the conditions are like this I take \( x \) to be \( F \).

In both Village Fair and Angry Patrick Stewart we have objects manifesting the distinctive look of some property but not in virtue of their exemplifying that property. In both cases, the instantiation of the property is dissociated from its distinctive look.

of another is to say that the latter brings about or is responsible for the former (Audi, 2012). So, the fact that \( x \) is \( F \) is what makes it the case that \( x \) looks \( L \) (this distinguishes the in virtue relation from the relation of necessary dependence that I appeal to in section 7, which is not a determination relation in this sense). Thirdly, the in virtue of relation is explanatory (Fine, 2001). Thus, to say that an individual manifests the look of anger in virtue of being angry is to say that his anger explains \( why \) he looks angry. Similarly, to say that an object looks like an apple in virtue of being an apple is to say that its being an apple explains \( why \) it looks the way it does. As should be clear from the Village Fair example, this is a complete explanation, unlike cases where a property is merely causally relevant to an explanation of an object’s appearance. For further discussion, see Audi (2012) and Rosen (2010).
Nevertheless, in both cases, the objects manifest the distinctive look of F non-accidentally. It is not a matter of luck that my wonky apples look like apples or that Patrick Stewart looks angry. Therefore, these cases illustrate that an object does not have to look F in virtue of being F in order for it to non-accidentally look F.

Despite having these points in common, there are two important differences between **Angry Patrick Stewart** and **Village Fair**. First, in **Village Fair**, the wonky apples look F only because something has causally intervened to change their natural look. Prior to my painting them, my wonky apples manifest the look $L(\text{wonky})$. So, in order for them to manifest $L(\text{apple})$, I had to do something to change $L(\text{wonky})$ to $L(\text{apple})$ and, importantly, the transition between these two looks could not be accidental. In **Village Fair**, this anti-luck condition is met because (a) I intend to change $L(\text{wonky})$ to $L(\text{apple})$ and (b) I know that my

---

7 What if my fruits manifest $L(\text{apple})$ but I nevertheless also paint them to have $L(\text{apple})$? Presumably, in that case, the way they look would not be accidental, but it seems that they would no longer have $L(\text{apple})$ in virtue of being apples. I think there are a couple of different ways to fill out the details of this case. We might think of it as a case of apple make-up, in which my painting the apples brings out the natural look of the apples, rather than changing it. In that case, it would be plausible to maintain that the apples have $L(\text{apple})$ (at least partially) in virtue of being apples. But we might instead imagine the case in such a way that I deliberately paint the apples in order to alter them to have $L(\text{apple})$, even though this look would be of the same type. In that sort of case, the appropriate in virtue of relation would plausibly not hold but, crucially, the details of the case would have to be spelled out in such a way that my action seemed like it was aimed at changing the overall look of the apples. If all I did was paint them with a shiny gloss, the case would be more like a kind of apple make-up.
fruits are apples. But we could also imagine some sort of mechanistic intervention achieving the same result as long as it too had the aim or goal of changing $L(wonky)$ to $L(apple)$. The point is that a material object can non-accidentally look F, but not in virtue of being F, only if some external thing has causally intervened in order to change $x$'s overall look to the look of F. Second, and relatedly, in cases where there is a counterfactual dependence between $x$'s looking F and $x$'s being F, as there is in Village Fair, whatever causally intervenes to change $x$'s look must be sensitive to the fact that $x$ is F. Thus, in Village Fair, it is because I know that the wonky fruits are apples that they would not look like apples unless they were apples. Without such knowledge, I could still change the overall look of my apples to $L(apples)$, but the look would not counterfactually depend on their being apples.

Neither of these features seems to be present in Angry Patrick Stewart. Stewart looks angry because he intentionally behaves in a particular way. He intends to exhibit the distinctive look of anger in the course of his performance and is generally successful in acting

---

8 What about an object which has a variable overall look because of something internal? Anil Gomes posed the following case to me. Suppose there is a certain type of apple that has a funny genetic mutation such that when one of these apples is placed in sunlight it has one of ten different looks, only one of which is the distinctive look of apples. Does this fruit look like an apple in virtue of being an apple? I think it depends on whether we think of the genetic mutation as part of the apple. If it is, then it is plausible that the overall look the apples manifests, though variable, is in virtue of its being a specific (unusual) type of apple. If not, then the mutation looks like a sort of causal intervention. It also matters whether or not the variability is random, because the look needs to be non-accidental.
precisely the way in intends. This is sufficient to ensure that the way he looks is not an accident. So, even in cases where he actually is angry, there is no reason to suppose that his observable behavior is determined by his underlying mental state rather than simply being the result of his intentionally acting a certain way. There is no reason to think that Stewart needs to be sensitive to, or aware of, his actual mental states in order to non-accidentally look angry.

More importantly, unlike material objects, it does not seem that Stewart’s intentional action needs to change or alter his overall look. We have been supposing that Patrick Stewart happens to be angry prior to his performance, so he may very well look angry before the performance without being aware of it. At that moment, it might be plausible to think that Stewart looks angry in virtue of being angry. Nevertheless, once he begins his performance, Stewart looks angry in virtue of his agency, not his anger. So although Stewart must intend to display the distinctive look of anger, he can plausibly do this without intending to alter or change anything about the overall way he looks. Stewart looks angry, but unlike Village Fair, it is not necessary that he intervene on any prior observable aspect of his behavior. This is because he does not need to intentionally target his appearance or aim to change the way he looks. Generally, it is possible for a person to transition from non-intentionally $\Phi$-ing to intentionally $\Phi$-ing without there being any difference at all in the way she looks.

So, in cases like Angry Patrick Stewart, it seems like the individual exemplifying F is able to dissociate the property of being F from its distinctive look without intervening on any aspect of the overall way she looks. This sort of thing does not seem possible in the case of a material object. There is no way for a particular object to be an apple but dissociate the distinctive look of an apple from that property. The only way that such an object could both, be an apple and non-accidentally look like an apple, but not in virtue of being an apple,
involve circumstances like those in *Village Fair*, in which something external causally intervenes to change the look of the object.

I therefore propose that the following principle holds for non-mental properties:

\[(\text{Virtuous Look}): \text{For any individual } x \text{ and property } F, \text{ if } x \text{ both is } F \text{ and non-accidentally looks } F, \text{ then, absent a causal intervention to change } x's \text{ look, } x \text{ looks } F \text{ in virtue of } x \text{ being } F.\]

As **Angry Patrick Stewart** demonstrates, \((\text{Virtuous Look})\) is not true for mental states. Because mental states are properties of agents, their subject can intentionally dissociate them from their looks even though nothing external to the agent intervenes with the aim of changing that look.

