

Reasons of which Reason Knows Not

Jennifer Church

"The heart has its reasons of which reason knows not " -- Pascal¹

There is something deeply puzzling about the idea that we can be ignorant of our own reasons. Whereas Pascal attributed certain reasons to the *heart* versus the mind, we may attribute such reasons to the *emotional* versus the cognitive mind or to the *unconscious* versus the conscious mind, but the central conviction remains the same: we may be moved by reasons without ever consciously thinking of them as reasons. How is it possible, though, to respond to something *as a reason* without consciously thinking it through? In what follows, I elaborate on why we should be puzzled by the possibility that our reasons are unknown to us, explaining why some simple solutions are inadequate, and developing a more promising line of response based on the nature of action in space. While I shall focus, for the most part, on unconscious reasons, in the end it should be clear how the proposed line of response helps to make sense of emotional reasons as well.

Two Possibilities

Pascal's quote may be interpreted in (at least) two different ways, depending on whether we assume that the "reasons of which reason knows not" are not *accessed* by consciousness or that, even when accessed, they are not *recognized as* reasons. The difference can be illustrated with an example: I find myself afraid of someone that I meet at a friend's house; I don't know why I am afraid of this person since he has not said or done anything threatening, and he is evidently a good friend of my friend; and yet I suspect that there is a reason for my fear. On the one hand, it is possible that this person actually has done something threatening -- a menacing look in my direction, or a secret smile of pleasure at another's discomfort, for example -- and I have noted this behavior unconsciously. If I were to become conscious of this behavior, I would recognize it as a reason for my fear and could

¹ "Le coeur a ses raisons que la raison ne connaît point." [1670, ed. L. Brunschwig, 1909, sec.4, no.277]

give it its due. On the other hand, it is possible that this person merely looks like someone else from my past -- someone who did endanger me -- and I have extended my fear from one to the other. If I were to become conscious of this extension, I would recognize it as unreasonable and I would try not to be affected by it. In the first case, I might be said to have reasons of which I am simply unaware; in the second case, I am affected by considerations that I (at least consciously) do not regard as reasons at all.

In his famous essay "The Unconscious", Freud [1915] introduces a similar distinction between what he calls a "topographical" model of the unconscious and what he calls a "functional" model of the unconscious. On the topographical model, information that is unconscious does not 'reach' consciousness because it is recorded in a different 'place' than information that is conscious. On the functional model, in contrast, information that is unconscious is processed in a different way than information that is conscious. As Freud realizes, these two models are not necessarily incompatible, as information stored in one place (a part of the brain, presumably) could also function in a different way than information stored in some other place. But Freud clearly favors the functional model for its appreciation of some apparent peculiarities of unconscious thought: its indifference to temporal order, its lack of negation, the fluid way in which it condenses and displaces contents, and its inability to distinguish psychic reality from external reality.² My tendency to respond to a new acquaintance in the same way that I respond to a past acquaintance with similar hair, for example, might be due to the fact that within the unconscious I fail to distinguish past and present and I tend to condense representations of two individuals into one.

Unconscious reasons may be *more or less accessible* to conscious thought, of course: I may realize that it is something about the way a man looks at me that makes me fearful without being conscious of what that way is, or I may have no idea why I am fearful; and I may become aware of what is bothering me after a moment's thought or I may remain ignorant despite intensive therapy. Likewise, unconscious reasons may be *more or less recognizable* to conscious thought as being reasons: I may view the assimilation of my lover to my parent as bizarre and wholly

² These peculiarities are mentioned at many points in Freud's writing, but listed most explicitly in Section V of "The Unconscious" [1915].

irrational or it may seem quite sensible, even if potentially misleading. The less access or recognition there is, the more puzzling things become; but, as we shall see below, even the milder cases of lack of access or lack of recognition challenge our usual assumptions about the nature of reasons.

Before turning to these challenges, it is worth noting that a failure to consciously access certain reasons could be closely related to their unusual ways of functioning and thus to their failure to be recognized as reasons. If consciousness consists in some kind of meta-level reflection on the contents of one's thoughts, and if hypothetical reasoning presupposes a capacity for reflective bracketing of these contents, then it would follow that unconscious reasoning, lacking the capacity for such meta-level reflection, would lack the capacity for hypothetical reasoning as well.³ Or if the distinction between past and present can only be registered within a language, and language requires consciousness, then it would follow that unconscious reasoning, lacking linguistic distinctions, would be lacking in temporal distinctions as well.⁴ Just how a failure of *access* might be related to a failure of *recognition* will depend, in part, on just what one takes consciousness to be -- a topic that is well beyond the scope of this paper. Here we will be content if we can solve some of the puzzles that surround both possibilities.

