

WHAT CAN COGNITIVISM TELL US ABOUT AGENTS'  
PATTERNS OF IDENTIFICATION?

Michael Smith

Valuing has a decidedly Janus-faced character.

On the one hand, to say that an agent values her performing some action is to convey her normative perspective on the world, a perspective which sustains the distinction between appearance and reality, and one for which the agent herself might have more or less justification. To this extent it seems that an agent's valuing some action is a cognitive matter, a matter of her having certain beliefs about the desirability of the particular action in question. On the other hand, when an agent comes to the conclusion that she values some action she seems to be in a state with the potential to lead her all the way to action. No additional desire to (say) do whatever it is that she happens to value doing is needed. Valuing an action thus seems to be a matter of desiring. On the assumption that no cognition can be a motivation — on the assumption, that is, that the belief that an action is desirable cannot be a desire to bring that action about — we appear to be impaled on the horns of a dilemma.

In recent times non-cognitivists have responded to this dilemma by insisting that we focus on an agent's pattern of emotional identification. Their idea has been, very roughly, that an agent values an action when she identifies with a motivation so to act, where such identification is in turn a matter of the agent's having a certain heirarchically ordered set of desires about that motivation's leading her all the way to action. Given that an agent who values performing some action therefore desires that she desire to perform that action, it is meant to be unsurprising that her valuing an action can lead her all the way to action. Valuing an action can lead an agent all the way to action because it can lead her, in the normal way desires that lead to the realisation of their content, to her desiring to perform that action.

The task that the non-cognitivists have set themselves has thus been to spell out their heirarchical account of identification without at any point assuming that the agent values anything. Without going into details, it seems fair to say that this has proved to be no easy task. At some point in explicating what it is for an agent to identify with a motivation non-cognitivists

either assume that the agent values something, or else they fail to give an explication that is plausible. (For the record, it seems to me that that is just the beginning of the non-cognitivists' problems with their account of identification, but more of that later.)

Accordingly, the time seems ripe not just to ask what a cognitivist might say about the original dilemma, but more importantly to ask what a cognitivist might say about the concept of identification itself. Non-cognitivists are, after all, surely right that we have an ordinary, everyday, concept of identification; and they are also surely right that the concept of identification connects up in some yet-to-be-determined way with the concept of valuing. Even if they have failed adequately to analyse the concept of identification themselves, the question still remains whether cognitivists can do any better. My task in the present paper is to address these issues.

In the first section I begin by offering a cognitivist response to the original dilemma: an account of what it is that an agent who values some action believes, when she believes an action is desirable. The account I give makes it plain why agents who have such beliefs are in a state that can lead them all the way to action. With that account of valuing as believing desirable in the background, in the second section I offer an account of what an agent's pattern of identification amounts to in the purely cognitive realm. The account I give of identification in the purely cognitive realm suggests a hypothesis about the nature of identification more generally. I offer this hypothesis at the beginning of the third section, and, in the remainder of the paper, I test that hypothesis against a range of examples. I draw some final comparisons with the non-cognitivists' account of identification in a brief conclusion.

### 1. A cognitivist response to the original dilemma

Let me explain how I think a cognitivist should respond to the original dilemma.

A cognitivist must, first and foremost, come up with a plausible account of what it is that an agent believes when she values some action, that is, an account of what it is for an action to be desirable. It seems to me helpful, in this connection, to begin from the more or less common sense assumption that for an agent to value her acting in a certain way in certain circumstances is for her to believe that so acting is advisable: that is, a matter of her believing that she would

advise herself to perform that act in those circumstances if she were herself in circumstances in which she was best placed to give herself advice. Two questions naturally spring to mind. First, what are these circumstances in which agents are best placed to give themselves advice, and second, what fixes the content of the advice that the agents in those circumstances would give themselves?

The answer to the first question is, I suggest, that agents are best placed to give themselves advice when their psychologies have been purged of all cognitive limitations and rational failings. The answer to the second question, the question about the content of the advice that agents would give to themselves, is that the content of such advice is fixed by the contents of the desires that they would have, were their psychologies thus purged, about what they are to do in the circumstances of action about which they are seeking advice. In other words, when I value my performance of a certain action that amounts to my believing that my performance of that act is advisable, where that, in turn, amounts to my believing that I would want myself so to act if I had a desire set that was purged of all cognitive limitations and rational failings.

