The Exemplification of Rules: An Appraisal of Pettit's Approach to the Problem of Rule-following

Daniel Watts [a]

[a] University of Essex , United Kingdom
Published online: 10 Feb 2012.

PLEASE SCROLL DOWN FOR ARTICLE

The Exemplification of Rules: An Appraisal of Pettit's Approach to the Problem of Rule-following

Daniel Watts

Abstract

This paper offers an appraisal of Phillip Pettit's approach to the problem how a merely finite set of examples can serve to represent a determinate rule, given that indefinitely many rules can be extrapolated from any such set. I argue that Pettit's so-called ethocentric theory of rule-following fails to deliver the solution to this problem he sets out to provide. More constructively, I consider what further provisions are needed in order to advance Pettit's general approach to the problem. I conclude that what is needed is an account that, whilst it affirms the view that agents' responses are constitutively involved in the exemplification of rules, does not allow such responses the pride of place they have in Pettit's theory.

Keywords: rules; rule-following; Pettit; exemplification; the ethocentric theory; Kripke

Philip Pettit has advanced a distinctive approach to a now familiar set of problems regarding rules and rule-following.[1] One distinguishing feature of Pettit's approach is the striking appeal it makes to a certain distinction between *instantiation* and *exemplification*. Pettit wants to show that a finite set of examples can *exemplify* a determinate rule, notwithstanding the fact that any finite set of examples *instantiates* indefinitely many rules. Further, Pettit seeks to ground this appeal to the instantiation/exemplification distinction in a novel theory of rule-following that gives explanatory primacy to agents' responses to finite sets of examples, that is, to agents' inclinations to regard such sets as exemplary for certain courses of action.

Pettit's account can be characterized, accordingly, in terms of two core claims:

Claim 1. A finite set of examples can exemplify a determinate rule, notwithstanding the fact that any finite set of examples instantiates indefinitely many rules; and

Claim 2. In order to exemplify a determinate rule, a finite set of examples depends on being regarded as exemplary for a certain course of action, where what it is for agents to regard examples in this way can be explained without reference to rules.

As we shall see in some detail below, Pettit's overall aim is to establish Claim 1 on the basis of Claim 2, and in this way to solve what he takes to be a radical problem.[2] The source of this problem – I'll call it *the representation problem* – is the following simple line of thought. Since indefinitely many rules can be extrapolated from any finite set of examples, and since rules have determinate applications in infinite or indefinitely large numbers of cases, it seems no finite set of examples could be capable of representing a rule. But since on the other hand we surely can and do follow rules on the basis of examples, it surely must be possible for a finite set to represent a rule.

Pettit sets out (inter alia) to solve this problem, and two features of his approach are worth highlighting at the outset. First, it is an important feature of his approach to the representation problem (as I am calling it) that he thinks it would be a serious mistake to conflate the question how examples can serve to represent a rule to an agent, with the less theoretically innocent question how examples can fulfil this function by *instantiating* a rule. To run these together, he thinks, would prejudice our thinking in favour of a questionable interpretation of what representation means in this context. Second, it is a central part of Pettit's aim to provide what he calls a 'non-sceptical' solution to the problem. In aiming for this, Pettit takes himself to be committed to a robust distinction between it merely seeming to an agent that he or she is following a rule on the basis of examples, and it really being the case that he or she is thereby 'put in touch' with a rule, conceived as a normative constraint that ranges over an infinite variety of cases.

My own aim in this paper is to provide an appraisal of Pettit's proposed solution to the representation problem. I shall argue that Pettit's general approach to this problem via Claim 1 is attractive and well-motivated, but that his theory of rule-following fails to deliver the kind of solution he wants, and that it fails in this way because it relies on Claim 2. I shall argue, in other words, that Pettit's theory fails to provide the appropriate grounds for his appeal to the exemplification/instantiation distinction; appropriate, that is, to the aim of providing in this way a non-sceptical solution to the representation problem. And I shall argue that at the source of this internal weakness in his approach is the basic explanatory role Pettit gives to agents, and their inclinations to respond to examples in certain ways, in his account of how examples represent rules. More constructively, however, I want also to consider

whether Pettit's general approach can be advanced in a way that does not rely on this problematic commitment to Claim 2; and I want tentatively to suggest that it can.

The structure of the paper is as follows. I shall begin by presenting Pettit's overall approach to the representation problem, drawing out three general requirements on the account of rule-exemplification this approach calls for (Sections I and II). I shall then argue that Pettit's own so-called ethocentric theory falls short when measured against this set of requirements (Section III). Finally, I shall consider, in a more programmatic way, what further provisions are called for if we are to advance Pettit's general approach to the problem (Section IV). What is needed, I shall argue, is a defence of the following claim:

> *Claim 3*. In order to exemplify a determinate rule, a finite set of examples depends on being regarded as exemplary for a certain course of action, where what it is for agents to regard examples in this way *cannot* be explained without reference to rules.

