

CAMBRIDGE
UNIVERSITY PRESS

Royal Institute of Philosophy

Playing the Rule-Following Game

Author(s): Julia Tanney

Source: *Philosophy*, Vol. 75, No. 292 (Apr., 2000), pp. 203-224

Published by: [Cambridge University Press](#) on behalf of [Royal Institute of Philosophy](#)

Stable URL: <http://www.jstor.org/stable/3751811>

Accessed: 10/11/2013 17:04

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Cambridge University Press and Royal Institute of Philosophy are collaborating with JSTOR to digitize, preserve and extend access to *Philosophy*.

<http://www.jstor.org>

Playing the Rule-following Game¹

JULIA TANNEY

I. Introduction

It has been suggested that in order to make the study of meaning more manageable, we ought to consider what must be known by an individual who is able to understand and speak a language. Philosophical questions about the nature of meaning have thus transformed into questions about the form that should be taken by an idealized theory of meaning, knowledge of which would suffice to explain a speaker's competence. Some scientifically-minded linguists, psychologists, and philosophers are attracted to this way of investigating language because they are optimistic that meanings—conceived as the contents expressed by the theorems of a meaning theory—can be understood as abstract representations in the brains of language users. Thus, the study of meaning, as such, is thought by some to invite a cognitive-scientific investigation grounded, optimistically, in the lower reaches of neuroscience. Recently, philosophers and psychologists have sought similar types of explanations of the ability to understand and predict rational behaviour by attributing to competent individuals a theory of mind, knowledge of which would explain these abilities, or correlatively, the lack of which would explain why certain individuals lack the ability to understand and predict the behaviour of others. Within the last decade, research in the cognitive sciences has focused on experimentation designed to test this hypothesis.

Rules or norms in some sense govern various social practices. An individual's knowledge of the norms is supposed to figure in an explanation of her ability to participate in the practice. But I suspect that there is something wrong—deeply wrong—with the attempt to give rule-following explanations of broadly rational activities and the problems inherent in such attempts do not seem to be solved by supposing that the rules are expressed in the contents of sub-cognitive states. Although I shall not be able to argue for it here, I suspect the difficulties in these projects extend to what is taken for granted in much contemporary philosophy of mind, as

¹ 'What this shows is that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call "obeying a rule" and "going against it" in actual cases.' Wittgenstein, *Philosophical Investigations*, § 201, trans. by G. E. M. Anscombe (Basil Blackwell, 1953).

they threaten certain contemporary accounts of reason-explanation and present a challenge to the very idea of causally-efficacious, content-bearing states.

In part II of this paper I shall develop a very general argument that raises a *prima facie* doubt about the coherence of certain attempts by theorists of meaning and mind that involve attributing knowledge of the relevant theory to an individual in order to explain her linguistic or rational abilities. (Henceforth, I call these ‘cognitive explanations’ and refer to the theorist who attempts this kind of explanation as a ‘cognitivist’.) The argument takes the form of a *reductio* in which the premises the opponent seems bound to accept lead to a dilemma, neither horn of which is viable. On one horn of the dilemma a vicious explanatory regress ensues. On the other horn, the regress may be halted only at the cost of destroying the normative nature of the practice, thus rendering attribution of *knowledge* of the norms pointless. I show how the dilemma arises by focusing on a norm-governed practice for which rule-following explanations make sense (I choose the game of baseball for this purpose) and argue that it is only when our broadly rational abilities are *presupposed* that such an explanation might even get off the ground. The attempt to abstract these broadly rational abilities and give systematic, cognitive explanations of them leads to the dilemma—a dilemma that confronts *any* attempt to give cognitive explanations of abilities whose broadly rational character is part of what needs to be explained. This includes, among others, the ability to understand and speak a language, the ability to understand and predict rational behaviour, and the ability to act rationally. I argue that the appeal to implicit knowledge does not manage to sidestep the regress, whereas the appeal to cognitive processes sidesteps it, only to become impaled on the other horn of the dilemma. In part III, I consider in more detail how this dilemma affects the idea that rational action is itself an ability that admits of rules or norms that govern it. If one supposes that knowledge of such rules is necessary for reasonable action, the same kind of dilemma arises and I show the negative implications of this dilemma for nomological accounts of reason explanation. In part IV, I consider the general ramifications of this discussion on rule-following for explanations involving the attribution of knowledge.

II. Cognitive explanations of rule-following abilities

1. A clear case of rule-following

In order to get a sense of when rule-following explanations make at least *prima facie* sense, it will be useful to consider ordinary games.

Playing the Rule-following Game

Games are relevant because they are typically governed by rules. The introduction of rules into the analysis of games seems indicated simply because they provide a standard in virtue of which it makes sense to judge particular moves in the game as 'correct' or 'incorrect.'² Human practices or activities of various kinds can be assimilated to games in so far as they, too, involve 'moves' that are apt for judgments of correctness or incorrectness. Consider the game of baseball. Call the moves that are made in the game of baseball—hitting the ball with a bat, running around the bases, etc.—'material' moves, or moves made in the 'object game'.³ The rules of baseball in some sense govern the material moves, but in what sense? Someone who knows nothing about baseball and whom we would be reluctant to describe as playing the game might make some of the material moves 'accidentally', as it were. Conversely, someone we would want to describe as playing the game might fail to make appropriate material moves (this person would be playing baseball incorrectly). In order to credit an individual with playing baseball we might seek some kind of 'internal connection' between her and the rules so that we can say that she is *making (or attempting) the moves as part of the game of baseball*. We might be satisfied if she were able to play what Sellars calls the 'metagame'; i.e., if she were able, using language, to cite the rules of baseball, and to make a case for the material moves being sanctioned by the rules.⁴ Perhaps here there would be nothing objectionable in describing this internal connection as a type of knowledge: someone who is able to cite the rules of baseball at the appropriate time might be said to *know* the rules and this knowledge might figure in an explanation of the moves she makes. Perhaps this knowledge might figure in an explanation of her ability to play the game.

² This is a minimal sense in which talk of rules is indicated. A serious topography of the various sorts of rules governing an activity such as baseball would be quite complex. It would involve, among other things, distinguishing regulative and constitutive rules; sub-rules governing professional games *v.* amateur games; National League *v.* American League; rules for umpires; rules for compiling statistics on the individual players, etc. My interest, in the main, is in the rules that are candidates for 'constitutive' rules—i.e., those that make baseball the particular kind of sports-activity it is—and in what the baseball player's relationship to those rules is supposed to be.