### 3. The Possibility of Acting

Recall that Cassam's fourth condition on perceptual knowledge (iv) describes a modal belief that a spectator must have in order to acquire perceptual knowledge. Specifically, it states: ‘believing that the conditions are like this I take \(x\) to be \(F\).’ It might seem like the problem posed by **Angry Patrick Stewart** has something to do with this condition. That is, it might seem like the case shows that a spectator cannot rule out the nearby possibility that a person is acting or pretending, and so cannot justifiably have the required modal belief specified by Cassam’s condition (iv). This is because it might seem like we must rely on Patrick Stewart’s testimony before we can have a justified belief about whether he is angry in virtue of his anger. In that case, however, there would be a kind epistemic dependence on the subject of a mental state that doesn't exist in the case of material objects.
To see this more clearly, suppose that the judges in Village Fair are in a situation in which the fruits in front of them are actually apples that naturally have the distinctive look of apples. But these judges mistakenly believe this look might be the result of some nefarious intervention. Perhaps they have some reason to think that competitors at the fair regularly make cosmetic alterations to their fruits. From their perspective, it might be the case that the fruits look the way they do because of intervention and so, in that context, they would reasonably believe (falsely) that condition (iii-r) had been violated: they would reasonably believe that the apples do not look the way they do in virtue of being apples. One might think that the judges in this scenario simply could not see that the fruits are apples, precisely because they believe that the look of the fruit is determined by some external causal intervention (cf. Millar, 2011).

But there is a more plausible way to think about the epistemic situation of these judges. Even if they did explicitly believe that Village Fair is a nearby possibility, they normally could take steps to determine whether or not it was actually the case. They could, for example, examine the facade of the apples more carefully, or scratch the surface to discern whether any paint was present. If things seemed really bad, they could even slice the fruits open or bite into them. Because the judges are usually in a position to rule out the possibility of external causal interventions, they are usually in a position to know that the fruits are apples based on the way they look. They are therefore in a position to know that the overall look of the fruits satisfies condition (iii-r) – which is to say that given the actual conditions, the apples look the way they do in virtue of the fact that they are apples. So, regardless of what a spectator happens to be thinking, in a good case it seems that she is in a position to rule out the nearby possibility that the look of F has been dissociated from being F, and so she is able to see that something is F.
This demonstrates how Cassam’s version of condition (iv) is slightly too strong. A spectator does not need to actually believe that condition (iii-r) holds (does not actually need to believe that conditions are such that the object in question looks F in virtue of being F) in order to see that an object is F. Rather, she need only have sufficient justification for this belief available to her. If there is adequate justification within her epistemic reach but she happens to mistakenly think condition (iii-r) is violated, what this shows is not that she fails to see that \( x \) is F, but that she does not know that she is seeing this (Williamson, 2000). So even if the judges mistakenly believe there is a significant risk of Village Fair, they can still see that my fruits are apples, provided that they are in a position to rule out the possibility that the look of being an apple has been dissociated from the apples. They simply fail to know that they are seeing this.

We should therefore revise condition (iv) for perceptual knowledge, as follows:

\[(iv\text{-}r): \text{I take } x \text{ to be F and am in a position to justifiably believe that conditions (i) – (iii-r) are satisfied.}\]

When it comes to acquiring perceptual knowledge of some fact on the basis of the way something looks, a spectator only needs to be in a position to rule out epistemic possibilities that are incompatible with her seeing that fact. In favorable epistemic circumstances, this will typically be rather easy. A spectator is usually in a position to look more carefully, or from a different side or angle, or pay closer attention to what she is looking at in order to rule out certain epistemic possibilities. These are basic skills associated with the capacity to recognize \( x \) to be F on the basis of the way it looks. Importantly, this does not mean that a spectator who is actually in a good case must be constantly double-checking the
deliverances of her senses. In good cases, \( x \) looks F in virtue of being F, and spectator can easily satisfy (iv-r).

Suitably revised, the four conditions for perceptual knowledge are as follows:

The sense in which I see that \( x \) is F is that

(i) I see \( x \),

(ii) \( x \) is F,

(iii-r) the conditions are such that \( x \) looks the way it does in virtue of \( x \)'s being F.

(iv-r) I take \( x \) to be F and am in a position to justifiably believe conditions (i) – (iii-r) are satisfied.

Nevertheless, even with these revised conditions, someone might still think that, when it comes to another person’s mind, a spectator is stuck relying on the agent’s assurances before she can be in a position to rule out the relevant alternative possibilities. Thus, someone might think that we can get ourselves into a position to rule out the possibility that a person is acting only by asking her whether or not she is. But asking is not sufficient because we would not acquire a reason to rule out the possibility that Stewart is intentionally manifesting the look of anger unless Stewart cooperates and informs us that he is not doing this. He may not do so. Perhaps what Stewart tells us about his intentions is also part of his performance or perhaps he feels that we just do not deserve to know the truth about what he is doing (Williams, 2000). Moreover, if he were to lack knowledge of what he is intentionally doing, perhaps because he was drugged, tired, or distracted, then his answer, even if sincere, would not help us rule out any relevant alternatives. Regardless, if we are in a position to know that Stewart is angry only because we have learned from him that he is not acting or dissembling, then our knowledge that he is angry would not be perceptual. It would rest (at least partially) on our knowledge of what he is intentionally doing.
A proponent of the perceptual view can reply to this concern by arguing that we do not need to know that Stewart is not intentionally manifesting the distinctive look of anger. Rather, we are simply entitled to assume that cases like *Angry Patrick Stewart* are not nearby possibilities, perhaps as a background condition on interpersonal conversation or social interaction. Both *Village Fair* and *Angry Patrick Stewart* are highly unusual cases in which the way an object looks is misleading. They are cases in which a spectator's epistemic circumstances are not as good as they could be. However, just because we cannot see that \( x \) is F in some bad case doesn't mean we can't see that \( x \) is F in good cases. Thus, one way to defend the perceptual view is by arguing that bad cases are rather distant possibilities and can safely be dismissed when our epistemic circumstances are favorable. The basic idea behind this line of response is that in good conditions, which are perfectly ordinary, perceptual experiences give us direct access to facts. In McDowell's words, in good cases, we are presented with 'the fact itself made manifest' (1982; pg. 214).

So if we are actually in a good case, we can legitimately ignore the possibility that our epistemic circumstances might be misleading. This is why in any ordinary case the judges at a fair can obviously see that the fruits are apples without having to carefully consider a wide variety of skeptical scenarios. In good cases there are no external interventions, so (*Virtuous Look*) implies that apples look the way they do (viz., like apples) in virtue of their being apples: their overall look is suitably connected to the kind of fruit they are. *Village Fair* is a possible case in which some apples would have the same look but that look would be insufficient for seeing that they are apples. However, the fact that the look of F cannot ground a perceptual presentation of some fact in some very distant non-actual world doesn't mean that it cannot do so in the actual one. The mere logical possibility of *Village Fair* does
not stand in the way of judges perceptually taking in the fact that fruits are apples in ordinary contexts.