To defend this claim would be to defend what I shall call a 'no-priority' account of rule-exemplification. Like Pettit's theory, such an account would affirm that agents' responses to examples play a constitutive role in the exemplification of rules. Unlike Pettit's theory, however, such an account would not give explanatory priority to agents, and their inclinations to respond to examples in certain ways. Though it initially seems obscure and questionable, I shall argue that the idea of a no-priority account is at least intelligible, and that such an account appears better suited than Pettit's own theory to advance his general approach to the representation problem. My appeal to the idea of a no-priority account shall be conditional: I shall claim only that, if we take up and seek to advance Pettit's general approach to the representation problem, we shall find ourselves in need of such an account.

I

The representation problem is one of many threads in the entanglement of issues routinely picked out in the literature on rule-following, and I shall be brief on its general form. The problem is perhaps best presented as a more or less informal challenge to show how it can be possible for a finite set of examples to represent a rule in a determinate way, in the light of the observation that indefinitely many rules can be extrapolated from any such set. Nonetheless, if only for ease of exposition, we can formulate a rough argument for the conclusion that it is not possible for

a finite set of examples to determinately represent a certain rule, as follows:

(1) It is possible for x to represent a rule only if it is possible for x to represent what the rule requires in any case to which it may be applied;[3] but

(2) For any finite set of examples S and for any rule R it is not possible for S to represent what R requires in any case to which it may be applied.

Following Pettit, the line of argument can be sketched out as follows.[4] Suppose that it is among the defining features of rules that, first, rules range across infinite or indefinitely large numbers of cases; and that, second, there is a determinate answer to the question what a rule requires (or permits, or mandates, or proscribes) in any of the cases to which it is appropriately applied. And suppose, further, that if something is to be capable of representing a determinate rule – that is, a rule conceived as having a determinate application across an infinite or indefinitely large range of cases – it must be capable of representing what the rule requires (or permits, or mandates, or proscribes) in any case to which it is appropriately applied. It will be important for us to keep in mind that Pettit regards these basic suppositions about rules, and the conditions of their representation, as part and parcel of what it is to believe in 'the reality of rule-following': this much, he thinks, must be preserved in any plausible attempt to vindicate this belief in the face of a sceptical challenge.[5]

But now the sceptical challenge arises more or less directly. For it seems a limited range of examples could only ever serve at most to represent what a rule requires for a limited range of cases: for any given set, we can readily envisage rival ways of extrapolating such that the set can be made out to represent mutually incompatible answers to the question what the rule requires in further cases. On the supposition that it is of the nature of a rule to have a determinate application in an infinite or indefinitely large number of cases, it seems moreover that what is implied by the failure of a finite set of examples to represent how a rule is to be applied in further cases is that it fails to represent a determinate rule as such. But then it seems to follow that no determinate rule can be represented as such by a finite set of examples.

This conclusion is perplexing not least because we naturally assume that examples do have an important place in the teaching and learning of rules. But certain background theoretical considerations may also come into play here. In the first place, there is the well-established philosophical tradition of viewing our ability to grasp conceptual and semantic content as distinctively rule-governed. In the classic, Kantian version

of this view, a concept functions as a set of rules or instructions which specify how disparate elements of experience can always be so combined as to represent a single, determinate unity.[6] But there are many versions of the view that do not rely on the machinery of Kantian synthesis, or indeed the identification of concepts with rules. For the general idea is often roughly this: an agent counts as possessing a particular concept, or as grasping the meaning of a given term, only if he is able to follow rules which specify how the concept or term is correctly to be applied. This sort of claim is not uncontroversial. For the purposes of this paper, however, I shall follow Pettit in regarding the representation problem against this background; that is, as most fundamentally raising the question whether finite sets of examples can serve to represent conceptual and semantic rules (rather than in connection with any more narrowly defined rules such as mathematical functions or moral rules or social conventions).[7] Certainly, our sense of the scope and the severity of the problem will be magnified to the extent that we do indeed take the ability to follow rules to be a pervasive feature of the exercise of our conceptual and linguistic abilities. Of the sceptical challenge in general to the reality of rule-following, Pettit can thus write that this 'is of the greatest importance in the philosophy of mind, though many practitioners seem to think they can ignore it' (Pettit, 2002: p. 26).