³ I am borrowing these terms, as well as some of the framework for discussion, from Wilfrid Sellars's 'Some Reflections on Language Games' in *Science, Perception, and Reality* (London: Routledge & Kegan Paul, 1963).

⁴ The suggestion is that the ability to cite the rules might be considered a (defeasible) sufficient condition for establishing the sought-after 'internal connection'. Whether or not it is also a necessary condition will be considered in the last section.

2. Playing the rule-following game

Suppose we are satisfied that an individual is making a move in the game because we have established the sought-after internal connection—this individual is manifesting *rule-following* as opposed to mere rule-conforming behaviour. She is not only able to make the material moves of the game, she is also able to correct or justify her moves by appealing to the rules of the game. But there is still a sense in which this individual, too, in being able to play not only the object game but the metagame as well, seems to be evincing general abilities—abilities involved in the very activity of following a rule—that might themselves go wrong.

Imagine, for example, someone who is being taught how to play baseball. Suppose she has a general understanding of the language her coach uses to describe the rules, yet she misjudges the domain of one of them: she takes it to apply to a situation to which it does not apply. Or suppose she can convince us of her general ability to apply rules in appropriate situations, but she cannot understand the meaning of some of her coach's words and therefore does not understand the rule she is supposed to be following. Or suppose she understands her coach's language, she understands the type of situations to which the rules apply, but she fails, at times, to act in the light of this understanding. Such errors are not happily described as failures to act in conformity with the rules of baseball for they may occur even if a person makes material moves that happen to conform to the rules of baseball. Rather, such errors (which may involve failure to understand the rule, failure to understand the relevant domain, or failure to implement one's understanding in action) thwart the attempt to follow the rules of baseball. People misfollow rules all the time and as a result tend to make mistakes: they bungle recipes and produce culinary disasters; they misinterpret, or misperceive signposts and travel in the wrong direction; and they make mistakes in following orders and get punished.

Following the rules of baseball or following recipes or maps—like making the material moves involved in playing baseball, or in cooking, or in travelling to a particular destination—are things that one can do correctly or incorrectly. This fact invites us to construe the activities that are involved in rule-following as themselves rule-governed, for it seems to be a platitude that error is diagnosable only in so far as a norm or standard has been violated. More concisely, we might say that the various kinds of activities involved in rule-following constitute a game, or a set of games. (By this I mean only that the appeal to rules, norms, or standards seems apt when a diagnosis of error can be made.) It should be clear, however, that if rule-follow-

Playing the Rule-following Game

ing is itself a kind of a game, it is not regulated by the sort of rules we have been considering—the rules of baseball do not tell you how to follow them. The rule-following game would consist in a different set of games which would require construing the appeal to the rules of baseball, or the appeal to a recipe, as a constituent ‘move’. Indeed, one might think that the abilities that are manifested might be construed quite generally, abstracted from any particular object game, like baseball or cooking. If, for example, the coach’s instructions are given in some language, then the abilities necessary for rule-following in this particular case would involve the ability to understand this language. If implementing the coach’s instructions involves recognizing the domains in which they apply, then the ability to follow a rule here would include the appropriate conceptual abilities.

The general abilities that are manifested in rule-following are those that are involved in implementing one’s understanding of a rule in action, in judgment, or in speech. These will include whatever conceptual, perceptual, interpretive, and linguistic abilities are needed to grasp the rule and to understand the situations in which it applies and whatever inferential and rational abilities are needed in order to act in the light of one’s understanding of it. Rule-following abilities include, that is, those abilities that are implicated in our ordinary notions of meaning, understanding, thinking, reasoning, and acting. Now, no one has explicitly attempted to give a cognitive explanation of the ability to follow a rule or play the rule-following game. However, many have attempted to provide systematic accounts (or theories) of language, rational thought, and action; indeed, this has been one of the central tasks of analytic philosophy. One of the central tasks has been, in other words, to give accounts of the general abilities that turn out to be, I am arguing, constitutive of, or part of what it is involved in, rule-following. Those who attempt to give cognitive explanations of any of these abilities broadly construed might as well be trying to offer cognitive explanations of the ability to play the rule-following game.

I suspect that there might be a problem in attempting to construe the constituent abilities that are involved in rule-following in a very general sort of way—in a way that abstracts from their role in an object game.⁵ I won’t pursue this line here. Whether or not it is a problem, it certainly makes no sense to attribute knowledge of the putative norms that govern these activities to an individual as part of a substantive explanation of her ability to participate in them. It makes sense to explain an individual’s material move in baseball (say, the batter’s merely tapping the ball with the bat and thus inviting the

⁵ See, for example, §§28, 29 and 43–49 in Wittgenstein, *On Certainty*, Anscombe and von Wright (eds) (Oxford: Basil Blackwell, 1960).

Julia Tanney

pitcher to throw her out at first) by attributing to her knowledge of some or other rules of baseball (e.g., about sacrifice strategies). But extending this kind of explanation to the case in which the material moves are the (generally and abstractly construed) constituent abilities involved in rule-following is incoherent. So I shall argue.

3. Why one cannot follow rules that govern the rule-following game

We were motivated to speak of someone's following a rule in the first place in order to discern some kind of internal connection between the individual who makes the moves and the rules that govern them. We were seeking this connection to assure us the individual whose behaviour accords with the rules is a participant in the practice (to distinguish her from someone who is able to make the moves, but only accidentally, or to distinguish a player who makes a mistake from a non-player). One way of making this connection, we decided, would be if the individual were able to appeal to the rules in justification, criticism, or correction of her moves.

But it makes no sense to attribute knowledge to an individual of the putative rules governing the general activities that are presupposed in rule-following in order to explain her ability to participate in these activities. This is because following a rule for any of the activities required in order to follow a rule presupposes the very abilities that the knowledge is supposed to explain. If someone did not know how to follow a rule, then showing her another rule would not help, since presumably if she lacked the ability to follow a rule, she would not know how to follow this other rule either. If, however, she knows how to follow a rule, then there is not any point in attributing to her knowledge of another rule to explain this ability, for certainly if an explanation were required in the first place, it would be required for the same ability that is required for the explanation to get underway.