Hidden Reasons: the Problem of No Access

We've come a long way since Descartes' insistence that consciousness is the very essence of mentality -- that it is not possible to have a belief or a desire or an intention or a sensation without at that very moment having knowledge of it.⁵ Our explanations of each other and of ourselves, both informally and within the science of psychology, are now replete with appeals to unconscious beliefs and desires. But if consciousness is no longer considered necessary for mentality, what is? As John Searle points out: "The explanatory power of the notion of the unconscious is so

³ Zoltan Dienes and Josef Perner [1999], for example, summarize various studies that suggest that "implicit" information cannot be used in hypothetical reasoning. One explanation of why this might be so is offered in Church [1987].

⁴ Defenders of the view that consciousness depends on language include philosophers Donald Davidson [1975] and Peter Carruthers [2000]. Psychologist Richard McNally [2003] argues that our pre-linguistic memories are so different in kind from our later memories as to be inaccessible to consciousness.

⁵ There is some dispute, still, over whether the requisite knowledge of my current mental state includes knowledge of its content. See Radner [1988].

great that we cannot do without it, but the notion is far from clear". [1992, p.151]
On what basis are we justified in attributing a particular mental state to someone in the absence of any first-person knowledge of that mental state?

The usual answer to this question relies on treating unobserved mental states like unobserved physical states -- presumed to exist on account of their observed effects. If I am behaving in just the way that I would if I consciously believed that my neighbor is a threat, then we may be justified in supposing that I have the unconscious belief that my neighbor is a threat -- just as we may be justified in supposing that there is an unseen moon if we find various objects behaving in just the way they do when a moon is observed to be present. This position concerning the *epistemology* of unconscious states is often, though not necessarily, accompanied by a functionalist *metaphysics* of mental states -- the view that mental states are constituted by their causal relations with the environment and with each other, that it is their causal relations rather than their intrinsic properties that make them what they are. On this view, states that *act* like mental states *are* mental states, by definition, and knowledge of how a mental state functions provides direct rather than indirect knowledge of what that mental state is. A functionalist, then, can accommodate a lack of first-person access to our mental states in several ways: (1) by treating self-knowledge as a non-essential causal effect, (2) by treating self-knowledge as just one of a cluster of defining causal effects, no one of which is necessary as long as most obtain, and (3) by treating self-knowledge as a necessary possibility, though special circumstances may prevent it from actually occurring. These three options parallel claims about the moon to the effect that (1) direct observation of a moon is irrelevant to its existence, (2) direct observation of a moon is one among many possible kinds of evidence, at least some of which must obtain in order for there to be a moon, and (3) direct observation of a moon must be at least possible, though special circumstances may prevent it from occurring.

Many have argued that while option (1) seems to be the most plausible in the case of physical objects and events, option (3) is more plausible in the case of mental states.⁶ The longstanding and familiar problem with option (3), however, is the

⁶ John Searle [1992], for example, suggests that the possibility of self-knowledge or consciousness is needed in order to lend opacity to the contents of our minds -- i.e. in order to allow concepts to have the same extension but not the same meaning. I think that Searle

problem of distinguishing between normal or ideal circumstances (in which there is self-knowledge) and special circumstances (in which there is not) in a way that is neither empty nor arbitrary. Consider the case of an unobserved moon: normal or ideal conditions seem to include our being appropriately situated, attentive, and without contextual interference. But these specifications will be empty if they simply mean that we must be situated in such a way that observation occurs, that we must be in a state of mind that registers the relevant information, or that nothing must prevent observation occurring. And the specification of ideal circumstances will be arbitrary if they require us to be within some specified distance from the moon, at some particular level of alertness, or without anything blocking our visual field. Likewise, in the case of unconscious mental states, it is not clear what normal or ideal conditions for self-knowledge should include: could they include an expanded capacity for short-term memory, training in biofeedback, or an absence of pain? Again, if all manner of helpful adjustments to our psychology are allowed, the constraint becomes empty; but any attempt to draw a line between acceptable and unacceptable adjustments seems arbitrary. Insisting that unconscious mental states must be capable of becoming conscious, therefore, does not really help us distinguish between legitimate and illegitimate attributions of unconscious content.