Similarly, even if \((\text{Virtuous Look})\) does not hold for mental states, in a good case an individual who is angry will not be intentionally dissociating the way she looks from her underlying mental state, which is to say that cases like \textbf{Angry Patrick Stewart} are normally not nearby epistemic possibilities. Indeed, when we are in favorable epistemic conditions, the person we observe will not be acting, pretending, or even trying to manifest a particular look. So, although \((\text{Virtuous Look})\) may show us that there are ways in which a mental state can be dissociated from its distinctive look which result from a person’s agency, the perceptual theorist can insist that we can ignore these misleading possibilities when our epistemic circumstances are favorable. They are, at any rate, not the most natural conditions in which we encounter and interact with others. So they do not raise any obvious threat to the perceptual model, especially on a conception of perception as openness to facts.\(^9\) As Ryle noted, ‘the menace of universal shamming is an empty menace.’ (1949, pg. 173)

In the following section, I shall suggest that the problem for the perceptual theorist that is raised by the cases presented in the previous sections is not that pretense or acting somehow present additional skeptical possibilities, but that these cases strongly suggest there is an ontological gap between a person’s mental state and the overall look the person displays, which arises because of the person’s agency.\(^{10}\) More specifically, I shall argue that because

\(^9\) As Gomes (2011) notes, this is why the mere possibility of pretending or acting does not pick out a special kind of philosophical problem for our knowledge of other minds.

\(^{10}\) This is contrary to the view of appearances presented in McDowell (1982). But it is worth noting that the reason there is an ontological gap between the appearances of mental states
(Virtuous Look) does not hold for mental states it difficult to see how the overall look manifested by an individual could really be in virtue of her being in a particular mental state. It is therefore unclear how the knowledge we have of another person’s mentality could be a form of perceptual knowledge.

4. The Basis of a Look

Village Fair shows us that we can change the overall look of an object without changing its kind, so we know that overall look of an object cannot supervene on its kind. As we have seen, whether or not a look \( L \) is distinctive of some class of objects (e.g., the ones that are F) will depend on external or relational facts like whether or not there are a significant number of schmapples nearby. But changing these relational facts does not change the overall look an object manifests. If I were to take my bright red Gala apple to a schmapple farm in some other possible world, it would have the exact same look, only this look would now be distinctive of schmapples.

It is of course possible to change the overall look of a material object, as we have also seen in the preceding sections. But it is very difficult to imagine how to do so without in some way altering the object’s basic observational properties, which are the basic properties directly presented in perceptual experience, such as the object’s color, shape, size, and motion. When faced with my wonky apples in Village Fair, the only way I have a chance at winning the blue ribbon is to either change or mask their color and shape, to alter their basic observational properties or to at least make it appear that I have done so. This strongly does not rest on accepting the ‘highest common factor’ conception of appearances that McDowell wants to resist.
suggests that the overall look of an object supervenes on some or all of that object’s basic observational properties, which explains why a sufficient causal intervention that specifically targets those properties can disrupt the natural connection between an object’s overall look and the kind of thing it is.

If it is true that the overall look of an object $x$ supervenes on $x$’s basic observational properties, it raises an important question.\(^\text{11}\) In what sense could it be true that $x$ looks $L$ in virtue of $\text{being } F$?\(^\text{12}\) Since we know that $L$ does not supervene on $x$’s being $F$ (from Village Fair), why should we think $L$ is ever the look of $Fness$? What precisely is the connection between $L$ and $x$’s being $F$ that would warrant the description that $L$ is the look of $F$?

These are important questions in part because condition (iii-r) for perceptual knowledge states that we can know that $x$ is $F$ on the basis of $L$, only if $x$ has $L$ in virtue of being $F$. However, if $L$ supervenes on some of $x$’s observational properties, then this implies that perceptual knowledge requires a rather strong connection between those properties and

\(^{11}\text{It is possible that the reason the overall look of an object supervenes on the object’s basic observational properties is because these are identical. For development of this sort of view, see Martin, 2010a.}\)

\(^{12}\text{I think there is a plausible anti-realist way of addressing this question. On the sort of view I have in mind, all it is for } x \text{ to have the distinctive look of } F \text{ is for us to be disposed to visually discriminate the } F \text{’s on the basis of their basic observational properties. On the sort of view I am imagining, there would be no deeper ‘in virtue of’ explanation of } x \text{’s look in terms of it being } F \text{. So the anti-realist would presumably resist condition (iii-r) for perceptual knowledge, perhaps in favour of a claim about what observers would visually discriminate in certain conditions. Unfortunately, I lack space to elaborate on this approach in this essay.}\)
$x$’s underlying nature (i.e., being F). Minimally, it would seem that $x$’s being F would have to determine or ground $x$’s basic observational properties, or at least those upon which $x$’s overall look supervenes, such that a complete explanation of $x$’s color, shape, or size could be given by reference to $x$’s being F. This would preserve the right sort of explanatory in-virtue relation between $x$’s manifesting $L$ and $x$’s being F, as is required by condition (iii-t). That is, we could maintain that $x$ manifests $L$ in virtue of being F, if and only if $x$’s basic observational properties, upon which $L$ supervenes, were determined by its being F.

As (*Virtuous Look*) suggests, we do find this sort of determination relation holding for the looks of material objects in cases where nothing external has intervened to change an object’s basic observational properties. Exemplifying a specific natural kind does seem like it fixes an object’s size, shape, and color, which is to say that it explains why the object has the basic observational properties that it does. In other words, the object’s basic observational properties are grounded in the kind of thing the object is, which is also to say that the object exemplifies those properties in virtue of the kind of thing it is.\(^\text{13}\) For example, as Schaffer

---

13 Much of the philosophical literature on grounding conceives of it as a necessary relation, such that if $x$ grounds $y$, then, necessarily, if $x$ then $y$. However, if the basic observational properties of an object are grounded in the object’s kind, then the relation is not necessary. This is for the simple reason that we can intervene directly on the observational properties to change them without altering an object’s kind. Option one (which is my preference): we could deny that grounding is a necessary relation (Leuenberger, 2014; Litland, forthcoming). Option two: we could say that strictly speaking the basic observational properties of an object are grounded in the conjunction of the object’s kind and the absence of external interventions
notes, ‘the red aspect of the apple exists and the round aspect of the apple exists because of how the apple is.’\textsuperscript{14} (forthcoming, p. 5) It is the nature of being an apple that explains why this particular apple is red and round. This is why we can say, as \textit{(Virtuous Look)} claims, that without any external intervention an apple looks the way it does in virtue of being an apple. Being an apple naturally determines the basic observational properties of the object, like its redness and roundness, upon which the overall look of the object supervenes.