In particular, if an individual does not know how to interpret symbols then showing her meaning theorems can be of no help until it is explained how any symbols in which the rules are delivered are themselves understood or until it is explained how she comes to be aware of the demands that they make.⁶ If an individual does not know how to make inferences then introducing her to inference rules can be of no help until it is explained how these rules of inference are to be applied.⁷ If she does not know how to act for reasons,

⁶ See Sellars 'Some Reflections on Language Games', *op. cit.*

⁷ A related worry (about justification) is raised in Lewis Carroll's 'What the Tortoise Said to Achilles', *Mind*, vol. 4 (1895), 278–80; reprinted in *Mind*, vol. 104, no. 416 (1995) pp. 691–93.

Playing the Rule-following Game

then invoking principles of rationality can be of no help until it is explained how the principles of rationality are to be acted upon.⁸ And if she does not know how to self-ascribe mental concepts, then having knowledge of a theory of mind can be of no help until it is explained precisely how she is to use the theorems to guide her.⁹ And so forth. If, however, she already has the ability to interpret symbols, make inferences, self-ascribe, or act rationally, then there does not seem to be any point in attributing to her knowledge of putative rules governing these abilities, because her ability to follow these rules would presuppose the very abilities that are to be explained.¹⁰

We have a puzzle. The ability to speak and understand a language, to understand oneself and others, and to act rationally are abilities that seem to demand some sort of explanation. These abilities—which are presupposed in rule-following, in so far as they involve implementing one's understanding of a rule in action—seem to be norm-governed and yet do not seem to be governed by the same rules that govern any particular object game. One might thus attempt to abstract them from any particular object game, attempt to find the rules that govern them, and then use these rules as part of a cognitive explanation of the generally construed abilities. But we have just seen that if there are a distinct set of standards that govern these abilities they cannot simply be invoked as part of a psychological explanation of those abilities. It makes no sense, that is, to effect the same kind of cognitive-explanatory relation here that we sought earlier between the person who is able to make the moves in the object game of baseball and the rules that govern those moves because of the regresses that threaten.

It might be useful to compare this regress—which is vicious—with another kind of regress that is not especially problematic.

⁸ This idea is developed further in part III and also in my 'De-individualizing norms of rationality' *Philosophical Studies* 79 (1995), 237–58.

⁹ This idea is developed in my 'Understanding oneself, understanding others', ms.

¹⁰ I doubt that these abilities are in any interesting way separable; I list them all in order to draw attention to the scope of the claim being made. Perhaps some of the confusion about what kinds of explanations are possible arises because some theorists *do* think that some of these abilities can be presupposed in order to 'explain' others (e.g., that inferential abilities can be presupposed in explaining linguistic abilities). I doubt that this is coherent. For a related discussion, see Dummett's claim that a meaning theory must tell us all that is involved in speaking a language (which, he claims, is *the* rational activity *par excellence*) in (most recently) Dummett, *The Logical Basis of Metaphysics* (Cambridge, MA: Harvard University Press, 1991) and John McDowell's response in 'In defense of modesty' in B. M. Taylor (ed.) *Michael Dummett* (Dordrecht: Martinus Nijhoff, 1987).

Consider, for example, that in playing baseball one uses skills that require other skills or abilities (e.g. the ability to hit the ball requires the ability to stand up or to grasp the bat). This ‘layering’ of skills is not a problem because the skills required are different from one another (and probably ‘bottom-out’ with abilities that do not need to be learned or with abilities that do not require the obeying of rules). But the problem with the skills involved in rule-following is that they are the very skills that are constitutive of cognitive explanations. If it even makes sense to talk about rules that govern rule-following, following these rules would involve *the very same abilities* that the rules were introduced to govern. There is thus no possibility that knowledge of the rules could serve as an explanation of these abilities.

4. Innate, implicit, or tacit knowledge

I have given a simple argument to show that it does not make sense to attribute to an individual knowledge of rules in order to effect an internal connection between the moves she makes in following a rule, and whatever, if any, standards govern those moves. I have suggested that this presents a challenge to those interested in giving cognitive explanations of our broadly rational (including linguistic) abilities. Now, those hoping to offer psychological explanations of these abilities will complain that I have placed too heavy a burden on what is to count as knowledge. They will agree that knowledge of the standards that govern these abilities cannot be *explicitly* represented in the sense that they can be consciously consulted. But they will insist that it is possible to appeal to innate, implicit, or tacit knowledge in trying to give an explanation of them.¹¹

Now, describing knowledge of the ‘rule-following rules’ as *innate* would seem to solve the problem of how the abilities that are presupposed in rule-following might be learned. It might be agreed that they cannot be learned for reasons given above: to learn them one would have to presuppose the very same abilities that needed learning. If we give up the idea that these rules need to be learned, and accept that they are intrinsically part of our makeup, we could avoid this problem. But our problem is not merely that it is impossible to *learn* how to follow a rule by being taught these other rules. The problem is that knowledge of these other rules cannot explain

¹¹ Perhaps the primary motivation for supposing that the theoretical knowledge that purportedly explains the abilities is tacit or implicit is simply the recognition of the obvious fact that the knowledge is not something that the individual herself is usually able to articulate. Here, I explore whether invoking tacit or implicit knowledge can, in any case, avoid the regress threats described above.

Playing the Rule-following Game

the kind of abilities that are involved in rule-following. The arguments show not merely that the acquisition of this knowledge is problematic, but that the *utilization* of it (even if it is innate) would involve presupposing the very abilities that the knowledge was supposed to explain.

One might be tempted to avoid the problem by talking about the possibility of 'implicit' or 'tacit' knowledge.¹² The rules governing the kinds of abilities that are manifested in rule-following would be 'cognitively grasped', yet not such that they are available for linguistic expression or conscious manipulation. Now, since the argument above is that there is nothing for knowledge of this sort to explain, it does not really matter *what* the vehicle of knowledge is supposed to be, whether or not it is accessible to consciousness, or what systemic/organizational role it plays. The problem that I am trying to call attention to does not have anything to do with the nature of the 'vehicle'. The problem is the explanatory poverty of putting the individual in any kind of cognitive relation to a 'content' or to a norm that governs an ability that is itself required for rule-following. That is the quick response. Still, a more detailed look might be in order to see more precisely how the appeal to tacit knowledge, in particular, goes wrong.