If something like this is along the right lines then all we need in order to get a full-blown analysis of desirability is to give an account of the conditions that need to be met by a desire set which is devoid of cognitive limitations and rational failings. My suggestion in this regard, developing an idea of Bernard Williams's, is that for a desire set to be devoid of cognitive limitations and rational failings is for it to be one which is maximally informed and coherent and unified. If this is right then it follows that what it is desirable for an agent to do in certain circumstances is a matter of what she would want herself to do in those circumstances if she had a set of desires that was maximally informed and coherent and unified. If we call the possible world in which the agent has the desires that she actually has in the circumstances of action she faces the 'evaluated' world, and the possible world in which she has that set of desires that is maximally informed and coherent and unified the 'evaluating' world, then, the suggestion is, what it is desirable for her to do in the evaluated world is fixed not by what, in the evaluated world she wants herself to do in the evaluated world, and not by what, in the evaluating world, she wants herself to do in the evaluating world, but rather by what, in the evaluating world, she wants herself to do in the evaluated world. This, accordingly, is the property that an agent must believe her act to have when she values the performance of that act.

Once this is agreed it seems to me that there is no difficulty at all in seeing why valuing has the Janus-faced character noted earlier. Since valuing is a matter of having certain beliefs about what is desirable it should come as no surprise at all to learn that an agent's values convey her normative perspective on the world, a normative perspective which sustains an is/seems distinction, and for the adoption of which she might therefore have more or less justification. For an agent's evaluative beliefs convey her normative perspective in exactly the same sense in which her beliefs about any subject matter convey her perspective on that subject matter; her evaluative beliefs convey how she takes the facts about the desirability of the actions she is contemplating performing to be, not merely how those facts seem to her to be; and, finally, her evaluative beliefs are capable of being more or less justified, depending on how well she has formed them in the light of the evidence available to her about what she would want if she had a maximally informed and coherent and unified desire set.

But nor should it come as any surprise to learn that an agent who values acting in a certain way has a belief that is capable of both causing and rationalizing certain desires without the aid of any further desire, for example a desire to do what she values. In order to see that this is so, imagine a case in which, on reflection, you come to believe that (say) you would desire that you keep a promise you made in the circumstances of action that you presently face if you had a maximally informed and coherent and unified set of desires, but imagine further that you don't have any desire at all to keep that promise. Now consider the pair of psychological states that comprises your belief that you would desire that you keep your promise in the circumstances of action that you presently face if you had a maximally informed and coherent and unified set of desires, and which also comprises the desire that you keep that promise, and compare this pair of psychological states with the pair that comprises your belief that you would desire that you keep your promise in the circumstances of action that you presently face if you had a maximally informed and coherent and unified set of desires, but which also comprises instead your aversion to keeping that promise. Which of these pairs of psychological states is more coherent?

The answer would seem to me to be plain enough. The first pair is much more coherent than the second. There is disequilibrium or dissonance or failure of fit involved in believing that you would desire yourself to act in a certain way in certain circumstances if you had a maximally informed and coherent and unified desire set, and yet being averse to the prospect of acting in

that way. The aversion is, after all, something that you yourself disown. From your perspective it makes no sense, given the rest of your desires. By your own lights it is a state that you would not be in if you were in various ways better than you actually are: more informed, more coherent, more unified in your desiderative outlook. There would therefore seem to be more than a passing family resemblance between the relation that holds between the first pair of psychological states and more familiar examples of coherence relations that hold between psychological states. Coherence would thus seem to be on the side of the pair that comprises both the belief that you would desire that you keep your promise in the circumstances of action that you presently face and the desire that you keep that promise.

If this is right, however, then it follows immediately that if you are rational, in the relatively mundane sense of having a capacity to have the psychological states that coherence demands of you then, at least when that capacity is exercised, you will end up having a desire that matches your belief about what you would want yourself to do if you had a maximally informed and coherent and unified desire set. In other words, in the particular case under discussion, you will end up losing your aversion to keeping your promise, and acquiring a desire to keep it instead. The belief that you would desire that you act in a certain way if you had a set of desires that was maximally informed and coherent and unified would thus seem able to cause you to acquire a corresponding desire when it operates in conjunction with the capacity to have coherent psychological states. Moreover, because acquiring the desire makes for a more coherent pairing of psychological states, it would seem to follow that the desire thus caused is rationalized as well. Finally, note that no causal role at all needs to be played by any desire. All that is required is a capacity to have coherent psychological states, a capacity whose exercise is ubiquitous across both the cognitive and the non-cognitive realms.

## 2. Identification and belief

With this account of what it is for an agent to value an action in the background, let's now consider what it might mean to talk of an agent's pattern of identification. One obvious suggestion springs immediately to mind, a suggestion that has everything to do with the fact that valuing turns out to be a matter of believing.