I think we are now in a position to see the real difficulty generated by the fact that \textit{(Virtuous Look)} fails to hold for mental states. It suggests that mental states do not fully determine the basic observational properties of a person. Although the basic observational properties of a person are more complex than those of simple material objects, they are also things like the person’s color, shape, size and motion. These are the properties that constitute an individual’s observable behavior, which everyone agrees is directly presented in perceptual experience. If mental states did naturally determine the basic observational properties of a person, then we would expect that, contrary to \textit{(Virtuous Look)}, in cases where nothing external intervenes to change a person’s basic observational properties, a complete explanation of why someone displays a particular look could be given in terms of that person’s mental state (since the mental state would fully determine the person’s basic observational properties). But we do not find this to be the case. Fixing a person’s mental

---

\textsuperscript{14} This does not imply that being apple determines only one set of observational properties. There are green apples and red apples but plausibly both are determined by, as Schaffer says, ‘how the apple is’.
state does not seem to fix the person’s observational properties. If that is right, however, then since the overall look of a person plausibly supervenes on those properties, it is hard to see how a person could manifest an overall look \textit{in virtue of} being in a particular mental state, rather than, for instance, \textit{in virtue of} intentionally displaying a certain pattern of observable behavior.

Consider anger. Being angry does not appear to determine the observational properties exemplified by an individual. Rather, it seems that someone can be angry but not manifest any observable behavioral response. Moreover, an angry person can typically alter her behavioral response on different occasions, and it may look very different each time.\textsuperscript{15} It is not as if there is a specific set of basic observational properties that Patrick Stewart must manifest on the day he happens to be angry. But since Patrick Stewart’s overall look supervenes on these properties, it is difficult to understand precisely how the overall look Patrick Stewart manifests to a spectator on a particular occasion is really \textit{in virtue of} his underlying mental state, rather than in virtue of the behavior he intentionally displays.

I think this is generally true. People often dissemble in social gatherings. We sometimes scowl intentionally in order to put someone off or intimidate them. During a lecture, we might try to look interested in a topic out of fear or respect; and certain forms of cognitive behavioral therapy are based partially on the idea that an individual can eventually affect the way she feels by intentionally manifesting the look of happiness. Indeed, what the

\textsuperscript{15} Note that the issue is not one of contextual variability, which could be addressed by a proponent of the perceptual model (see Dretske, 1969). Rather, the issue is that mental states do not seem to naturally determine any specific way of looking. The overall look a person manifests is always under her control and mediated by her agency.
earlier discussion suggests is that a person’s overall look is always mediated by her agency, even in cases where the subject is actually exemplifying the mental state in question. Perhaps this is partly constitutive of being an intentional agent.

5. Ways We Can’t Help Looking

Naturally, if there were some mental states that did naturally determine specific observable behavioral responses, responses which were not under an agent’s control and could not be helped, for example if one could not help looking a particular way whenever one was truly happy or in extreme pain, then it would be more plausible to think that a person looked L in virtue of her underlying mental state. This is because it would be easy to see how that mental state alone was what determined the person’s basic observational properties, upon which the person’s overall look L supervenes. Therefore, someone who wishes to defend a perceptual view might argue that there are some mental states that do determine observable behaviors, such that no matter how much a person might try to repress or hide them, some behavioral responses simply fall outside the scope of human agency, even for the very best of actors.

Emotional states like extreme anger or intense happiness come to mind as the sort of mental states that would naturally determine certain patterns of observable behavior. Indeed, it might seem that much of the experimental work done by Paul Ekman and colleagues over a number years showing the existence of ‘micro-expressions’ demonstrates that this is the case. ‘Micro-expressions’ are extremely brief facial expressions that are displayed in experimental contexts where a person is asked to try to conceal, hide, or repress her true feelings (Ekman, 2003; Ekman and O’Sullivan, 1991; Hurley, et. al. 2014; Yan, et. al. 2013). For example, if someone who is actually experiencing strong negative emotions tries to conceal them by deliberately smiling, they will display facial muscular activity that is associated with negative
affect (Ekman, et. al. 1988). There appears to be plenty of evidence for the existence of micro-expressions for a range of different emotions (Ekman, 2003; Yan, et. al. 2013). Thus, one might conclude that when a person experiences a particular emotional state, it does determine a certain pattern of observable behavior which is beyond the agent’s control, namely the facial micro-expressions characteristic of the emotion. If this is right, then a defender of the perceptual model of our knowledge of others’ minds could claim that these clearly visible micro-expressions are naturally determined by mental states, and perhaps even constitute the distinctive looks of certain emotions.

However, it seems to me that the evidence we currently have concerning facial micro-expressions does not support this conclusion. First, micro-expressions are displayed only in experimental conditions where an individual is asked to mask or conceal her emotions. That is to say, they are only displayed by individuals who are engaged in some sort of deliberate act of deception. As Yan and colleagues note, this introduces a significant amount of ‘noise’ into the experimental conditions, which makes it hard to rule out the possibility that a subject’s facial muscular movements are ‘just indicating the action of repression.’ (Yan, et. al. 2013, pg. 228) There is therefore no way to really tell whether a specific micro-expression is caused by an agent’s attempt to conceal her emotion, rather than by the emotion itself. Thus, it seems premature to conclude that micro-expressions are determined by emotional states. Secondly, it seems that several of the micro-expressions studied by Ekman and colleagues could not possibly ground the distinctive look of an emotional state because the majority of people who they studied do not actually display them (Ekman, 2003; Ekman, et. al. 1988). So even if there were some look L that did supervene on these micro-expressions, it would not be shared by a sufficient number of people to be the distinctive look of an emotion. Finally, even for micro-expressions that are more prevalent, it does not seem that they are really
determined by a person’s emotional state. In all of the experimental work on micro-expressions, there are individuals who do not display micro-expressions, which suggest that the display of a micro-expression, though involuntary, is not really outside of an agent’s control. Indeed, Ekman himself seems to agree with this, claiming that ‘there is no single source within demeanor that is completely trustworthy, impervious to efforts to disguise; nor is there a source that should be ignored because it is completely untrustworthy’ (Ekmann, 2003, pg. 217). I therefore think there is no reason to believe that any of the observable behavior we might associate with mental states cannot be controlled by agents.

Other empirical research seems to confirm this. For a long time it was thought that a person could not deliberately produce or repress a Duchenne smile because it required the use of certain facial muscles, specifically the orbicularis oculi muscle which surrounds the eye, that were not under a person’s control. Traditionally, it was thought than only spontaneous smiles that were caused by genuine happiness (i.e., Duchenne smiles) involved these muscles. However, recent studies have shown this to be incorrect. A significant number of people can voluntarily produce a Duchenne smile, even when they are not experiencing any positive affect (Gunnery, et. al., 2013; Krumhuber and Manstead, 2009). There is some evidence that subjects are a bit less successful at voluntarily producing genuine expressions of pain (Craig, et. al. 1991; Hill and Craig, 2002; Larochette, et. al. 2006). Yet, Craig and colleagues report that subjects who were given instructions to ‘fake or pose a facial display of pain’ produced responses that ‘tended not to differ qualitatively’ from those of subjects who were actually in pain (1991, pg. 168). Although, this response tends to be more exaggerated, especially in children (Larochette, et. al. 2006), it does not seem to affect the look of a subject in a way to which spectators are sensitive. It therefore seems to me that experimental psychology and cognitive science have not produced any unambiguous evidence indicating that certain looks
could only be grounded in the existence of a mental state, rather than in an agent’s voluntary behavior.