Note that if there is to be any substance to the claim that knowledge of the norms plays an explanatory role, then this 'explanatory role' cannot simply collapse into a description of what would constitute a correct performance. In particular, unless we had evidence that the agent sub-cognitively 'grasps' the norm and evidence that it is in virtue of her grasping it that she acts in accordance with it,

¹² The normal understanding is that implicit knowledge is available to consciousness once it has been made explicit and that tacit knowledge is not normally accessible to consciousness at all. Dennett suggests using 'implicit representation' in a different sense: as information that is logically implied by something that is stored explicitly. (See 'Styles of mental representation'. *Proceedings of the Aristotelian Society*, **83**, 213–6; reprinted in *The Intentional Stance* (Cambridge, MA: A Bradford Book, MIT Press, 1987)). Both versions make the notion of implicit knowledge dependent on explicit knowledge. For discussions of tacit knowledge, see Gareth Evans, 'Semantic theory and tacit knowledge'. In S. Holtzman and C. Leich (eds) *Wittgenstein: To Follow a Rule* (Routledge and Kegan Paul, 1981); reprinted in his *Collected Papers* (Oxford University Press, 1985), 322–42. See also Martin Davies, 'Tacit knowledge, and the structure of thought and language. In C. Travis (ed.) *Meaning and Interpretation* (Blackwell, 1986); 127–58; 'Tacit knowledge and semantic theory: Can a five per cent difference matter?' *Mind* **96**, 441–62; and 'Tacit knowledge and sub-doxastic states' in A. George (ed.) *Reflections on Chomsky* (Blackwell, 1989); 131–52.

we would not be able to distinguish between her merely acting in accordance with the norms and her having the critical or justificatory abilities that tacit knowledge of the norms would ostensibly explain. The methodological difficulty of deciding which non-linguistic behaviour is to count as 'justificatory' or 'correction' behaviour should not be underestimated.¹³ But even if such a difficulty could be surmounted, it should be evident that the same regress problems would arise. This is because the abilities will involve 'sub-conceptualizing' the norm, 'sub-conceptualizing' the stretch of behaviour to which it applies, 'sub-judging' its applicability, and 'sub-acting' in the light of it. To perform these tasks *correctly* presupposes the very same abilities, operating sub-cognitively, that these tacit representations are, on the hypothesis we are considering, posited to explain.

Someone might be tempted to respond that, at the level of tacit representations, the question of (in)correct manifestations of abilities does not arise. She might suggest that 'tacit knowledge' should be understood simply as whatever could be read off a person's dispositions under certain circumstances, where the circumstances themselves are circumscribed in such a way as to ensure that error will be ruled out. Thus, she might argue, the standards governing the abilities that are manifested in rule-following behaviour might be seen as recoverable from, or embodied in, laws of human psychology or as part of the 'hardwiring' of individuals who act in accordance with them.

But this picture is confused. On the one hand, it wants to include the role of a *norm*, whose whole *raison d'être* was allow us to make (logical) space for moves that are in violation of it, and yet are still moves made within the activity or practice that these norms in some sense govern. On the other hand, it attempts to force the opposite picture so

¹³ If the notion is coherent, we ought to be able to say what could count as evidence that a person sub-cognitively conceptualizes the norm and guides her behaviour in the light of it. But how would we distinguish this from the case where she mis-conceptualizes the norm, but acts in accordance with it none the less? And how would we distinguish, for example, the case where she succeeds in conceptualizing the norm yet fails to implement her understanding of it in action from the case where she simply fails to act in accordance with it? The difficulty in answering these questions puts pressure on the very coherence of the notion of 'sub-cognitive conceptualization'. (See W. V. Quine, 'Methodological Reflections on Current Linguistic Theory', *Synthese* 21 (1970), 393 for similar doubts.) It is arguable that these sorts of distinctions are at least *prima facie* plausible when 'knowledge of the norm' is understood as shorthand for a person's ability to defend, justify or correct herself by citing a rule. But the linguistic ability that these abilities would require depends upon explicit knowledge and, of course presupposes, and thus cannot figure in an explanation of, the ability to follow a rule.

Playing the Rule-following Game

that the existence of the norm, or knowledge of it, somehow determines that the moves will be in accord. In specifying the conditions under which the dispositions are to count as manifesting tacit knowledge, the theorist would be ruling out, for example, the cases in which someone fails to act in accordance with the norm. But this would be ruling out too much. We introduced norms in the first place, not only to distinguish individuals—as a player or non-player—both of whom act in accordance with the rules, but also to distinguish individuals—as a player or non-player—both of whom fail to act in accordance with a norm. The suggestion never was that if we could attribute knowledge of the norms to the individual we would thereby ensure that she was playing the game correctly, or even playing the game at all; it was merely that if we could attribute to her knowledge of the norms, we might have at least a *prima facie* case that she was playing it.

To sum up, the cognitivist faces a dilemma. On one horn of the dilemma, the subject manifests behaviour that can be described as 'correct' or 'incorrect', in which case trying to explain her ability to make correct moves by attributing to her knowledge of a putative rule governing her behaviour goes badly wrong when the moves involved are moves within the game of rule-following. A vicious regress ensues *whether the knowledge is attributed to the person or to one of her subsystems*. On the second horn of the dilemma, the subject does not manifest behaviour that in any relevant sense can go wrong. In this case it is not clear why we would invoke a standard in the first place, let alone attribute knowledge of it to her. This dilemma would seem to threaten attempts to provide a cognitive explanation of any of the general abilities that are constitutive of rule-following.