Where does this leave us? If, contra Descartes, there are mental states to which we lack first-person access -- in actuality if not in principle, then they must be known (and perhaps also constituted) by their characteristic ways of functioning; and legitimate attributions of unconscious content will be those that are supported by sufficient evidence of states that function in these ways. Their characteristic ways of functioning will be determined, in the first instance, by determining how conscious mental states function. And what ways are those? Most would agree that beliefs function to record one's surroundings, that desires function to motivate one's actions, and that beliefs and desires interact with each other to generate still other beliefs and desires. (It is these latter interactions, between mental states, that mark the presence of internal states with content, as opposed to a mere series of reinforced responses to one's environment.) Such interactions are, moreover, *reasonable* interactions in so far as the information of one's senses *ought* to be

underestimates our ability distinguish between co-extensive concepts on the basis of their causal effects, however; and I think that he overestimates the determination of content provided by first-person descriptions (privileged, yes, but not perfect).

reflected in one's beliefs, the objects of one's desires *ought* to be pursued in one's behavior, and one *ought* to develop new beliefs and new desires on the basis of old.⁷ Reasonable interactions between beliefs, desires, and actions are precisely those that are described by the canons of practical and theoretical reasoning, and one thing that is immediately evident from these canons is the fact that premises must share some but not all of their contents with each other and with a conclusion (i.e. they must have structure). A desire for X will lead to an action Y only in conjunction with a belief that Y will achieve X; and a belief that X is always accompanied by Z will lead to the further belief that Y will achieve Z; and so on. Unless there is some shared content between premises and conclusion, it is hard to see how the premises can be seen to inform the conclusion. And unless there is some difference between the contents of premises and conclusion, the conclusion cannot tell us anything new.⁸

The fact that unconscious beliefs and desires function in reasonable ways is not yet sufficient to show that, in their unconscious form, they can be *my* reasons (I return to this concern shortly). But it is sufficient to show that in order to make sense of unconscious reasons, we must make sense of unconscious contents with the sort of structural complexity described above. And although this requirement may seem straightforward if we simply think of unconscious contents as written in a 'language of thought', determining when and whether there is such an inner 'language' will depend, ultimately, on determining when and whether one's responses to the environment are sufficiently structured to warrant the attribution of such an inner structure. Kant [1787] was probably the first to address the way in which the structure of traditional syllogistic reasoning coincides with the structure of our experience of objects in space and time. More recent philosophers, largely inspired by P.F. Strawson's [1966] influential interpretations and appropriations of Kant, have continued this project of grounding the structures of reason in the

⁷ There is, of course, a sense in which the same might be said of stimulus-response conditioning: one ought to continue behavior that is associated with pleasure and one ought to avoid behavior that is associated with pain. But seeking pleasure and avoiding pain are things that we ought to do *other things being equal* -- that is, only if there are no longer-term consequences or competing priorities to factor in; and factoring these things in depends on acquiring a network of beliefs and desires that can interact with each other as well as with the environment.

⁸ The nature of and need for structure is a large topic, explored most extensively perhaps by Gareth Evans [1982].

structures of spatio-temporal experience. Christopher Peacocke [1983], for example, has offered a fairly detailed account of a type of behavior that makes it appropriate to attribute a network of beliefs and desires quite apart from any self-description that the subject may offer. He describes the way in which various non-linguistic animals will follow novel paths between two points (taking the shortest way home from a distant point that has been reached circuitously, for example, or making complicated detours around a suddenly problematic route); and he suggests that this behavior only makes sense on the assumption that such animals have an internal 'map' of the position of various objects (including themselves) within a single space -- a space that admits of infinitely many paths between each of the objects in that space. When confronted with such behavior, then, it makes sense to suppose that a subject believes that its home behind an obstacle, that it wants to go home, that it believes that by veering to the right and then back to the left it will bypass the obstacle, and so on.⁹

Peacocke does not think that the beliefs and desires that are revealed by such complex spatial behavior count as "judgments", however. The difference between animal beliefs and human judgments, he suggests, is that non-human animals can't withhold assent from experience, nor can they lie.¹⁰ This seems right since the ability to temporarily 'bracket' a representation that is prompted by perception, or to 'detach' a linguistic representation from the mental state that it expresses, seems to depend on an ability to register representations *as* representations -- an ability not only to represent the world to oneself but an ability to recognize oneself as doing so. But granting this distinction, we can now ask: does a belief only become a reason when it becomes a judgment as well? must I recognize myself as having a particular belief in order for that belief to count as my reason?

Far from being questions merely about our use of the word "reason", these are questions about how we understand the connection between mentality and

⁹ For another and rather different description of animal reasoning in space, see Bermudez [2002], pp. 251-257.