Consider what it means to talk of a subject identifying with (say) one philosophical theory rather than another. Suppose, for example, that you are a philosopher of action and that I ask you whether you think that the causal theory of action is plausible or not. Let's suppose that, by way of response, you begin to tell me all about Davidson's version of the causal theory of action, perhaps in lavish detail and with a keen eye to all of the subtleties. It seems to me that I might well insist that you aren't answering the question I asked. I didn't want you to tell me what reasons other people who believe the causal theory have for their belief, I asked you whether you believe the causal theory, and, if so, why. Suppose that you take the point, but then proceed to tell me all about your teachers' various ingenious objections to Davidson's version of the causal theory. My objection would be exactly the same. If this pattern were to continue then I would probably conclude that you don't have any view about the plausibility of the causal theory of action. Indeed, I might conclude that you don't have a view about the nature of action at all. You are merely a walking and talking catalogue of other people's views. You have, as we might say, lots of 'book knowledge' about theories of action, but, in one perfectly ordinary sense of the concept of identification, you do not identify yourself with any theory: not with the causal theory, and not with any alternative to it either. You do not identify yourself with any particular view as regards the nature of action because you do not stand up for one theory rather than another.

As I said, the idea of identification at work here would seem to be familiar enough. To say that a subject identifies with one philosophical theory rather than another registers the fact that she has given the matter a certain amount of reflective consideration, and that her considered opinion is that that philosophical theory is correct. These are the philosophical theories that the subject identifies with, or stands up for, in the sense of being willing and able to defend them in rational argument. Moreover, what goes for a subject's belief in a philosophical theory goes for all of her beliefs. With regard to any subject matter whatsoever, we can always ask whether people identify with certain views about that subject matter, in the sense of having come to those views as a result of giving matters a certain amount of reflective critical scrutiny. These will be views that they are willing and able to defend in rational argument. These view will contrast with those that the subject does not identify with at all: those she has no view about, say, because she hasn't given the subject any thought whatsoever.

Of course, in everyday life we rarely have the opportunity to give any subject matter our undivided attention, so the answers that subjects give to questions about the views that they have about some subject matter will reflect matters of degree. A particular subject may have given (say) gun control a certain amount of thought, and so be willing and able to defend the view she has arrived at to that extent, but she might be quite prepared to admit that there could be considerations that she hasn't taken into account, considerations that she could and should have taken into account if she had given the matter her undivided attention. She might therefore begin her defence of her views by saying something like 'Look, I might be completely wrong, but it seems to me that....' Such a person would therefore be identified with that view about gun control to a corresponding extent.

This, in turn, makes it plain why it is at least possible for people to have views with which they do identify, but not all that much — or perhaps it will sound more appropriate to put that in the passive voice: this explains why subjects can have views with which they can be identified, but not to any great degree. These are views about subject matters that people have, not because they have subjected the matter to much in the way of critical scrutiny, but rather because (say) they were taught to believe these things at school, or in the context of being brought up by their family, or by watching television, or whatever, and simply haven't bothered asking whether that learning process resulted in their having true beliefs. In other words — and this is the crucial point — these will be views that a subject will only be willing and able to defend in argument to a very limited extent. Of course, a particular subject might be lucky enough to find that, when she does finally subject a particular such belief to reflective critical scrutiny, it just so happens to pass muster. The crucial point, however, is that that would be purely a matter of luck. The subject herself is in no position to take any credit for the fact that she has a belief which survives the process of critical reflection, because the fact that it survives that process plays no role in an explanation of why the subject has the belief in the first place. This, accordingly, is why, even in such cases, we cannot identify subjects with these beliefs to any great extent.

Note that the facts of identification that are in play here are not facts about which subjects will themselves necessarily be authorities. For example, a subject might be identifiable to a certain extent with a particular view and yet deny that he is identified with that view at all.

Imagine someone who makes it plain in all of his actions that he believes that women are inferior to men, and yet who, when asked outright whether he believes that women are inferior to men insists that he believes no such thing. This agent has a view about women alright, perhaps not one that he would be willing and able to defend in argument to any great degree, but a view none the less, and yet he mistakenly denies this to be so. Of course, we would certainly expect agents to be authorities about those views with which they identify fully, but that is merely a consequence of the fact full identification requires so much in the way of reflective critical scrutiny. Once subjects have given the claims that they believe that much in the way of reflective critical scrutiny it is hard to imagine how they could fail to be aware of the fact that they believe the claim in question.