One thing that serves to render the representation problem still more pressing is a familiar argument – and another thread in the entanglement of issues surrounding rule-following – which purports to show, not merely that examples occupy an important place in contexts of teaching and learning, but that examples are in general necessary to account for the very possibility of one's learning to follow and apply rules. Briefly, the argument is as follows.[8] On pain of a vicious regress, it cannot be a general requirement of my learning to follow a given rule that I must do so on the basis of my grasp of some further rule. If it is in general to be possible for me to learn to follow rules it must, therefore, be possible that I do so on the basis of something other than my prior grasp of rules; in other words, it must be possible to grasp a rule in a way that is not mediated by further rules and which is, to that extent, immediate and direct. But the outstanding candidate for what might make possible the immediate and direct grasp of a rule is, precisely, exposure to *examples* of its application. These appear to be the considerations, for example, behind Kant's well-known saying that examples are the go-cart of judgement (Kant, 2003 [1787]: B 172–B 174: pp. 177–8).

It will be a *desideratum* of any proposed resolution of the representation problem, then, that it takes into account the motivation for thinking of our conceptual and linguistic abilities as distinctively rule-governed and for thinking that the function of examples is crucial in explaining how these abilities can be learnt or acquired.

II

Pettit does not wish to deny that any finite set of examples instantiates indefinitely many rules. On the contrary, he affirms this as a fact:

> The fact is that any finite set of examples, mathematical or otherwise, can be extrapolated in an infinite number of ways; equivalently, any finite set of examples instantiates an infinite number of rules. It appears then that I cannot be in touch with a particular rule just on the basis of finite examples. (Pettit, 2002: p. 33)

It soon becomes clear however that Pettit does not mean to endorse the idea that it really is an implication of the fact that any finite set of examples instantiates indefinitely many rules that 'one cannot be in touch with a particular rule just on the basis of examples'. He thinks, rather, that the observation at the heart of the sceptical challenge can be accommodated within a robustly non-sceptical account of the place of examples in our learning to follow and apply rules. How so?

As we have already indicated, Pettit's first move towards establishing 'the reality of rule-following' is to advance the following claim:

> *Claim 1*. A finite set of examples can exemplify a determinate rule, notwithstanding the fact that any finite set of examples instantiates indefinitely many rules.

Pettit writes:

> . . . the relationship that is of concern to us when we ask whether a finite set of examples can represent a determinate rule is not instantiation but exemplification. Exemplification is a three-place relationship, not a two-place one: it involves not just a set of examples and a rule but also a person for whom the examples are supposed to exemplify the rule. Although any finite set of examples instantiates an indefinite number of rules, for a particular agent the set may [sc. may, possibly] exemplify just one rule. (Pettit, 2002: p. 36)

This passage indicates a distinctive diagnosis of what is amiss with the inference from the observation that finite sets can be extrapolated in indefinitely many ways to the conclusion that we cannot be 'put in touch' with rules on the basis of examples. What is amiss, on this diagnosis, is that this inference runs together the general question whether or not a set of examples can 'represent a determinate rule' (Pettit's

phrase) with the specific question whether or not such a set can represent a determinate rule merely in virtue of its *instantiating* the rule. This diagnosis is further developed by the thought that the specific question about instantiation is, in fact, radically different from the more general one, since if we were to formulate the latter question properly it would ask about exemplification rather than instantiation. And Pettit goes on to offer an instructive account of what the relevant difference is: namely, that whereas the instantiation of a rule is a relation between a set of examples and a rule, exemplification involves a crucial further element, namely those 'for whom' a set of examples is exemplary in a certain way.[9] This, in turn, indicates the following prognosis: our problem will remain unresolved all the while we fail properly to take into account the place of the person or agent, the place of the rule-*follower*, in our thinking about the relationship between rules and examples.

It should be clear, then, that the point of Pettit's appeal to this distinction between exemplification and instantiation is to rule out any interpretation of the initial notion of *the representation of a determinate rule* merely in terms of instantiation. However else one might want to distinguish between instantiation and exemplification, the critical contrast, for Pettit's purposes, is between a conception of what it is for a set of examples to represent a determinate rule which appeals to a relationship between just two entities – viz. a rule and a set of examples – and one that appeals not to these two *relata* alone but also to the ways in which agents respond to examples.

Two general features of the idea that we can respond to the representation problem by appeal to Claim 1 make this an attractive proposal. First, the proposal is, in a certain sense, modest: Pettit does not seek to refute the claim that indefinitely many rules are instantiated by any finite set of examples but rather to show that this observation does not have the sceptical implications it seems to have. Second, the proposal is, in another regard, ambitious and takes seriously the sceptical challenge associated with the representation problem. The task, as Pettit sees it, is to defend Claim 1 as a solution to this problem, not merely to invoke the distinction between instantiation and exemplification. As he conceives it, moreover, the task is in this way to deliver a so-called 'straight' or 'non-sceptical' resolution. Pettit writes:

> Any non-sceptical response to the challenge about rules has to vindicate the idea that we intentionally try to conform to entities that satisfy the objective condition: constraints that are normative over an infinite variety of cases. (Pettit, 2002: p. 35)

That is, a non-sceptical resolution, as Pettit understands it, must preserve a robust distinction between it merely seeming to me that I am following a rule on the basis of examples, and it being the case that I am thereby 'put in touch' with a rule, conceived as a normative constraint that ranges over an infinite variety of cases. Part of what is attractive about a non-sceptical response, so understood, is that it promises to take into account the motivation both for thinking that our conceptual and linguistic resources are rule-governed and for thinking that the function of examples is critical in explaining how it is possible for these abilities to be learnt or acquired. A non-sceptical resolution would take these considerations into account in a straightforward way, by vindicating the idea that agents really do follow constraints that are normative across an infinite range of cases on the basis of finite sets of examples.