¹⁰ Peacocke discusses this concern on p.77 [1983]. Bermudez [2002], expresses a similar concern with his claim that non-linguistic animals cannot "access validity" because they cannot "understand the appropriateness of an action relative to an instrumental belief in terms of the constructibility of a valid argument." (p. 257 and p.261)

normativity. Most everyone would agree that, at least in general, in order for something to count as a representation, it must be capable of *misrepresentation*, and in order for something to count as the content of a belief, it must be capable of being *false*.¹¹ If I can judge of a representation that it is a misrepresentation, or of a content that it is false, then clearly I am aware of the distinction between correct and incorrect use; i.e. I am aware of its normativity. But if I am not aware of my beliefs and desires, then how can they be normative for me? And if they are not normative for me, how can they be my *reasons*? We may consider certain sorts of belief-desire interactions to be reasonable or normative from the 'point of view' of evolution, or from the 'point of view' of a surrounding community, but that does not make them reasonable or normative from my point of view. Indeed, I may think that evolution often favors false beliefs over true, or that the community that surrounds me represents the world in systematically misleading ways. Nor will it do to suppose that my unconscious point of view is whatever point of view makes most of my unconscious beliefs true and most of my unconscious desires wise, for there will always be *some* way of interpreting my behavior in such a way that all of my beliefs *are* true. (This is the point made by Wittgenstein [1958] in his extended discussion of the difficulty of determining, from behavior alone, that I am making a mistake as opposed to following some idiosyncratic rule.)

I think that a more promising response to the puzzle of how unconscious reasons can be normative for me begins with the observation that we often do sense that we have made a mistake, and act as though we have made a mistake, even when we remain unaware of what our mistake was or how to correct it. I might realize that I have the wrong 'take' on someone I've just met without knowing what my take is; I simply sense that I have got it wrong. And my disconcerted behavior around that person may signal to others that I have become uneasy with my usual assumptions without even knowing what those assumptions are, or how to replace them. Likewise, in the case of non-human animals, there is every reason to think, on the basis of its behavior, that a dog can recognize that it has made a mistake even though it cannot represent, even to itself, what that mistake is.¹² Davidson's

¹¹ Exceptions to this rule may include the contents of necessarily true propositions in math or logic and self-reports on ones' own sensations

¹² One of the most common examples of our ability to register mistakes as mistakes without knowing what the mistake is occurs when we display an awareness that we have spoken ungrammatically, or have used the wrong word for something, even when we are unable to

[1982] suggestion that a non-linguistic animal could be startled but not surprised seems oblivious to the clear difference between the way a dog reacts to a sudden clap, for example, versus the absence of a recently-buried bone.¹³ (I do not mean to suggest that there is a distinctive sensation associated with the recognition of having made a mistake; only that certain feelings tend to accompany behavior that is indicative of our recognition of a mistake. In the final section of this paper, I say more about the relation between the structure of feeling and the structure of behavior.)

Strange Reasons: the Problem of no Recognition

So far we have assumed, and indeed counted on, a considerable overlap between the functioning of reasons of which we have first-person knowledge and reasons of which we lack first-person knowledge. For it was just such overlap in manner of functioning, as revealed in characteristic patterns of behavior, that justified the attribution of unknown reasons at all. There is, however, a long tradition of supposing that both emotional reasoning and unconscious reasoning function in ways that are systematically different from those of conscious thought. Their operations are not only hidden from conscious thought but are different in kind from those of conscious thought. Pascal seemed to believe that the very 'logic' of the heart is different than the logic of thought; and Freud certainly believed that the unconscious functions in ways that are systematically different than those of consciousness. Is it possible to admit such systematic differences in functioning while continuing to treat emotional states and unconscious states as reasons?

The fact that emotional states and unconscious states have systematic and predictable effects on our beliefs and our behavior -- that anger makes us believe the worst of another, or that unconscious registration of a thing's strangeness makes us avoid it -- does not in itself make these states into reasons, nor does it

specify what the mistake is. Speech becomes halted and self-conscious, a word or phrase is repeated with a slight alteration (which may or may not itself be appropriate or grammatical), there is a small shrug, and more confident speech resumes. In such cases, people have no trouble reporting that they have spoken in a way that is somehow mistaken, but they may not be able to say why it is mistaken, or their attempts to describe the mistake may be wildly off the mark (focussed on the wrong word even). Thus, the feeling of having made a mistake, and the behavior that indicates an awareness that one has made a mistake, are enough to show that beliefs of which I remain ignorant may nonetheless be reasons for me.

¹³ Davidson's argument is discussed in my [1990].

ensure that they have content. For our conscious thoughts and actions are causally affected by various things, internal and external, that do not have any normative force or content: hunger may cause us to forget an appointment, for example, and caffeine may cause a flurry of cleaning, but hunger is not a reason for forgetting and caffeine is not a reason for cleaning. Any defence of emotional reasons or unconscious reasons as differently-functioning reasons must explain why this different functioning has normative as well as causal force.