The fact that subjects can be identified to a greater or lesser extent with certain of their beliefs is, in turn, an illuminating fact about them, for it tells us something about how far they can go, or are prepared to go, in the exercise of their rational capacities, and, in so doing, it gives us a standard against which we can measure subjects' actual beliefs. To the extent that a particular subject fails to have beliefs with which we can identify her to any great degree — in other words, to the extent that she does not have the beliefs she has because she has subjected them much in the way of critical scrutiny, but rather through inertia after having acquired them in some less less critical way — we have grounds for thinking that she doesn't live up to her responsibilities in intellectual matters. She doesn't think for herself, but rather lets others do her thinking for her.

On the other hand, to the extent that a subject can be identified with beliefs which others see glaring reasons for rejecting, objections which the subject herself overlooks or dismisses lightly, we have (defeasible) grounds for either thinking of her as once again failing to live up to her responsibilities in intellectual matters, or, perhaps, as having reached the limits of her intellectual capacities, and so the limits of what we can reasonably expect of her in the way of belief. The problem with such an agent is not that she doesn't think for herself, she does, the problem is rather that, given her limited abilities, thinking for herself isn't enough to get her all the way to a well justified belief.

Finally, to the extent that a particular subject can be identified with beliefs which she has subjected to high degree of critical scrutiny, and to the extent that she is able to get others to share those beliefs with her by explaining why they are indeed the ones that are rationally defensible, we will (again defeasibly) think of her as not just having lived up to her responsibilities in the formation of her beliefs, but as positively deserving of congratulation and admiration for having formed her beliefs so well. She won't just be deserving of congratulation and admiration in so far as she thinks for herself, but also in so far as the thoughts she thinks are the thoughts to think.

Will identification, in the sense in play here, correlate with anything distinctive in the phenomenology of belief? For example, will agents who cannot be identified to any great extent with their beliefs feel some discomfort or lack of satisfaction with their beliefs, by contrast with those who have beliefs with which they identify? This is an extremely vexing question. Certainly there are expressions we commonly use to describe those who have beliefs with which they cannot be identified much at all, in the sense in play here, expressions like 'There's no one at home', 'The lights are out', 'He's got such a vacant look about him', and the like, and these expressions contrast markedly with those that we use to describe people who have beliefs with which they can be identified, expressions like 'She's switched on', 'You can see the sparkle in her eye', 'She's so bright eyed and bushy tailed!', and the like. Moreover, these expressions do seem to be suggestive of some associated phenomenological difference between the people thus characterised. I am, however, reluctant to hazard a guess as to what exactly that phenomenological difference might consist in. The crucial point I want to emphasise, however, is that such phenomenological differences as there are, if indeed there are any, will at best be correlated with the different patterns of identification. They will not themselves be constitutive of those differences.

Given what we have just about identifying subjects with particular views, it should be plain what we have in mind when we talk of agents identifying with particular evaluative claims. To say that an agent identifies with particular evaluative claims registers the fact that she has given those matters a certain amount of reflective consideration, and that her considered opinion is that those evaluative claims are correct. The evaluative claims that an agent identifies with are thus those which she is prepared to stand up for in the sense of being willing and able to defend

them in a rational argument. The evaluations an agent identifies with are thus to be contrasted with she doesn't identify with at all — these are evaluative claims that she has no view about one way or the other— and they also contrast with those evaluative claims with which agents can only be minimally identified, those which, say, the agent learned while growing up, but hasn't ever really subjected to much in the way of reflective critical scrutiny.

Facts about how identified agents are with certain evaluative claims will, in turn, serve as grounds for judgements about the extent to which those agents live up to their epistemic responsibilities in the evaluative realm. Those who can only be minimally identified with the evaluative claims they accept would seem to have failed to live up to these epistemic responsibilities: they tend not to think through evaluative matters for themselves, but are rather prepared simply to continue believing whatever it is that they were taught at home, or what they have been exposed to through media, or whatever. Those who are maximally identified with certain evaluative claims would seem to have discharged their epistemic responsibilities to a correspondingly high degree: they do think through evaluative matters for themselves, rather than being prepared simply to believe, without question, whatever it is that they were taught. And those who are not just maximally identified with certain evaluative claims, but who are able to convince others in the process of rational argument that these are the evaluative claims that they too should accept will be (defeasibly) deemed to be especially deserving of congratulation and admiration. For these people don't just think for themselves, the evaluative claims that they are led to believe through the process of thinking for themselves are the evaluative claims to believe.