In fact, rule-following explanations will not even get off the ground unless we can presuppose that the individual to whom the knowledge is attributed knows what to do with this knowledge; unless we can presuppose, that is, that she has the conceptual, inferential, cognitive, linguistic, or in short, the rational abilities that following a rule requires. But the arguments above show that any attempt to give cognitive explanations of these presupposed abilities themselves is doomed to failure precisely *because* they are presupposed. These reflections ought to raise *prima facie* doubts about the coherence of offering cognitive explanations of the ability to understand or speak a language or the ability to understand and predict rational behaviour.¹⁴ In the following section, I'll explore in more

¹⁴ There are a number of *prima facie* targets, including:

(1) Those philosophers who suggest that a fruitful task in philosophy of language would be the construction of a theory of meaning knowledge of which would (suffice to) explain a person's ability to understand and speak a language. I suspect that the discussion in this paper uncovers at least a

detail how these reflections might be used to challenge the idea that principles of rationality might be invoked as part of a cognitive explanation of rational action and how this affects certain contemporary accounts of what is involved in acting for reasons.

prima facie tension in some of the requirements that Dummett, for example, places on a theory of meaning. I have in mind the constraints that a theory of meaning should 'describe, without making any presuppositions, what it is that we learn when we learn to speak' (p. 91, *LBT*) and that the knowledge ascribed to the speakers is genuine, propositional knowledge and not a mere theoretical representation of a practical ability. These suspicions, of course, have to be examined in detail. See Dummett, 'What is a theory of meaning?' (II), in Hotzmann and Leich (eds) *Wittgenstein: To Follow a Rule* (London: Routledge & Kegan Paul, 1976), 99–137; Dummett, *The Logical Basis of Metaphysics*, op. cit. and John McDowell, 'In defense of modesty' op.cit. For a discussions about the possible difficulties with theories of meaning in the light of Wittgenstein's reflections on rules, see Crispin Wright, 'How can the theory of meaning be a philosophical project?' *Mind and Language*, **1**, (Spring, 1986), 31–44; 'Theories of meaning and speaker's knowledge' in *Realism, Meaning and Truth* (Oxford: Basil Blackwell, 1986); and 'Wittgenstein's rule-following considerations and the Central Project of Theoretical Linguistics' in A. George, (ed.) *Reflections on Chomsky*, op. cit.

(2) Linguists and psychologists, following Chomsky, who believe that it is the task of theoretical linguistics to formulate a grammar (a set of rules or principles) the tacit knowledge of which would explain a speaker's competence. See, for example, *Reflections on Language* (Fontana/Collins, 1976); *Rules and Representations* (Blackwell, 1980); see also A. George (ed) *Reflections on Chomsky*, op. cit., for philosophical discussions surrounding the general theme.

(3) A corresponding group of theorists from philosophy, experimental and developmental psychology and cognitive anthropology who have suggested attributing to adults a theory of mind in order to explain their ability to understand and predict rational behaviour. See, for example, Davies and Stone (eds) *Folk Psychology—The Theory of Mind Debate* (Blackwell, 1995); Astington, Harris, and Olson, (eds) *Developing Theories of Mind*, (Cambridge University Press, 1989); Andrew Whiten (ed.) *Natural Theories of Mind* (Blackwell, 1991).

(4) Philosophical discussions in self-knowledge that suppose that a person's special access to her own mind genuinely explains her ability to self-ascribe mental states. I discuss some of the issues related to this in 'A constructivist account of self-knowledge', *Philosophy*, **71**, (1996), 405–22 and in 'Understanding oneself, understanding others' (ms.). Paul Boghossian and Crispin Wright both note and take on board the regress threat posed to traditional accounts of self-knowledge. See Wright, 'Wittgenstein's Later Philosophy of Mind: Sensation, Privacy, Intention', *Journal of Philosophy*, **86**, 11 (1989), 622–35; extended version in *Meaning Scepticism*, Puhl, Klaus (ed.), (de Gruyter, 1991), 126–47 and Boghossian 'Content and Self-knowledge', *Philosophical Topics* **xvii** (1989), 5–25.

III. Reason explanation

1. Back to baseball

It was suggested that attributing to an individual knowledge of the rules that govern some practice might figure in an explanation of the person's ability to participate in the practice, but it is not at all clear *how* such an explanation would work. We wanted to seek some kind of internal connection between the rules of baseball and the person who makes the material moves so that we could say that she is *making the moves as part of the game of baseball*. We agreed that we might be satisfied if she were able to cite the rules of baseball at the appropriate time, say, in an attempt to justify a material move. And we agreed that perhaps there would be nothing objectionable in describing this as a type of knowledge: someone who is able to cite the rules of baseball at the appropriate time might be said to *know* the rules. Then it was suggested that one's knowledge of the rules of baseball might figure in an explanation of one's ability to play the game.

Someone might object that *if* we require that the ability to justify or correct one's moves in the light of the rules is necessary for one to be considered a real baseball player, then at best, this sub-ability, or knowledge how to appeal to the rules to justify, etc. would be part of a complete description of the (full-fledged) ability to play the game, which would partly *consist* in the individual's sub-ability to appeal to the rules. Still, it is easy to see how an individual's knowledge of the rules might figure in an account of her ability to play the game. It might be supposed, for example, that knowledge of what the game requires would explain *why* the player made a certain material move. Her having this knowledge might be part of a reason-explanation for making the move.

Suppose that reason-explanation works in virtue of there being a logical or justificatory relation between a sentence describing the action and the sentences describing the beliefs and desires that are attributable to the agent.¹⁵ We could say the individual we are

¹⁵ This version of what is involved in reason-explanation is accepted by most of the people whose views I go on to criticize, although it seems to need reconsideration in light of the conclusions of this section. The important point is that on this model (unless it is supplemented with causation) a reason does not determine, or provide a sufficient condition for, the action that it rationalizes. Note that nothing about the ensuing argument will change significantly if values, judgments, and intentions (or statements expressing them) are added to the model of reason-explanation. Because on this more complex model, either intentions do not determine actions (see Pears, David *Motivated Irrationality* (Oxford: Clarendon Press, 1984)) or all things considered judgments do not determine intentions (see Davidson,

Julia Tanney

considering makes a particular move (she runs to first base) because she wants to play the game of baseball and she knows that the rules require her to run to first base under certain circumstances (and she knows that these are the circumstances, etc.) This might give us at least one way of cementing the internal connection we are looking for.

2. Having reasons for acting reasonably.

Next we noticed that just as one might sometimes make mistakes in playing baseball, so, too, might one sometimes make mistakes in following its rules. This led us to suppose that one might sometimes make mistakes in following rules generally construed, abstracted from any particular object game like baseball. One might, for example, misperceive, misintuit, or misread a symbol expressing a rule, one might mistake a situation as one that falls (or does not) within the domain of the rule, and one might fail to act in accordance with the rule, even if one understood it and its domain properly. This invited us to construe the abilities that are involved in rule-following on the analogy of a 'game'; i.e., as activities that are themselves rule-governed.