It may be worth keeping in mind that the minimal sort of empirical evidence that a perceptual theorist would need to support her view is evidence showing that the occurrence of a particular mental state was reliably correlated with a certain pattern of observable behavior. And there may be other relevant experimental work that I have failed to consider in this essay, or, indeed, future work that might support such a conclusion. ¹⁶ Since I lack space

¹⁶ Some embodied perception theorists claim that there is evidence that certain behaviors are necessary for experiencing certain emotions (and so they would be always be correlated with the occurrence of that emotion). For example, Joel Krueger and Soren Overgaard appeal to a number of studies indicating that individuals who mimic expressive behaviors tend to feel corresponding emotions (Duclos and Laird, 2001; Laird and Bresler, 1992; Niedenthal, 2007). In addition, they appeal to research on Moebius syndrome, a type of congenital facial paralysis, which appears to show the converse is true: namely, individuals who are unable to manifest expressive behaviors report having diminished emotional experiences (Cole, 1997; Krueger and Michael, 2012). They also appeal to work which suggests that inhibiting expressive behaviors impedes performance on cognitive tasks that require processing emotion specific information (Niedenthal, 2007; but see Bogart and Matsumoto, 2010 for counterevidence). However, it does not seem to me that this evidence does demonstrate the existence of the sort of correlation a perceptual theorist would need. The work of Laird and colleagues on facial mimicry indicates that there is significant variation in how much expressive behavior affects emotional experience, including several subjects ‘whose facial expressions do not affect their feelings’ (Laird, et. al. 1994 pg. 233; cf. Laird and Bresler, 1990;
to rule out all of the potentially relevant experimental work, I would like to register one additional reply to this line of objection.

   Even if it turns out to be true that some observational properties are correlated with specific types of mental states, this would not show that we generally have perceptual knowledge of others' minds, only that we could perceive someone to be in a particular mental state in those cases where the distinctive look associated with that state is not something a person can control. The resulting perceptual view would be severely restricted. It would not, for example, be applicable to any cognitive states such as belief or desire, which even perceptual theorists acknowledge lack a distinctive look (Green, 2007; McNeill, forthcoming).

   There just is no way it looks to believe that London is a wonderful place to live. So even if we were able to acquire perceptual knowledge of some of a person’s emotions, this would be only a partial explanation of our knowledge of others’ minds. Moreover, it would be an explanation that needed supplementation with some account of how we know about the broad range of mental states that fall outside the scope of perception. This may be acceptable Duncan and Laird, 1980). So even if exogenously produced expressive behaviors cause a related emotional experience in some individuals, this would not show that experiencing the emotion requires the expressive behavior. The anecdotal evidence from subjects with Moebius syndrome is similarly inconclusive. Although some subjects do report having either an overly intellectual or somewhat diminished experiences of their emotions, they appear to acknowledge that the emotions are nevertheless present (cf. Cole, 1997). Additionally, some subjects report being primarily concerned with ‘being out of control with emotions’ and so clearly do not report lacking them (quoted in Cole, 1997).
to some perceptual theorists but it seems to me that a unified account of our knowledge of others’ minds would offer a simpler and more plausible picture.

Nothing I have said in this section denies that someone who is overjoyed or in extreme pain is often easy to spot. Nor am I denying that many people, especially small children, have a hard time concealing certain observable behavioral responses. However, the fact that it is often quite easy for us to tell when someone is in extreme pain or very happy does not entail that our way of knowing this is perceptual.

6. What We Say and Embodied Perception

Advocates of the perceptual model often emphasize that it accords with our ordinary ways of speaking. For example, Mitchell Green claims, "we often say such things as that we could see the anxiety on someone’s face, feel the trepidation in her handshake, and hear the exuberance in her voice." (2010, pg. 45; cf. Green 2007) And, according to McDowell, it is a piece of commonsense that 'one can literally perceive, in another person's facial expression or his behavior, that he is in pain." (1978, pg. 305; cf. Dretske, 1973) If, as I have argued, there are problems with the perceptual model of our knowledge of others' minds, how can we make sense of these locutions? How are we to understand our commonsense talk that seems to refer to our perceiving another person's mental states?

Notice that the locutions Green and McDowell are primarily concerned with have the form of 'perceiving the x in y’, as in 'seeing the sadness in her eyes'. This is peculiar since

---

17 One easily finds similar locutions in the embodied perception literature. For instance, Gallagher and Varga claim that we perceive emotions and intentional states directly 'in the bodily movement, expressions, and actions of others.' (2014, pg. 187, cf. Gallagher, 2008).
these are not reports of directly seeing a mental state but of seeing it in some aspect or other of a person's behavior or anatomy. Generally, however, this syntactic structure does not suggest that the 'x' one sees 'in y' is itself seen. Such locutions therefore do not typically commit a speaker to even the existence of 'x' let alone to perceptual acquaintance with it. For example, when I report that I 'see Granny Smith in the photograph' or 'see the Battle of Waterloo in the painting,' nothing I have said implies that I am perceptually acquainted with either Granny Smith or the Battle of Waterloo, both of which may not exist. Rather, a more plausible analysis of such statements is that they are reports in which a speaker claims to be perceptually acquainted with y as a representation of x (by way of implicit contrast with some other description of y). I might report that I see Granny Smith in the photograph in a context in which the photograph is old, damaged, or otherwise of poor quality - a context in which it isn't obvious that the photograph does represent Granny Smith. I do not mean to suggest that a person's facial expressions are representations of her emotions in the way that photographs represent only that reports of 'seeing the sadness in her eyes' does not imply that the speaker is perceptually acquainted with anyone's sadness (Martin, 2010b).

Nevertheless, there are other locutions that might seem to directly refer to the looks or appearances of mental states. People often say things like 'John looks angry' or 'Sally looks sad', neither of which have the 'sees x in y' structure. How can we make sense of these utterances, unless mental states have distinctive looks? Throughout this essay, we have been

---

18 On some accounts, I do see Granny Smith herself when I see her in a photograph. I cannot discuss this sort of view in this essay but will only point out that it is substantive and controversial view about the nature of photographic representation and not entailed by a speaker's reports.
assuming that mental states do have distinctive looks. So if that assumption is correct, then these sorts of ‘looks’ locutions could refer directly to the distinctive looks or appearances of mental states. As we have seen, however, if mental states do have distinctive looks this does not mean that we have perceptual knowledge of others’ minds.