Identification with evaluative claims can thus, in these ways, be seen simply to be an instance of the more general phenomenon of a subject's identifying with one view rather than another, or perhaps with no view at all. Though it seems to me that it cannot be overestimated just how important this phenomenon of identification with evaluative claims is to moral theory — think of how impoverished an account of autonomy would be if it didn't, at some point, make reference to the extent to which agents identify with the evaluative claims which they accept, in the sense of identification in play here — I can well imagine non-cognitivists becoming impatient at this point and objecting that, interesting and important though the phenomenon might be, it doesn't really compete with their account of identification, because theirs is an

account of what it is for an agent to identify with a motivation, not an account of what it is for a subject to identify with a view. In the next section I address this objection. It seems to me to radically underestimate how readily we can generalize the account of what it is for a subject to identify with a view.

### 3. Identification and the exercise of rational capacities

The foregoing discussion of identification in the purely cognitive realm suggests a hypothesis about the nature of identification as such. As we will see, that hypothesis contrasts starkly with the non-cognitivist's account of identification.

The hypothesis is that when we talk of agents' patterns of identification we are quite generally talking about the extent to which we can see those agents as exercising their rational capacities. In other words, when we say that agents are identified fully with something or other, what we have in mind is that we can see in that thing the full exercise of the agents' rational capacities; when we say that they aren't identified at all with something or other, what we have in mind is that we can't see their rational capacities exercised in that thing at all; and when we say that they are identified to a certain extent with something or other, then what we have in mind is that we can see in that thing an exercise, to a corresponding degree, of their rational capacities.

If this hypothesis is correct then we would expect to find that talk of agents' patterns of identification is appropriate virtually anywhere that agents can display their rational capacities. Now, we certainly do find ourselves quite happily talking not just about agents being more or less identified with particular views, or beliefs, but also with their being more or less identified with particular motivations, emotions, intentions, and actions. The idea, according to the hypothesis, is that what makes such talk appropriate is in each case the fact that we can see in those motivations, or in those emotions, or in those intentions, or in those actions, a corresponding expression of the agent's rational capacities. Let me consider some examples in order to test the hypothesis.

#### (i) Motivations

Consider first a case of identification with motivations.

Imagine an agent who believes it is desirable to (say) keep a secret that he has been told in so far as he was told the secret in confidence. However, at lunch with some acquaintances, and, desperate for his lunch partners to think that he is fun to be with, he also comes to believe that it is desirable for him to tell them that secret, at least in so far as it would make his acquaintances think that he is fun to be with. So far he is faced with a conflict of evaluations. Let's suppose, then, that when he weighs these two values against each other, he concludes that it is more desirable that he keeps the confidence than it is for him to make his lunch partners think that he is fun to be with. In other words, he believes that it more desirable to keep the secret than it is to tell it to his lunch partners. If, in such a case, the agent has some motivation to tell his lunch partners the secret, and no motivation at all to refrain from doing so, what should we say about his pattern of identification with regard to his motivations?

The account of valuing offered earlier looks tailor made to provide an answer to this question that dovetails with the hypothesis. For, according to that account it seems that what the agent in this case believes is that if he had a maximally informed and coherent and unified desire set, then he would have conflicting desires about what he is to do in the circumstances of action in which he finds himself. On the one hand, he believes that he would have some desire to do what's necessary to make his lunch partners think that he is fun to be with, a desire that would give rise to a corresponding desire that he tells them the secret. But on the other hand he also believes that he would have a stronger desire that he keeps the confidence, a stronger desire that would in turn give rise to a correspondingly stronger desire that he keeps the secret. Overall, then, it seems that what the agent believes is that, if he had a desire set that was maximally informed and coherent and unified, then, though conflicted, he would more strongly desire that he keeps the secret that he was told in confidence than that he tells that secret to his lunch partners.

If this is right, however, then it follows immediately that the beliefs and desires this agent possesses departs from the rational ideal. It departs from the rational ideal because, to the extent that the stronger desire to keep the secret is absent from his psychology, his overall psychological state lacks coherence. If the hypothesis is along the right lines then we should conclude that, to the extent that the agent is more strongly motivated to tell the secret to his lunch partners than to refrain, he is not to be identified with a desire with that particular strength. He is

not to be identified with a desire of that strength because, much as in the purely cognitive case, he would not be prepared to stand up for the fact that he has a desire of that strength in the sense of being prepared to defend his possession of a desire with that strength in a rational argument. Indeed, quite the opposite is the case. The agent would almost certainly disown the strength of the desire, insisting that no rational defence of his possession of a desire with that strength is possible.