However, even though it seems to make sense to say that someone might make a mistake in following these rules, and that avoiding this sort of error might be considered an achievement, it is difficult to see how this could be understood as a cognitive or a rational achievement. On the model we are presently considering, knowledge would be attributed as part of a reason-explanation for making the material moves, and the material moves we are now considering are the moves made within the game of rule-following. The motivation for attributing reasons here would be to effect an internal connection that would allow us to see the moves an individual makes within the rule-following game as reasonable instead of as moves made merely by accident. According to the hypothesis we are considering these moves would be reasonable only if the individual has reasons for playing the rule-following game.

But rule-following, under the proposal we are considering, *is* acting for reasons. Thus, to generalize the abilities constitutive of rule-following, and to construe these general abilities themselves as subject to standards would require us to attribute reasons in order to see the moves an individual makes within the *game of reason* as themselves reasonable instead of moves made merely by accident. According to this proposal, an individual's moves would only be *reasonable* if she

'How is Weakness of the Will Possible?' in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980)). It is this gap that reason-explanation exploits; it is this gap that makes irrationality possible.

Playing the Rule-following Game

has reasons for acting reasonably. Perhaps this would require the introduction of 'norms of rationality': a person might be said to have reasons for acting reasonably only if she knows what the norms of rationality require and she wants to act in accordance with them. Indeed, some have supposed that our psychological/interpretive practices are governed by norms or principles of rationality and some have argued that these norms are contained in a theory, knowledge of which might explain the ability to participate in these practices.¹⁶

Knowledge of these ostensible principles, however, could not figure in a substantive explanation of a person's ability to act rationally. This line of approach, like the approach in part II, also invites a regress. The regress is vicious because the moves the individual makes that ostensibly needed explaining by attributing to her knowledge of the rules of reason are exactly the same type of moves that she would be making if she were to follow these rules of reason. So, again, the kind of moves that were thought to need explaining are presupposed in the very kind of explanation on offer. If it were necessary to make out an internal connection in order to obviate mere accidental rule-conformity in the explanandum, then it would be necessary to make out an internal connection in order to obviate mere accidental rule-conformity in the explanans since the moves are of exactly the same type.

We seem to have gone wrong in supposing that we need to attribute reasons for acting reasonably to a person in order to see her actions as reasonable. *If* we must effect an internal connection between reasons and behaviour, then being able to describe a person's behaviour as an action (let alone in a way that displays the right kind of logical connection with statements expressing her beliefs and desires) *is* to effect it; we do not *also* need to see her actions as acquiring an independent status *as reasonable* by attributing to her knowledge that governs how to reason reasonably.¹⁷

¹⁶ See Davidson's discussions of interpretation in his *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984) and the large amount of literature discussing his interpretation strategy. See also Davidson's discussions of irrationality; in particular, 'Paradoxes of Irrationality', in R. Wollheim and J. Hopkins (eds), *Philosophical Essays on Freud* (Cambridge University Press, 1982) where he introduces the 'principle of continence' as an example of a norm of rationality and Annette Baier's response in 'Rhyme and Reason: Reflections on Davidson's Version of Having Reasons', in Lepore and B. McLaughlin (eds) *Actions and Events—Perspectives on the Philosophy of Donald Davidson* (Oxford: Blackwell, 1985). For discussions involving knowledge of a theory of mind, see, for example, M. Davies and T. Stone, (eds) *Folk Psychology*, op. cit. and M. Davies and T. Stone, (eds) *Mental Simulation* (Oxford: Blackwells, 1995).

¹⁷ See 'De-individualizing norms of rationality', op. cit.

3. The limits of reason-explanation

In response, one might claim that this line of argument shows the impossibility of giving cognitive explanations for rational abilities and their kin only if it assumed that cognitive explanations are exhausted by reason-explanation, and only if it assumed that reason-explanation fits the model just described. On this model, there is a logical gap between what the agent has reason to do and what she will do. And, of course, if there is the logical latitude to fail to act in accordance with one's reasons, then attributing reasons for acting reasonably will not work, since this, too, involves a logical gap. But, so the response continues, this does not mean that we cannot explain a person's ability to act rationally on a different model. We could, for example, make scientific hypotheses concerning the laws in accordance with which she operates.

Indeed, the response might continue, if reason-explanation were to be truly explanatory, we would have to extend its domain by introducing laws or something like them into the picture anyway. We would have to do this since it has just been shown that reason-explanation has to make certain presuppositions (that a person acts rationally) *and* it has been acknowledged that a person's rational abilities might, themselves, go wrong. That they do not go wrong in a particular case cries out for explanation. It cannot be a kind of explanation that itself leaves room for error or regress ensues. So rather than attempt to explain the person's rational action with higher-order rules, we might cement the connection between reason and action with causation. On certain widely accepted assumptions about the nature of causal explanation, we would thereby obtain explanation involving nomological subsumption.

Providing this sort of relation between reason and action, however, destroys the possibility of error, and it was this possibility that encouraged us to adopt the analogy of a game. Without the analogy of the game, we have lost the reason for seeking an internal connection, and with it, we have lost any reason to attribute knowledge of the reasons.¹⁸ To see this, consider that the possibility of psychological laws (strict or not) depends on the possibility of true, non-empty sentences expressing generalizations couched in psychological terms. Of course, rational people do not always act rationally and thus any putative psychological law will not always hold. The common response to this is to argue that irrationality is not a problem. For just as we attempt to specify natural laws by circumscribing the

¹⁸ In 'How to resist mental representations' (a critical notice of Tim Crane's *The Mechanical Mind*), *International Journal of Philosophical Studies* vol. 6 (2), (1998) I attempt to develop this argument in a way that challenges the existence of content-bearing, causally-efficacious states.

Playing the Rule-following Game

set of conditions under which they hold, we might do the same for psychological laws. Among the conditions to be ruled out are the circumstances in which a person is irrational. Of course, these conditions would have to be specified independently—or in a way that does not reintroduce the laws themselves—in order to come up with truths that are non-empty.