For this very reason, however, it is quite hard to see why we should bother holding onto the assumption that mental states have distinctive looks. Indeed, for many mental states, it seems clear that there is no distinctive look displayed by most of the people who exemplify them. I therefore think we might do just as well without the assumption that mental states have distinctive looks and shall drop it for the remainder of this essay. But then we need some alternative way to understand locutions that seem to refer to the looks of mental states.

I propose that these locutions are elliptical and that rather than referring to the look or appearance of a mental state, they refer to the look or appearance of an expression of a mental state. For example, the utterance 'John looks angry' should be interpreted as 'John looks like someone who is expressing anger' or ‘John’s is behaving in a way that looks like an expression of anger’. Thus, the appearance or look to which these sorts of statements refer is not a property of John’s anger but of his behavioral expression of anger. That is to say, there is a way that John looks, $L$ and $L$ is identical to, or relevantly similar to, the distinctive way that an expression of anger looks (cf. Martin, 2010a). Unlike mental states, which many feel are simply imperceptible, it is widely agree that expressive behavior is perceptible. So it is naturally the sort of thing that would manifest an appearance or look and could therefore easily function as the reference of locutions such as 'x looks sad'.

My suggestion then is that we are directly acquainted with the looks or appearances of expressions of mental states and we opaquely refer to these with statements like ‘John looks angry’. One point in favor of this proposal is that it allows us to easily explain our reaction to
**Angry Patrick Stewart.** Recall that Stewart is angry and is behaving in a way that manifests the distinctive look of an expression of anger. Yet, by hypothesis, Stewart is not expressing anger; he is merely acting so as to manifest the distinctive look of such an expression. Thus, the case is exactly analogous to the case of schmapples, the foreign fruits that look just like apples. Stewart looks like he is expressing anger but he is not, just as the schmapples look like apples, but are not. Stewart's behavior therefore has the same overall look as a genuine expression of anger, just as a schmapple has the same overall look as a genuine apple. It is just that in both cases the look is misleading, which is why either one can easily fool us. If this is right, then it should be clear why we cannot see that Stewart is expressing anger (since he is not expressing anger, condition (ii) for acquiring perceptual knowledge is not even satisfied).

Appealing to some notion of ‘expression’ is familiar in philosophical discussions of knowledge of others’ minds (e.g., Austin, 1946; McDowell, 1982). But doing so raises an obvious question. How do expressions of mental states relate to mental states? One possibility, which has recently been defended by advocates of an embodied perception model of our knowledge of others’ minds, is that it is a kind of part-whole relation, for instance an expression of anger would be a proper part of being angry. Thus, Joel Krueger and Soren Overgaard write:

Taking “expression” in a constitutive sense is the idea that certain bodily actions are expressive of mental phenomena in that they actually make up proper parts of some mental phenomena. In other words, some mental phenomena have a hybrid structure: they straddle internal (i.e., neural) and external (i.e. extra-neural, gross bodily) processes. (2012, pg. 245; cf. Krueger, 2012)
If, as Krueger and Overgaard claim, the relation between expressive behaviors and mental states is one of proper parthood, then this would seem to support the claim that we have direct perceptual acquaintance with mental states themselves. This idea that we are perceptually acquainted with the mental states of others via expressive behavior is the central thesis of the embodied perception model and it is meant to capture one of the insights of classical phenomenology, which many embodied perception theorists draw on for inspiration, namely that mental states are typically ‘embodied’ in some sense rather than hidden ‘internal’ states of a person (e.g., Gallagher, 2008, 2012; Krueger, 2012; Krueger and Overgaard, 2012; Smith, 2010).  

In addition to accurately characterizing the phenomenology of social interactions, Krueger and Overgaard also believe that a significant amount of empirical research, from a variety of sources, supports their proposal that expressive behaviors are proper parts of some mental states. Since I have already discussed this work in footnote 17, I shall not do so again. But it is worth noting that even if evidence does show that certain expressive behaviors are necessary for experiencing some emotions, this would not yet demonstrate that the behavior in question was a proper part of an emotional state. It would only show that the latter depends on the former such that if take away the expression we would, in these cases, take away the emotion. Parthood is one way to explain this dependence but there are a number of different relations that could also explain it. Thus, one hypothesis would be that temporally extended emotional experiences require feedback from the autonomic nervous system, such that if something inhibits signals from that system, a subject would have either fewer or less vivid emotional experiences. Similarly, exogenous stimulation to the autonomic nervous would produce misleading feedback as if one were experiencing an emotion (cf. Kret, 2015). But
I think the proposal that expressive behaviors are proper parts of mental states faces a number of challenges. First, if *Angry Patrick Stewart* is possible, then we know that someone can be angry without expressing anger, which means that the expression is not an essential part of being angry. So Krueger and Overgaard must think that it is only in cases where a subject is actually expressing anger that the expression is a proper part of the subject’s anger. However, this goes against the thought that expressions are responses to anger or are somehow caused by anger. The embodied perception theorist wants to reject this conception of the mental, but then it isn’t clear how exactly we should understand everyday interactions that appear to presuppose it. When, for example, we ask our friend why she screamed at us and she replies 'because I was very angry about your being late to my recital', we naturally take her to be offering us a kind of causal or reason-giving explanation of her behavioral response. This type of response is perfectly natural but it would be a bit more puzzling if the expression were simply a proper part of her anger.

Second, it is not clear to me from Krueger and Overgaard’s remark about ‘straddling’ internal and external processes what precisely the other parts of an emotional mental state are supposed to be.\(^{20}\) Are they nothing more than neural processes, or is something more

---

\(^{20}\) One might be inclined to take this language of ‘straddling’, and similar remarks about ‘affect programmes’, to be an indication that Krueger and Overgaard’s parthood relation is more like set membership than mereological constitution. If expressive behaviour were a part of a
traditionally ‘mental’ supposed to be a proper part of emotions as well? This is an important point because the temporal character of expressive behavior is not obviously of the right kind for it to be a proper part of a mental state. Several philosophers writing on the ontology of mind have noted that differences in the temporal profiles of states and events indicate that the two belong to fundamentally different ontological categories (e.g., Soteriou, 2013; Steward, 1997). States are entities that are, in some sense, wholly present at each moment of the interval of time during which they exist. This is connected to their lacking a progressive verb form (i.e., one cannot answer the question ‘what are you doing’ with ‘being happy’). Events, on the other hand, seem to be entities that unfold or develop over time and are not wholly present at each moment of the interval during which they exist. If it ordinarily takes 10 minutes for me to walk to the store, then I have not walked to the store after only 5 minutes (although I may be in the process of walking there, which is why it would be appropriate to answer the question ‘what are you doing?’ with ‘walking to the store’). Given that there seems to be this difference in the temporal structure of events and states, it isn’t straightforward that an expression, which would be an event that unfolds over an interval of time, is the type of thing that could partially compose an emotional state, which is the kind of thing that endures through time. So, at the very least, the embodied perception theorist needs

mental state in the way that the number 2 is a part of the set of even numbers, then the questions I am raising about ontological categories could be avoided. But this would not help the perceptual theorist because one does not generally perceive a set in virtue of perceiving one of its members.
to give us a more complete picture of how precisely event-like constituents like expressions could compose whole entities like emotional states.21