Consider again the case where I am unexpectedly fearful of a new acquaintance because he reminds me of someone who has harmed me in the past. Does his similarity to this other person give me a *reason* for being afraid, or is it just the *cause* of my fear? There are good evolutionary reasons for grouping like things with like, and for responding to new things as one has responded to similar things in the past. But, as we saw above, responding to like things in like ways is something that can be accomplished through stimulus-response conditioning alone, and the reasons of evolution are not necessarily *my* reasons. Unless it makes sense to posit a network of interacting beliefs and desires by which I can feel normatively constrained, it makes no sense to attribute reasons to me.¹⁴ Could it be that my fearful response is not merely a conditioned response but one that is derived from an emotional conviction or an unconscious belief to the effect that this new acquaintance is actually the same person as the one who harmed me in the past? That is certainly possible, but attributing such a belief is only warranted on the assumption that this belief interacts appropriately with various other beliefs and desires. If I believe that the past offender has died, for example, then that belief ought to interact with the belief that the person before me is the offender to make me abandon my belief about his death, or add a belief about ghosts, for example. That is how such a belief would function if it were available to consciousness. That is not what happens in the case we are considering, however; in this case, I retain my belief that this is a new acquaintance *and* I continue to be afraid. (Freud claims, at times, that there is no contradiction in the unconscious; but surely the very idea

¹⁴ Likewise, while it may be useful to describe the 'heuristics' that guide my reasoning, unless these are rules of which I have representations and by which I feel normatively constrained, they will not function as reasons for me.

of content depends on the possibility that some contents are incompatible with others.¹⁵⁾

Could it be, then, that although I know that this person is a different person than the offender from my past, I unconsciously believe that anyone who shares certain facial features shares certain personality traits as well? By hypothesis, this is not a belief that I consciously endorse or one that would stand up to critical scrutiny (which does not mean that it is false); but it may be a belief that interacts with many other beliefs and desires to generate a larger set of prejudices and a wider range of biased behavior. In short, it could have a place in a rational network of beliefs and desires that explain my actions, questionable as those beliefs and desires may seem to my more reflective self. This is a real possibility, but it is no longer a case of emotional reasons or unconscious reasons functioning in some different way from those of conscious thought; it is another case of my failing to know what my reasons are, not a case of my failing to recognize my reasons as reasons. So we haven't yet described a case whereby emotional states or unconscious states can provide reasons for what we think and do without also conforming to the norms of reasoning that characterize conscious thought.

Let us return to the case of spatial reasoning. The sensitivity to spatial relations indicated by intelligent movement through space -- the ability to find novel paths home, for example -- was presented as evidence for an internal structure of beliefs and desires. Such behavior, it was suggested, indicates the existence of an internal 'map' that structures the acquisition of new beliefs and the generation of new actions much as more linguistic syllogisms structure the acquisition of new beliefs and new actions. Indeed, the structure of objects and movements in space may be viewed as grounding the structure of thought. This does not mean that spatial reasoning is like syllogistic reasoning in every respect, however; so if emotional reasons or unconscious reasons adhere to the structure of space yet lack the added capacities that come with language, there may indeed be ways in which their functioning is systematically different than that of conscious thought.

¹⁵ Cavell [1996] is right to reject Freud's claim about contradiction; but more sympathetic interpretations of Freud are possible here -- interpretation that emphasize a certain ease the unconscious has in 'getting around' contradiction through condensation and displacement, for example.

How does spatial reasoning differ from linguistic reasoning? One way to approach this question is to consider how various concepts can or cannot be captured in a purely spatial representation. Various abstractions -- notions of justice, the idea of a right, etc. -- seem to defy spatial representation (even if they are derived from a consideration of quite concrete objects and events that can be rendered spatially). If this is right, then reasoning that relies on maps only cannot include reasoning that involves such abstractions. But this seems to be a difference in subject matter rather than in manner of reasoning. More interesting for our purposes is the fact that certain logical constants -- *if/then*, *either/or*, *not* -- also seem to defy spatial representation; while it is possible to picture a scene in which there are *both* cats *and* dogs, there is no scene that represents the presence of *either* cats *or* dogs, and no way to picture a situation in which *if* there were cats, *then* there would *not* be dogs (which is not equivalent to a picture of only cats). It is also the case that pictures cannot represent an absence, and thus cannot portray something as existing in the past or the future rather than the present. So reasoning that depends on representing something as a mere possibility rather than an actuality, or representing something as past (or future) rather than present, is not a kind of reasoning that is available to purely spatial forms of reasoning.