But even if it were possible independently to specify the conditions under which a person is, for example, self-deceived or akratic so that exceptionless laws or generalizations could be produced, ruling out cases of irrationality would be ruling out too much. For when a person acts *irrationally* as opposed to non- or a-rationally, she is acting in a way that is, at least to some extent, rational. Or, returning to the analogy of a game, she is making a move in the 'game' of action that can be described as recalcitrant (or 'incorrect') in the light of some or other norm of rationality. It is still, however, a move within the game of action, just as an *error* is a move within the game of baseball. But if the psychological laws or generalizations are meant to apply only when someone acts rationally as opposed to irrationally, or only to the cases in which someone acts *correctly* in making moves in the game of action, then a major subset of the moves we consider to be moves made within the game of action will have been left out of the account. In any case, unless error of this kind is possible, there is no reason to construe rational moves as norm-governed as opposed to law-subsumed or pattern-instantiating and thus no reason to suppose cognitive grasp of the norms is necessary to explain anything.

The upshot is this. First, a person's acting rationally cannot be explained by attributing to her knowledge of higher-order reasons (norms of rationality) in the service of providing a reason-explanation of her reasonable action. Second, it cannot be explained by adverting to laws if they are thought to provide a determinate connection between an agent's reasons and her actions.¹⁹ The appeal to

¹⁹ *A fortiori* it cannot be explained by the obtaining of causal relations between mental events, if this relation is supposed to cement the logical gap that exists between an agent's reasons and her actions. I suggest elsewhere that this motivation lies behind Davidson's claim that something is missing in reason explanation if we consider only the purely justificatory relation between reasons and action and that this intuition motivated his introduction of a 'causal element' into his account of reason explanation. I also argue that his doctrine of anomalous monism—which commits him to the view that the only way to introduce laws into an account of mental causation is via a physical description of mental event-particulars—does not protect him from the charge that in buying into a nomological account of psychological explanation (however the laws are described) he is not leaving space for the possibility of error. See 'Why reasons may not be

higher-order reasons leads to a vicious regress and the appeal to laws that subsume the causal processes leads to the dissolution of the normative nature of the phenomenon to be explained. These considerations suggest that there is something wrong with the idea that reason-explanation is deficient unless it provides us a way of securing a law-like connection between what the agent believes, desires and does (or statements describing this). Reason-explanation was never supposed to provide us with the wherewithal to make strict predictions about what a person will do, because even if we know everything there is to know about what she wants and believes (which certainly is not required for reason-explanation) she still might act impetuously (non-rationally) or for other reasons she values less (irrationally). Reason-explanation was meant to provide us a way of understanding her behaviour by showing how it would make sense in light of certain facts. But it does not (and should not) purport to tell us that she *will* behave in a way that is understandable in light of these facts.²⁰

IV. Reconsidering Attributions of Knowledge

I. Baseball again

Perhaps we should reconsider the game of baseball to see why it was thought that attributing knowledge of the rules was necessary in the first place.²¹ Again, we were seeking an internal connection between

²⁰ I have not distinguished explicitly between different kinds of standards. Dworkin (in 'Is law a system of rules?' in R. M. Dworkin, (ed.), *The Philosophy of Law* (London: Oxford University Press, 1977)), makes a logical distinction between rules and principles. Rules dictate or determine results even though they admit of exceptions; an accurate statement of the rule would take these exceptions into account. If a contrary result is reached, then the rule must have been abandoned or changed. Principles are like reasons: they incline a decision one way, though not conclusively. They do not necessitate a particular result, and they survive intact when they do not prevail. The discussion above, then, might be construed as suggesting that the norms that govern our broadly rational practices are principles instead of rules.

²¹ It has been assumed that the rule-follower (as opposed to the mere rule-conformist) must grasp the rule, understand the situation in which it applies and act in the light of her understanding of it. Rather than exploring in more detail what it would take to manifest rule-following behaviour (a central case of which occurs when one is able to cite a rule in the justification or correction of a 'material' move) the discussion has centred on causes', *Mind & Language*, **10**, nos. 1 and 2 (1995), 105–28. Note that nothing is materially changed in the argument if the 'determinate connection' is thought to hold between overriding reason and action.

Playing the Rule-following Game

the rules and the person who makes the moves that accord with them so that we could say that she is making the moves as part of the game of baseball. This was because it was thought possible for someone to make the moves of baseball (say, run to first base) even though she knew nothing about the game and should not be credited with playing it. We thought that we might be satisfied that the performance is not accidental if the player has the ability to cite the rules of baseball, in an attempt to justify or correct her moves. It was suggested that perhaps there would be nothing objectionable in describing this internal connection as a type of knowledge: someone who is able to cite the rules of baseball at the appropriate time might be said to know the rules. It was then suggested that attributing knowledge of the rules of baseball to someone might figure in part of an explanation of that person's ability to play the game. This seemed reasonable if the knowledge was to figure as part of a reason-explanation for her making a particular move. All of this seemed straightforward in the case of baseball, but we ran into trouble when we considered attributing knowledge as part of an explanation of the abilities involved in the game of rule-following itself because these abilities are presupposed. But this criticism relates only to the attempt to give a cognitive explanation of the abilities that are presupposed in rule-following. Where does this leave cognitive explanations in general, and explanations involving implicit, or tacit knowledge, in particular?

Let us look at baseball one more time. In allowing that someone's ability to justify or correct her moves might warrant the attribution of knowledge, we agreed that all we meant by this is that she knows how to appeal to the rules to justify or criticize certain moves. So, *if* we decide more is needed to establish that a person is playing the game of baseball than the fact that she acts in accordance with its rules and *if* we decide that this 'something more' can be supplied by attributing to her knowledge of the rules, then all we are entitled to say so far is that such knowledge (knowledge how to appeal to the rules to justify, etc.) would be part of a description of the full-fledged ability to play the game.

Someone might argue that such knowledge is not necessary for playing the game, if what we mean to require is that she be able to cite the rules and if this requires that she have certain linguistic abilities. It would be enough, someone might argue, to credit her with playing by the rules, even if she had no linguistic abilities. Surely we could imagine a pre-linguistic child playing baseball; or, the circumstances in which it is necessary to invoke the distinction between rule-conformity and rule-following in order to credit an individual with making a legitimate move.