The fact that the agent cannot be identified with his having a stronger desire to tell the secret to his lunch partners, rather than to keep that secret, in turn grounds facts about the extent to which he lives up to his responsibilities in the motivational realm. For, on the assumption that the agent has the capacity to control himself — that is, on the assumption that he has the capacity to lose and acquire those motivations required for him to have a more coherent psychology — his failure to have a stronger desire to keep the secret is itself rationally criticizable. It is rationally criticizable because he could and should exercise self-control and acquire a stronger desire to keep the secret. On the other hand, if that assumption is incorrect and the agent does not possess the capacity to control himself, then we must conclude that though the agent isn't rationally criticizable, he is none the less radically deficient in terms of the rational capacities that he possesses.

In all these respects it thus seems to me that the hypothesis gets things exactly right. It doesn't just correctly diagnose the fact that the agent would not identify with his having a stronger motivation to tell his lunch partners the secret, but offers the right kind of explanation of that fact. The agent does not identify with his having a motivation to tell the secret of that strength because he deems its strength not to be rationally defensible. The account also correctly locates these facts about the extent of the agent's identification with his desire in the context of an assessment of the extent to which he lives up to his motivational responsibilities, in other words, his obligations to exercise self-control.

## (ii) Emotions

Consider next the case of emotional identification.

Gary Watson asks us to imagine a case in which an agent is quite decisively defeated in a game of squash. Feeling angry and humiliated, he desires nothing more than to smash his opponent in the face with his racquet. As Watson tells the story, we are not to suppose that the agent thinks that smashing his opponent in the face with his racquet is in any way desirable. Indeed, we can suppose that the agent thinks that everything is to be said against his doing so. In so far as the agent's evaluative beliefs are concerned, then, what he believes is that smashing his opponent in the face with his racquet is altogether undesirable. None the less, angry and humiliated as he is, he desires very strongly to do smash his opponent in the face with his racquet. The question to ask is thus whether the agent identifies with his emotions, in such a case.

Note that, at least in so far as his motivations are concerned, this case is in fact simply another version of the case described above, the case in which the relative strength of the agent's motivations fails to correspond with his estimations of the comparative evaluative significance of doing what it is that he is thus motivated to do. As such, for the same reasons as were given in that case, we should conclude that the agent Watson asks us to imagine is not identified with his motivations. He is not identified with his motivations because, given that he believes that it is undesirable for him to smash his opponent in the face with his racquet, coherence demands, if anything, that he be averse to doing any such thing. His desire to smash his opponent in the face with his racquet is thus, once again, one with which he cannot be identified because it is a desire that he is in no position to defend in a rational argument.

The difference between this case and the earlier case, however, is that, in this case, the agent's motivation is placed in the context of an emotion which typically gives rise to such a motivation. Anger is, after all, an emotion that typically manifests itself in a desire to damage the object of the anger. But now, to the extent that this is the case, it seems to me that we should conclude that the agent cannot be identified with his anger either. He cannot be identified with his anger because he is no more in a position to give a rational defence of his being angry than he is to defend his desire to smash his opponent in the face with his squash racquet. His anger, manifesting itself as it typically does in a desire to damage the object of the anger — in this case, the desire to smash his opponent in the face with his racquet — fails quite dramatically to cohere

with his belief that it is undesirable to smash the person who has defeated him in the face with his squash racquet.

If this is agreed, then note that there may even be a sense in which the agent can be held responsible for his anger. For while it might not be up to the agent himself whether or not he suffers the characteristic affect of anger — though, in other words, it might not be up to him whether or not, say, his blood boils — it might none the less be up to him whether or not he displays his anger. This is because, to the extent that the agent has the capacity to acquire and lose those motivations required for him to have a more coherent psychology — in other words, to the extent that he has the capacity to exercise self-control — then, given that he believes that it is undesirable to smash his opponent in the face with his squash racquet and notwithstanding the fact that anger is a state that typically gives rise to a desire to damage the object of the anger, it follows that he could and should exercise self-control and so prevent his anger from being, in this respect, typical. It might not be up to an agent whether or not his anger has its characteristic affect, but it might well be up to him whether or not his anger gives rise to the desires that are characteristic.

Note that if something like this account of what it is for an agent to be identified with his emotions is correct then one upshot is that it enables us to understand an otherwise puzzling feature of the emotions.