Finally, my biggest concern with Krueger and Overgaard’s appeal to parthood is that it does not even guarantee that we have perceptual knowledge of others’ minds. For one thing, having perceptual acquaintance with a part of something does not mean that one perceives the whole thing (cf. McNeill, 2012). Martin (2010a) offers an example wherein one looks through a slit and sees a flash of turquoise that is part of a scarf. As Martin rightly concludes, one does not thereby see the scarf. In response to this sort of concern, Overgaard has recently argued that, although it is true that seeing a part of an object does not generally entail that one sees the object, it is plausible that one ‘counts as seeing’ the whole in cases where one is perceptually acquainted with a sufficiently significant part (2014, pg. 139; cf. Neta, 2007). I do not know what would settle whether or not expressive behaviors were sufficiently significant for seeing an entire mental state. But regardless, the broader concern is that seeing an object is an extensional relation. From the fact that a perceiver is perceptually acquainted with a certain object, nothing follows about which facts involving that object the perceiver sees. To use Dretske’s example, ‘the cat can smell, and thus be aware of, burning toast as well as the cook, but only the cook will be aware that the toast is burning.’ (1993, pg. 266) So even if we suppose that the embodied perception approach is right, this would only

21 One thought would be to analyse mental states as a series of events (perhaps some combination of neural and behavioural events). However, this would not be a simple characterization of our everyday point of view on the world but a substantive metaphysical thesis in need of further argument. For further discussion, including some worries, about this sort of analysis, see Soteriou, 2013 and Steward, 1997.
show that we are perceptually acquaintance with some mental states, not that we have perceptual knowledge of others’ minds. It would seem that the embodied perception theorist would need to say more, at least if they intend their proposal to be an account of our knowledge of others’ minds.\footnote{Surprisingly, it seems that not all embodied perception theorists do intend this. For example, Overgaard writes that ‘the embodiment thesis is not supposed to constitute a full account of how we can see(i) that others are angry. Rather, it is supposed to supply one significant piece to this puzzle…’ (2014, pg. 137) If the argument presented in this essay is right, then there are reasons to think this puzzle cannot be completed. However, this should not be misconstrued as an argument against the embodied perception ‘piece’ of the puzzle.}

7. Expression and Inference

One thing to admire about the embodied perception approach is that it takes seriously the idea that expressions are the source of our knowledge of others’ minds. Many of the more traditional approaches to our knowledge of others’ minds, in spite of superficial discussion of expressive behavior, do not. For example, a number of philosophers have claimed that our knowledge of another person’s mind is based on some kind of probabilistic inference from expressive behavior, but nothing about this proposal requires that the basis of the inference be an expression. Indeed, any observable feature that was reliably correlated with the occurrence of a mental state would suffice to ground such an inference. Thus, this approach does not really help us understand why it would matter for the behavior a person exhibits to be an expression, or why it might be significant that we ordinarily take people to be expressing, or failing to express, their mental states. Indeed, the most natural ways to
characterize many of our social interactions, including our way of knowing about others’ mental states, suggest that the expressive nature of a person’s behavior is crucially important.

Within the confines of this essay, I cannot develop a complete account of the expressive but in the remainder of this section I shall sketch a proposal for thinking about the relationship between expressive behaviors and mental states, which builds on Austin’s idea that ‘there is a peculiar and intimate relationship’ between a person’s mental states and their natural expressions (1946; pg. 180).

It will be useful to contrast the type of view that I want to introduce in this section with a standard way of thinking about expressions as behaviors that are merely caused by a person’s mental states. Sellars calls this the ‘causal’ sense of expression and takes its canonical form to be that of a thought causing a particular utterance (1969, pg. 520). For example, we naturally take a spontaneous utterance of ‘London is a wonderful city to live in’ to be caused by, and therefore express, the thought that London is a wonderful city to live in. Although this seems right, notice that the behavior Sellars focuses on is not caused by just any underlying mental state but it also has the same semantic content as the state which caused it. This is interesting because it suggests that there may be some additional criteria for genuinely expressive behavior.  

---

23 It seems easy to imagine how my belief that London is a wonderful city to live in might cause a range of behaviors, including, for instance, casually browsing real estate advertisements in The Guardian. To my ear, it sounds odd to call this an expression of my belief. It at least seems to be importantly less of an expression than my assertion of ‘London is a nice place to live’ (cf. Author Suppressed).
What I want to focus on, however, is the fact that according to this causal conception of ‘expression’ the relation between an expression of a mental state and the state being expressed is completely contingent. If we think that the occurrence of a mental state is nothing more than the cause of a particular pattern of observable behavior, we allow for the possibility that the very same kind of behavior could be caused by something else, perhaps even by something not mental. So even if an expression of M is normally caused by M, the causal sense of expression allows for occurrences of an expression of M without its normal cause. Keeping Austin’s remark in mind, this seems to be an insufficiently intimate relation.

By contrast, the type of view that I want to introduce does not think that an expression of a mental state could occur, as an expression, without the occurrence of the mental state it is expressing. Rather, it holds that there is a kind of necessary connection between the expression of a mental state and the mental state being expressed. Another way to put this is to say that an expression of M metaphysically depends on the occurrence of M, so the former could not occur without the latter.

To say that an expression of M metaphysically depends on the occurrence of M is to say more than that the former supervenes on the latter. It is to make a claim about the nature of expression, about what it is for something to be an expression (whereas supervenience would merely note a necessary covariance). Thus, on the view I am proposing, it is the nature of an expression of M to be appropriately related to an occurrence of M. So, even though it is plausible that an expression of M is causally related to M, this would not exhaust the way in which the former depends on the latter.\(^{24}\)

\(^{24}\) For the purposes of this essay, I wish to leave it open whether or not M must be a fully sufficient or partial cause of an expression of M. It seems plausible to me that the subject
As long as we are individuating patterns of observable behavior in terms of the category ‘expression’, something cannot occur as an expression of M without M also occurring. It is important to realize, however, that this is not to deny that there may be other ways of individuating the very same pattern of observable behavior, which do not appeal to its nature as an expression. For example, we may be able to individuate an expression of anger in completely non-mental terms, as a sequence of precise bodily movements. But if we do individuate it this way, then we allow that the behavior could exist without the occurrence of anger. On the view I am recommending, this means it would not be an expression of anger and we would therefore be individuating it at a level of description that overlooked an essential aspect of what it is.