Julia Tanney

better, think of the pre-linguistic children who play chess. It is very tempting to slide from this to the conclusion that pre-linguistic children who play chess have implicit knowledge of the rules of the game.

But if we acknowledge the propriety of describing pre-linguistic children as playing baseball or chess, this gives us a reason to retract the requirement that the child have the ability to cite the rules in justification, etc., of her moves. Since this is what we meant by 'knowledge of the rules', it gives us a reason to retract the requirement that she have this knowledge. Perhaps we might still seek to assure ourselves that it was no accident that the child was playing in accordance with the rules of baseball or chess. We might rule out accidental performance if the child is able to repeat her successful performance. Or we might rule out accidental performance if the child has been trained by someone who was able to play the metagame (someone with linguistic abilities who was able to cite the rules in justification and criticism), or by someone who herself was trained by such a person. But in imagining this, we have ruled out accidental rule-conformity without attributing knowledge of the rules to the individual who makes the moves. Again, the fact that we do not always require linguistic-dependent justificatory abilities of someone in order to construe her as a participant in some rule-governed activity gives us a reason to retract the requirement that she know the rules—not to attribute the rules to her as the content of implicit knowledge.

Still, it might be argued that all we *mean* when we attribute implicit knowledge of the rules to someone is that she acts in accordance with the rules and this is no accident since she was trained to act in accordance with them or because she is able to repeat her performance. So far, there is still nothing wrong with describing her as having implicit knowledge. Note, however, that training is not necessary for someone to be considered a participant in some rule-governed activity. (In order to bring this point home we will have to change the example). Somebody might be able to solve Rubik's cube, even if she had not been trained by anybody. Indeed, someone might be able to solve the puzzle even if she could not repeat the performance. It is very tempting to slide from this to the conclusion that the person has implicit knowledge of the rules for solving Rubik's Cube, where 'implicit knowledge' cannot simply refer to an ability that has come about as a result of training or to the fact that successful performances can be repeated.

But if we agree with the thought that someone might be able to solve Rubik's Cube even if she never had been trained by anyone, then this gives us a reason to reject the idea that there must be an

Playing the Rule-following Game

internal connection between the rules that govern an activity and the individual who makes the moves. We can say that it is sometimes enough to credit someone with playing the game if she acts in accordance with the rules. Knowledge (implicit or otherwise) has dropped out of the picture. To insist that someone cannot solve the puzzle unless she somehow conceives the rules (even if she cannot articulate them, even to herself) and acts in the light of her conception of the rules is simply dogmatic. What would justify such insistence? If this person were suddenly entered in a contest and produced the cube with the colours in the right places, we would not withhold the prize because she merely acted in accordance with, but did not follow, the rules. Acting in accordance with the rules *is* solving the puzzle in certain cases. Of course, the story could be filled in to make it plausible that she was not solving the puzzle at all; say she was colour blind and had the habit of fidgeting nervously with objects of any kind. This might give us a reason for requiring of *her*, or others like her, more than just the ability to act in accordance with the rules. Although it would be in order to state what additional abilities would be required in particular circumstances such as these, it would not be in order to import the requirement under the guise of ‘implicit knowledge of the rules’ to other circumstances in which the manifestation of such additional abilities would not be required.

To insist that someone must conceive the rules *somehow*—even if what it would be for her to conceive these rules is inaccessible to us—is misguided; it fails to explain anything. Recall that one reason for supposing that ‘grasping the rules’ was necessary was to allow us to rule out the possibility that an individual who acts in accordance with the rules does so accidentally. But if she conceives them incorrectly then she probably will not act in accordance with them. And yet if she cannot help but conceive and act in accordance with them correctly, then the possibility of mere accidental rule-conformity would be blocked from the beginning, and again there would be no point in attributing to her knowledge of the rules.

Sometimes, in certain circumstances, we might require more than a mere ability to act in accordance with the rules that govern a practice in order to rule-out mere accidental conformity. We might be satisfied if the individual in question is able to play the metagame and cite the appropriate rule in the appropriate circumstances (think about exams in school). But this is not always the case. Somebody could have the ability to cite the rule in the appropriate circumstances even if she did not have the ability for which citing the rule was supposed to be evidence (suppose she cheated on an exam). Sometimes we do not need explicitly to rule out the

Julia Tanney

possibility of accidental conformity. Sometimes the circumstances in which rule-conformity is in question are those in which it is presupposed that the possibility of accidental conformity is ruled-out (think about contests involving skill).

These reflections suggest that the conditions under which we credit an individual with participating in a norm-governed practice vary. When the practice is thought, action, or language the conditions under which we credit an individual with *thinking*, *intending*, *understanding*, or *meaning* vary. The criteria for applying these concepts do not remain fixed when the circumstances surrounding their application change. Problems arise when this is forgotten. For instance, if one imagines that in order to be credited with meaning something by a sign, one must have relevant justificatory abilities, then in the cases where no obvious examples of this occur in the normal way through language use, one might be tempted to describe these as cases where something like language use is going on prelinguistically none the less. If one imagines that to be credited with thinking, one must have the kinds of abilities that are evinced when one deliberates consciously, then one might be tempted to describe cases in which thinking seems to occur (without deliberation) as cases of unconscious or tacit deliberation. I have suggested in part IV that in many cases involving the attribution of tacit knowledge it will be sufficient to describe the achievement as one that accords with a rule. It might not be necessary in some of these cases to suppose that the rule is internalized, conceptualized, and followed by the one whose behaviour conforms to it. I have argued in parts II and III that for certain abilities—those that are presupposed by the kind of explanation on offer—the supposition is not even coherent. Arguments like this ought to give some support to the view that epistemic or rational norms are part of the ‘bedrock’.²²

University of Kent at Canterbury

²² I have read versions of this paper for the philosophy societies at King’s College, London; Trinity College, Dublin; University of Bristol; University of Kent; and University of Sussex. The first half was presented at the European Society for Philosophy and Psychology in Barcelona and the University of Hertfordshire. I received many valuable comments on these occasions. In particular I would like to thank Helene Buerger, John Flower, Simon Glendinning, David Pears, Tony Skillen, Helen Steward, Crispin Wright, and the Editor of this journal for their comments. Finally, I would like to express my gratitude for the research leave granted to me by AHRB, during which I completed the final version of this paper.