How might this help us understand the case in which I seem to have an unconscious reason for my fear of a new acquaintance? If the similarity between the appearance of this person and some past person who endangered me merely triggers a similarity in my response, it will be a matter of conditioning rather than reasoning. To make my response reasonable, it must be supported by an unconscious belief to the effect that the person before me is the same person as that from my past. But why would I come to have that unconscious belief? The answer, I think, lies with the fact that unconscious beliefs (or, at least, many unconscious beliefs) may be represented in purely spatial form so when the perception of one person standing before me triggers an unconscious image of the other standing before me, the unconscious is confronted with two spatial representations that must be consolidated into a single image. Unconsciously, I must replace one image with the other or I must find a way to amalgamate or "condense" the two. These are perfectly reasonable adjustments within the unconscious insofar as it is unable to distinguish, spatially, between what is present and what is past, or between what is perceived (actual) and what merely imagined (possible). Thus, I have an

unconscious reason for my fear despite the fact that this reason would not be accepted as a reason if various non-spatial distinctions were also available to me.¹⁶

Condensation, so understood, is not simply a matter of grouping things in loose or non-standard ways. Marcia Cavell [1996] has suggested that many of Freud's cases are best understood as involving atypical and rather childlike ways of forming categories rather than any distinctive way of employing the categories one has.¹⁷ But my conscious concept of a new acquaintance and my conscious concept of a past offender are neither loose nor non-standard; it is only when I am restricted to a spatial type of reasoning -- a type of reasoning that characterizes much of our unconscious reasoning -- that the two combine into one. Nevertheless, doesn't such unconscious condensation of contents result in unconscious concepts that are different than their conscious counterparts? Yes and no. Certain concepts will be different in the unconscious, so if one individuates concepts solely on the basis of their functioning within a wider field of concepts -- i.e. if one is a thorough-going holist, then the spatialization of my thoughts will transform the very identity of the concepts involved. On the other hand, since reasoning itself depends on the possibility of holding one's concepts fixed while discerning new relations between them, it is hard to see how such thorough-going holism is compatible with the normativity of content.¹⁸ And as long as it is possible, as I think it must be, to distinguish between content-constituting relations between concepts and more strictly logical or formal relations between contents, then it will be possible to

¹⁶ Note how vivid imagining, caused (not justified) by a strong desire, can also result in an unconscious map that represents the world as being (more) as one would like it. Wishful thinking, so understood, is reasonable within the unconscious. This, I think, accords with Richard Wollheim's [1984] understanding of fantasy.

¹⁷ I think that Cavell is right about the case of the child "Hans," with respect to his concept of a 'widdler', since he evidently lacks the more specific concepts of 'penis' and 'udder'. And the more one focusses on visual similarities between the referents of one's concepts, the more liable one is to visual amalgams. Freud's adult patients, however, clearly possess the distinct concepts that Freud says get condensed in the unconscious. Cavell rejects so-called primary processing, including condensation, as constituting reason at all, but she accepts the existence of odd contents in the unconscious and what I am calling spatialized reasoning is related to what she would call proto-reasoning.

¹⁸ See above discussion of normativity. Some important objections to holism are discussed in Fodor and LePore [1992].

understand the peculiarities of spatialized reasoning not merely as peculiarities of content but peculiarities of function as well.

So far, spatial reasoning has been characterized as a relatively primitive form of reasoning, available to non-human animals, impoverished by its inability to use logical constants, its inability to distinguish possibility from actuality, and its inability to make temporal distinctions. Spatial reasoning also offers certain advantages over more linguistic forms of reasoning, though. The demand that all objects of thought be representable in space gives such reasoning a kind of concreteness and precision that is often lacking in more abstract forms of thought; and the need to portray multiple relations simultaneously in a single space can serve to ensure that they are consistent with one another. Also, and importantly, spatialized reasoning forges a kind of visceral connection to the objects of belief and desire -- a connection that secures the transition from belief and desire to action (the move from the premises of a practical syllogism to its conclusion). For in order to be "moved" by reasons, reasons need to be grounded in the structures that govern our movement in space. (Note how often we need to vividly imagine what we are talking about, or hearing about, before we are moved to act on it.)

It is not only those objects and relations that are inherently spatial that are representable in space. One of the functions of metaphor and of art more generally is to give spatial representation to things that lack it -- to cast a near-death experience as a "brush" with death or as a particular sweep of the hand through space, to portray one's home as the "earth" or as a magnet, and so on. Transitions between thoughts are also spatialized when we speak of one thought as "leading to" or "following from" another, "containing" or "excluded by" another, being "grounded in" or "transcending" another.¹⁹ Such creations and transformations make our language and our experience 'live', not only in the sense that they add newness and interest but, more profoundly, in sense that they forge a connection between our minds and our bodies, between what we think and what we do.

Emotions and the Unconscious

I have described the emotion of surprise as being an indicator of unconscious reasoning -- one that helps to show that I can distinguish between correct and

¹⁹ Mark Johnson [1987] catalogues many such metaphors.

incorrect applications of unconscious contents. And I have indicated that engaging the imagination, which is central to unconscious reasoning, will also engage us emotionally. But I haven't said much about the nature of emotional reasons -- our counterpart to Pascal's reasons of the 'heart'. In this final section, I want to suggest that the very nature of emotion can be understood in terms of imagined movements in space, with the result that the spatialized nature of unconscious reasoning is the nature of emotional reasoning as well.