It is often remarked against the judgement theory of the emotions — the theory that an emotion is, *inter alia*, an evaluative judgement — that though those who are (say) afraid of spiders very often believe that spiders are dangerous, others, those who suffer from an irrational fear of spiders, may know full well that spiders are not dangerous, and yet be terrified of them none the less. The reason that this is supposed to be possible is that those who suffer from an irrational fear are unable to get themselves past the fact that spiders persistently appear dangerous to them, notwithstanding their knowledge that they are not dangerous. The suggestion, in other words, is that emotions like fear make evaluative illusions possible, where these illusions are supposed to be illusions in much the same sense as the Muller-Lyer is an illusion. Just as, in the case of the Muller-Lyer, we can know full well that the two lines are the same length even though they persistently appear different lengths to us, so, in the case of

irrational fear of spiders, we can know full well that spiders are not dangerous even though they persistently appear dangerous to us.

The puzzle, however, is to explain why we label the fear that people experience in such cases irrational. We do not, after all, label the perceptual state in which people find themselves when they suffer the Muller-Lyer illusion an irrational perceptual state. Why is there this difference? The answer is suggested by the account just given of identification with emotions. The irrational fear that people suffer in such cases has three distinguishable components. First, there is the belief that spiders are not dangerous. This is the analogue, in the case of the Muller-Lyer, of the belief that the lines are the same length. Second, there is the appearance that spiders are dangerous. This is the analogue, in the case of the Muller-Lyer, of the appearance that the lines are different lengths. But then third, and most importantly, in the case of an irrational fear of spiders, there is the additional fact that those to whom it appears that spiders are dangerous have the desires characteristic of fear, that is, they have the desire to flee. There is no analogue of this in the case of the Muller-Lyer illusion. Yet this is what forces us to label the fear irrational. For while the desire to flee coheres well with the belief that spiders are dangerous, it fails altogether to cohere with the belief that spiders are not dangerous, the belief that is possessed by those who have an irrational fear of spiders. The irrationality of the fear is, in this way, explained by the very same facts that explain why the fear something with which the agent cannot identify.

### (iii) Intentions

Consider next a case of identification with intentions.

Imagine two agents, A and B, who go into a supermarket looking for washing powder. They reach the relevant aisle and are confronted with hundreds of identical boxes between which neither can discriminate in terms of desirability. In other words, they each find themselves in a situation much like that in which Buridan's Ass found itself. However, let's suppose that A goes on to form an intention to take a particular box of washing powder, takes that box, pays for it at the check-out, and leaves the supermarket, whereas B remains in the aisle, paralysed by the fact that he can think of nothing that would make it more desirable to take one box rather than

another. What should we say about the extent to which A and B identify with their intentions, or lack there-of?

I take it to be uncontroversial that B, unlike A, would find his experience deeply alienating. Indeed I take it that B, unlike A, would probably think that he was going completely crazy. It seems to me that these uncontroversial claims about A and B themselves already suggest that the hypothesis that when we talk of agents' patterns of identification we are quite generally talking about the extent to which we can see those agents as exercising their rational capacities is along the right lines. The reason is that, as reflection on Buridan's Ass cases shows, among the various capacities that rational agents have we must suppose that there is a capacity simply to form intentions in cases of evaluative ties.

Of course, we must be careful in the characterisation of this capacity. For example, though it might look like a capacity to act for no reason at all, it would be a mistake to suppose that it is in fact simply such a capacity. It would be a mistake because someone who has the capacity to act for no reason at all has a capacity that could be manifested on all sorts of occasions, including those on which there is clearly, from the agent's point of view, a uniquely most desirable thing to do. This plainly isn't a capacity that any rational agent would possess. Instead, it seems to me, we should characterise the capacity in a much more tightly constrained way. It is a capacity to act for no reason at all when it comes to choices between alternatives which are tied in terms of desirability. It is this capacity that A manifests in forming his intention to take a particular box of washing powder, and it is B's lack of this capacity that explains why he remains paralysed in the aisle.

As I said, it should be plain why this account of the difference between A and B dovetails so well with the hypothesis that facts about the extent of agent identification are facts about the extent to which agents manifest their rational capacities. For the hypothesis suggests that A is indeed to be identified with his intention because, in forming his intention, he thereby simply manifests this rational capacity, and it suggests that B finds his experience deeply alienating because he is in a situation which requires the possession of a rational capacity that he seems to lack. Moreover, and perhaps somewhat paradoxically, note that what this claim about the extent to which each can be identified with their intention, or lack there-of, in turn entails is that A has

an intention that he is willing and able to defend in rational argument, by contrast with B. That is to say it is because A thinks, perfectly correctly, that a rational agent would form an intention for no reason at all in cases of evaluative ties that he can give a rational defence of the intention he forms, and it is because B thinks, incorrectly, that a rational agent could never form an intention for no reason at all that he can give no rational defence of his failure to form an intention.