It is worth highlighting that this proposal claims that there is a type of constitutive relation between an expression of M and M. But notice that this constitutive relation is not identity, as would be true for the behaviorist, nor is it parthood, as was suggested by the embodied perception theory we considered in the previous section. There is reason to think that this is a positive feature of the view. In the previous section, we saw that considerations about the temporal character of events and states suggest that an expression could be neither identical to nor part of a mental state. So if there is to be some kind of constitutive relation between the two (a relation stronger than causation), it must hold between existentially distinct states. The proposal that expressions of mental states metaphysically depend on the mental states they are expressing offers precisely this kind of necessary relation.

herself may be causally involved in an expression of M, as is suggested by some ‘neo-expressivist’ theories of self-knowledge (e.g., Bar-On, 2004; Finkelstein, 2003).
This way of thinking about the relation between expressions and mental states is similar to a view about the nature of conscious experience defended by Matthew Soteriou in *The Mind's Construction*, in which he states ‘there appear to be states whose obtaining necessarily requires the occurrence of events *while* the state holds—states that obtain for a given period of time only if certain kinds of events occur *during* that period of time.’ (2013, pg. 105) The converse could be said of the proposal I am outlining in this section: there appear to be events, namely expressions of mental states, whose occurrence necessarily requires the obtaining of a state *while* the event unfolds—events that occur only if certain types of states obtain during the time they occur. There is obviously more to be said about the nature of the expressive, including its relation to a subject’s agency, but the suggestion that a subject’s expressive behavior is necessarily connected to her mental states strikes me as a plausible way of elaborating upon Austin’s point that there is a ‘peculiar and intimate’ relation between the two.

What implications does this sort of view have for our knowledge of others’ minds? If the arguments presented earlier in this essay are correct, then it is not plausible that we have perceptual knowledge of others’ minds. One might think the only alternative is that our knowledge is based on some type of inference. The idea that our way of knowing about another mind must be either perceptual or inferential is a peculiar feature of many philosophical discussions of our knowledge of mentality. The two paths are usually presented as exhaustive options: either our knowledge of others’ minds is perceptual or it is inferential.

However, in a recent essay, Anil Gomes rejects this dichotomy and claims that ‘when one reflects on the way in which we know about the mental lives of other people, it’s not perception or inference which stands out as the central means by which we come to know about others’ mental lives.’ (2015, pg. 173) Instead, Gomes notices that we often seem to gain
knowledge of others’ minds because of what they tell us and this observation motivates his intriguing proposal that testimony, which he argues is neither perceptual nor inferential, can be a basic source of our knowledge of others’ minds.

I do not wish to dispute Gomes’s assertion that we sometimes acquire testimonial knowledge of another person’s mind. What I do want to say is that his rejection of the two standard options is driven in part by the fact that he is responding to a philosophical tradition with a very particular way of understanding what is involved in inferential knowledge. The standard conception of ‘inferential knowledge’ that figures in discussions of our knowledge of others’ minds is one which takes it to be knowledge based on a kind of probabilistic inference. On this conception, a legitimate inferential basis lends epistemic support to a proposition \( P \) by making \( P \) more likely or probable. Thus the most popular inferential accounts of our knowledge of others’ minds take it to be based on either an analogical or an abductive inference (Gomes, forthcoming).

But we do not have to think of all inferential knowledge in probabilistic terms. Indeed, there are some cases where it seems strange to think of a warranted inference along these lines, namely cases where an inferential basis entails the truth of the proposition that one comes to know on that basis. For example, suppose that we learn, perhaps through testimony, that Samantha’s car is either crimson or scarlet. It seems that we can justifiably infer that the car is red. But it would be odd to think that our basis for this inference, our knowing that Samantha’s car is either crimson or scarlet, did nothing more than make it probable that her car is red. Indeed, it rather seems that it guarantees or entails that the car is red. This recommends adopting a broader conception of inferential knowledge on which one’s knowledge that \( P \) counts as inferential if it requires one to know some distinct proposition \( Q \) (cf. Pryor, 2005).
If we adopt this broader conception of inferential knowledge, then the proposal sketched earlier in this section implies that our knowledge of others’ minds is inferential. The central claim of that proposal is that an expression of a mental state metaphysically depends on the state it is expressing, but the two are nevertheless existentially distinct. This naturally suggests that our knowledge of a person’s mental states is based on knowledge of her expressive behavior. For instance, if we know that someone is expressing anger, then we are able to know that she is angry. Since, on this picture, our knowledge of the person’s mental state would be based on knowledge of a distinct fact, it would be inferential.

But I think it would be a mistake to think of this knowledge as a kind of probabilistic inference. The epistemic basis upon which we would come to know about another person’s mental state (her expressive behavior) would be necessarily connected to the fact we come to know on that basis. Unlike the probabilistic model, our knowledge that one is in a particular mental state would not be based on evidence that is compatible with the falsity of what we come to know. Rather expressive behavior would be a kind of fact-entailing evidence, which is perhaps the reason that linguistic constructions of the form ‘x expresses Ψ’ seem factive. Thus, there is a sense in which the model I have proposed means that our knowledge of others’ minds is inferential, but the type of inference involved would more closely resemble deductive proof or demonstration than inductive argument. It therefore deviates from the standard conception of inferential knowledge found in the philosophical literature on others’ minds, which Gomes wants to resist.

It is quite natural to think that behavioral expressions have a major role to play in securing our knowledge of others’ minds. This may be part of the reason that most people outside of philosophy are quite reluctant to accept the perceptual model. If we could straightforwardly see that someone is angry, then her expressive behavior would be
something like a distraction from a much more direct way of knowing what she is thinking or feeling. However, if a person’s expressions did nothing more than increase the probability of her being in a particular mental state, they wouldn’t really have any greater epistemic value than any other type of sign, symptom, or indicator of what the person is thinking or feeling. So if we want to respect the intuition that expressions are uniquely important for acquiring knowledge of what someone is thinking or feeling, then we need to envision expressive behavior along the lines Austin recommends, as something that stands in a peculiar and intimate relationship to a subject's mentality.\(^{25}\)

References


---

\(^{25}\) Many of the ideas in this essay were inspired by several enlightening conversations with Anil Gomes, who was also kind enough to offer valuable comments on earlier versions of the essay. I am extremely grateful to him. For additional inspiration and for many helpful suggestions on previous drafts, I am also very grateful to Anita Avramides, Bill Brewer, Mike Martin, Will McNeill, and Josh Shepherd. Parts of this essay were presented at Cardiff University, King’s College London, and Ruhr-Universitat Bochum. On each occasion, I was asked very helpful questions for which I am grateful. Finally, I would like to thank two anonymous referees and the editors of *Mind* for their thoughtful comments.


Gomes, A. forthcoming: Skepticism about Other Minds. In D. Machuca and B. Reed (ed.) *Skepticism: From Antiquity to Present*


Krueger, J. and Overgaard, S. 2012: Seeing Subjectivity: Defending a Perceptual Account of


Pryor, J. 2005: There is Immediate Justification.


Ryle, G. 1949: *The Concept of Mind*. University of Chicago Press,


Siegel, S. 2006: Which Properties are Represented in Perception?