Theories of emotion tend to divide between those that emphasize sensations and those that emphasize beliefs. Sensation theories identify each emotion with a specific sensation, noting that sensations are not under our control, they are not 'about' anything, and they may or may not accompany any particular belief.²⁰ Anger, it is said, is a distinctive feeling that can overtake us despite ourselves, that need not be about anything at all, and that may or may not accompany the belief that someone has wronged us, for example. Belief theories, in contrast, identify each emotion with a particular belief or judgment, whether or not a particular sensation is present.²¹ Anger, on such an account, might be identified with the belief that we have been wronged, or that an injustice has been done, even when there is no particular feeling that accompanies this belief. While the sensation theories can be seen as favoring the physicality and causality of emotion, the belief theories favor the intentionality and rationality of emotion. But part of what is so remarkable about emotion is the way in which it seems to bridge the two: emotions must be felt *and* they must be about something, they involve *both* sensations and beliefs -- not as an accidental pair but as a unified experience.

Several theorists of the emotions (arguably extending back to Spinoza) have suggested that sensations are themselves intentional insofar as they are about our bodily states. Antonio Damasio, most recently, has offered such a theory, summarizing his reflections as follows:

"what made the feeling deserve the distinctive term feeling and be different from any other thought, was the mental representation of parts of the body or of the whole body as operating in a certain manner. Feeling, in the pure and

²⁰ The theories of Descartes [1649] and James [1884] are classic examples of sensation theories.

²¹ Solomon [1983] and Lyon [1980] offer detailed defenses of belief theories.

narrow sense of the word, was *the idea of the body being in a certain way*. In this definition you can substitute idea for "thought" or "perception." Once you looked beyond the object that caused the feeling and the thoughts and mode of thinking consequent to it, the core of the feeling came into focus. Its contents consisted of representing a particular state of the body." [2003, p.85] Note, however, that the "idea of the body being in a certain way" is not an idea that can interact *rationally* with other ideas about the body; and insofar as it generates characteristic thoughts about the world around one, these thoughts are *caused* by the feeling, not *justified* by it. It seems misleading, therefore, to refer to feelings as representations with contents; or, in any case, they are not the sorts of representations or contents that can serve as reasons.

Suppose, instead, that one thought of emotions as internal versions of external movements. Fear, at its most basic, would be the internal counterpart to the external act of fleeing -- involving the retraction of various small muscles to produce a kind of fleeing inward that mimicks the outward act of fleeing danger. Disgust would be an internal counterpart to pushing objects away from oneself, and affection would be the internal counterpart to pulling objects towards oneself. More sophisticated emotions (or more sophisticated versions of these basic emotions) could be understood as internal forms of more sophisticated actions: pride as an internalized form of strutting or self-display, hate as an internalized form of destruction, shame as an internalized form of hiding, and so on.²² These internal 'acts' may replicate (and exaggerate) the internal changes that typically anticipate an outward version of the same (as in Darwin's [1872] analysis of angry barring one's teeth as appropriate preparation for an attack); or they may imitate an external act in ways that actually interfere with the successful completion of that act (as when we are so frozen by fear that we are unable to flee). Emotions, so understood, would retain the intentional structure of the actions from which they derive -- for they would be directed towards certain objects in space, but they would also retain the physicality of action -- since they would necessitate appropriate movements within the body.

This is not the place to offer a full account of such a theory, which I defend elsewhere [1995]. What is important, for our purposes, is the realization that the

²² de Rivera [1977] offers a grid of emotions along these lines.

structure of emotion, and of emotional reasons, may be precisely the structure of reasons that guide movement in space and that make the possibility of unconscious reasons intelligible. If this is right, then our earlier presentation of spatial reasoning as a kind of reasoning that we may confidently attribute even in the absence of consciousness, and as a kind of reason that is different in kind from forms of reasoning more typical of consciousness, will carry over to the case of emotional reasoning as well. Our frequent lack of access to emotional reasons and their distinctive functioning can be explained in the same way that we explained our lack of access to unconscious reasons and its distinctive functioning.

In closing, it is appropriate to note how Pascal himself aligned reasons of the heart with our prereflective understanding of space, admonishing us to consider such knowledge foundational, if also incomplete:

The heart feels that there are three spatial dimensions and that there is an infinite series of numbers, and reason goes on to demonstrate that there are no two square numbers of which one is double the other. Principles are felt, propositions proved, and both with certainty though by different means. It is just as pointless and absurd for reason to demand proof of first principles from the heart before agreeing to accept them as it would be absurd for the heart to demand an intuition of all the propositions demonstrated by reason before agreeing to accept them.