#### (iv) Actions

Consider, finally, a case of identification with actions.

Again imagine two agents, C and D, but this time suppose that each of them desires to play the piano as if nervous, and each of them believes that they can play the piano as if nervous by playing an E at a certain crucial point in the piece that they are playing instead of playing an F, as the music requires. However, let's imagine that whereas C goes on to play that E instead of an F, as planned, and so manages to fake being nervous, the fact that D has the particular desire and belief he has makes him so nervous that, at the crucial point, he plays an E instead of an F, not as a pretence of nervousness, but rather as an expression of nervousness. What should we say about C's and D's patterns of identification with their respective playings of an E rather than an F?

Again, it seems to me uncontroversial that C would identify with his playing of an E rather than an F, and that D would not. Moreover, this is again just what we would expect, given the hypothesis that facts about the extent of agent identification are facts about the extent to which agents manifest their rational capacities. C identifies with his playing of an E rather than an F because, in playing an E, he manifests his rational capacity to satisfy his desires given his beliefs. D, by contrast, does not identify with his playing of an E rather than an F because his playing of an E rather than an F is not a manifestation of his rational capacity to satisfy his desires given his beliefs, but is instead a manifestation of his nervousness. Again, facts about agent identification reflect facts about the extent to which agents are willing and able to defend themselves in argument. C, unlike D, is able to give a rational defence of his playing an E rather than an F in because he can cite his reasons for so acting: his desire and his belief. D, by contrast, is not able

to give a rational defence of his playing an E rather than an F because that his doing so isn't susceptible to rational explanation at all.

### Conclusion

My aims in this paper have been two-fold. My first aim has been to give a cognitivist response to the original dilemma about the nature of valuing. My second aim has been to explain what cognitivism can teach us about the nature of agents' patterns of identification.

As regards the first aim, my suggestion has been that cognitivists should suppose that agents who value their acting in certain ways in certain circumstances have beliefs about the desirability of their acting in those ways in those circumstances, where facts about the desirability of actions performed in certain circumstances are, in turn, facts about what the agents of those actions would want that they themselves do in those circumstances if they had a set of desires that was maximally informed and coherent and unified. I have argued that accepting this analysis of desirability enables us to avoid the original dilemma because agents who have such beliefs, and who in addition have the capacity to have a coherent set of psychological states, are thereby able to acquire corresponding desires without the aid of any further desire.

As regards my second aim, my suggestion has been that we can learn a lot about the nature of agents' patterns of identification if we begin by looking at the purely cognitive phenomenon of agents' identifying with certain claims or views. In particular, I have suggested that, generalising what we learn by looking at this purely cognitive phenomenon, we see that facts about the extent to which agents can be identified with certain claims, or beliefs, or motivations, or emotions, or intentions, or actions, or whatever, are simply facts about the extent to which agents those claims, or beliefs, or motivations, or emotions, or intentions, or actions, or whatever, are manifestations of those agents' rational capacities.

I said earlier in this essay that the account I was going to offer of agents' patterns of identification contrasts markedly with the non-cognitivists' account of identification. Non-cognitivists, you will recall, offer an analysis of something quite specific, namely, what it is for an agent to identify with a certain motivation. Their suggestion, remember, is that identification with a motivation is a matter of an agent's having a certain hierarchically structured set of desires

about that motivation. There are three main points of contrast with the account of identification offered here.

First, the non-cognitivists' account of identification doesn't have anything much to do with the manifestation of rational capacities, whereas, as our initial focus on the purely cognitive phenomenon of an agent's identification with a certain view makes plain, that is one place where the concept of identification is most certainly at home. Second, whereas the account of identification offered here makes it plain why identification comes in degrees, it is not at all obvious how the non-cognitivists' account of identification is supposed to explain that fact. And third, it is difficult to see how to extend the non-cognitivists' account of identification with a motivation to the full array of cases in which we suppose agents to be capable of identification: cases of identification with a view, or identification with a belief, or identification with an emotion, or identification with an intention, or identification with an action.

In all these respects, then, it seems to me that the account of identification on offer here is far superior to the non-cognitivists' account. Having said that, however, let me close by saying what I do not take myself to have done in this paper. I do not take myself to have said everything there is to say about the concept of identification. Rather, I take myself to have isolated the core of that concept, at least in so far as the concept of identification has been employed in philosophical contexts to date. There is no doubt much more to say about the nature of identification. My hope is simply that the more that there is to say will emerge as a fairly natural extension of what has been said here.