Our inability must therefore serve only to humble reason, which would like to be the judge of everything, but not to confute our certainty. As if reason were the only way we could learn! [1670, ed. A.J.Krailsheimer, 1966, sec 6, no 282]

I have tried to vindicate Pascal's conviction that we can feel the force of a genuine reason without being in a position to articulate its content, and without being able to subject it to critical scrutiny. Also, like Pascal, I have looked to our prereflective grasp of the complexity of space in order to elucidate some of the peculiarities of "reasons of which reason knows not." If Pascal is right, and if I am right, both unconscious reasoning and emotional reasoning are not only possible, they are a fundamental part of our lives as reasoners.

References

Bermudez, Jose Luis. "Rationality and Psychological Explanation without Language", *Reason and Nature: Essays in the Theory of Rationality*, eds. J.L. Bermudez and A. Millar. Clarendon Press, 2002.

Carruthers, Peter. *Phenomenal Consciousness*. Cambridge University Press, 2000.

Cavell, Marcia. *The Psychoanalytic Mind: From Freud to Philosophy*. Harvard University Press, 1996.

Church, Jennifer. "L'émotion et l'intériorisation des actions". *Raisons Pratiques 6: La couleur des pensées*, eds. P. Paperman & R. Ogien. Paris: Éditions de l'École des Hautes Etudes en Sciences Sociales, 1995.

_____, "Reasonable Irrationality". *Mind*, vol.96, no. 3, 1987.

_____, "Judgment, Self-Consciousness, and Object Independence". *American Philosophical Quarterly*, v.27, no. 1, 1990.

Damasio, Antonio. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Harcourt, 2003.

Darwin, Charles. *The Expression of Emotion in Man and the Animals*, 1872, University of Chicago Press, 1965.

Davidson, Donald. "Thought and Talk", *Mind and Language*, ed. Samuel Gettenplan, Oxford University Press, 1975.

_____, "Rational Animals", *Dialectica*, 1982. Reprinted in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. Ernest LePore and Brian McLaughlin, Blackwell, 1985.

de Rivera, Joseph. *A Structural Theory of the Emotions*. International Universities Press, 1977.

Descartes, René. *Passions of the Soul*, 1649, trans. Norman Kemp Smith, *Descartes' Philosophical Writings*, Macmillan, 1952.

Dienes, Zoltan and Perner, Josef. "A Theory of implicit and explicit knowledge," *Behavioral and Brain Sciences* 22 (1999).

Dummett, Michael. "Thought and Language", *Origins of Analytic Philosophy*, Duckworth & Company, 1993.

Evans, Gareth. *Varieties of Reference*, Oxford University Press, 1982.

Fodor, Jerry and LePore, Ernest. *Holism: a shopper's guide*. Blackwell, 1992.

Freud, Sigmund. *Standard Edition of the Complete Psychological Works of Sigmund Freud*. Trans. and ed. J. Strachey et al. Hogarth Press, 1953-1974.

_____ "Analysis of a Phobia in a Five-year old Boy", vol. 10 [1909]

_____ "The Unconscious", vol. 14 [1915]

_____ "Mourning and Melancholia", vol. 14 [1917]

_____ "New Introductory Lectures on Psychoanalysis", vol.22 [1933]

James, William. "What is an Emotion?" *Mind* 9, 1884.

Johnson, Mark. *The Body in the Mind*. University of Chicago, 1987.

Kant, Immanuel. *Critique of Pure Reason*, 1787. Trans. Norman Kemp Smith, Macmillan, 1929.

Kripke, Saul. *Wittgenstein on Rules and Private Language*. Harvard University Press, 1982.

Lyons, William. *Emotion*. Cambridge University Press, 1980.

McNally, Richard J. *Remembering Trauma*. Harvard University Press, 2003.

Pascal, Blaise. *Pensées*, 1970, ed. A.J. Krailsheimer, Penquin, 1966.

Peacocke, Christopher. *Sense and Content: Experience, Thought, and their Relations*. Clarendon Press, 1983.

Radner, Daise. "Thought and Consciousness in Descartes", *Journal of the History of Philosophy*, 1988.

Searle, John. *The Rediscovery of the Mind*. MIT Press, 1992.

Solomon, Robert. *The Passions*. University of Notre Dame Press, 1983.

Strawson, P.F. *The Bounds of Sense: an essay on Kant's 'Critique of Pure Reason'*. Methuen, 1966.

Wittgenstein, Ludwig. *Philosophical Investigations*, Blackwell, 1958.

Wollheim, Richard. *The Thread of Life*. Harvard University Press, 1984.