What is called for by Pettit's proposed approach, then, is an account of what it is for a rule to be exemplified by a finite set of examples, appropriate to the aim to provide a non-sceptical solution to the representation problem based on Claim 1. And we can specify three general requirements on the appropriate account of rule-exemplification, as follows:

(R1)   *The account shall make ineliminable reference to agents' responses to finite sets of examples.*

(R2)   *The account shall specify some suitable relationship between agents' responses to finite sets of examples and the determinate rules represented by such sets.*

(R3)   *The account shall not rely on the supposition that the instantiation of a rule in a finite set of examples is sufficient to represent the rule.*

R1 derives from Pettit's claim that the representation of a rule by a set of examples is a relationship of exemplification, where the notion of exemplification is understood, by contrast with instantiation, as making ineliminable reference to agents' responses. R2, on the other hand, derives from the constraints on a defence of Claim 1 for the purpose of providing a *non-sceptical* solution; specifically, from the need to accommodate a robust distinction between it merely seeming to an agent that he or she is following a rule on the basis of examples, and it really being the case that he or she is thereby 'put in touch' with a rule, conceived as a normative constraint that ranges over an infinite variety of cases. For this implies that agents' responses to examples can enter into relationships with determinate rules; and we need some account of how

this can be so. Finally, R3 is motivated by the supposition that any finite set of examples instantiates indefinitely many rules, together with the supposition that to represent a determinate rule is to represent a constraint that is normative across an infinite variety of cases. We must accept R3, that is, if with Pettit we concede to the sceptical challenge that these two suppositions jointly render questionable whether the instantiation of a rule in a finite set of examples can be sufficient to represent the rule.[10]

In characterizing Pettit's overall approach in terms of these requirements, I do not claim that an account of rule-exemplification that meets R1–R3 would be sufficient to provide a non-sceptical solution to the representation problem based on Claim 1, only that it is necessary. And it remains to be seen what exactly it would mean for an account of rule-exemplification to meet these requirements. Let us turn then to the question whether Pettit's own theory is satisfactory when measured against R1–R3. I shall argue that it is not.

III

First, a brief overview of Pettit's so-called ethocentric theory.[11] The theory relies on two key postulates. In the first place, humans naturally form inclinations to extrapolate on the basis of similarities they regard as salient in the examples to which they are exposed. When properly spelled out, this postulate is supposed to explain how an agent can come to regard a set of examples as exemplary for a certain course of action. And second, humans naturally form second-order dispositions to seek convergence, across persons and across times, by discounting divergent responses if these can be put down to perturbing factors. This postulate, when properly spelled out, is supposed to explain how an agent's natural responses to examples can come to be suitably associated with a certain rule. For the rule will correspond to just that course of action which is made salient to the agent when his inclinations operate under (what the theorist may describe as) favourable conditions, i.e. those conditions, whatever they are, which survive the process of standardization.[12]

Let us not quarrel with Pettit's postulates as such: the question I want to press is whether they can furnish an account that is adequate to the requirements set up by his general approach to the representation problem. Whilst the theory obviously meets R1, I think there are reasons to doubt that Pettit's theory can satisfactorily meet R2 and R3.

The ethocentric theory obviously meets R1: for it tells a detailed story about the constitutive role played by agents' responses, inclinations dispositions and the like in the exemplification of rules. Indeed – and this point is crucial – Pettit's theory clearly affords a certain explanatory priority to the agent in this regard: that is, it gives priority to the notion

that which agents regard as exemplary for a certain course of action over the notion of being that which represents a determinate rule. It is this priority that allows Pettit to present his theory as offering a kind of 'genealogy' which explains how rule-following 'gets going', as he puts it, on the basis of agents' dispositions to respond to examples in certain ways.[13] There is, then, a reasonably clear sense in which Pettit's theory is *agent-based*.[14] That is, it relies on the following claim:

> *Claim 2.* In order to exemplify a determinate rule, a finite set of examples depends on being regarded as exemplary for a certain course of action, where what it is for agents to regard examples in this way can be explained without reference to rules.

But we should consider whether, in conforming to this agent-based explanatory framework, Pettit's theory is capable of meeting the requirement (R2) to specify some suitable relationship between agents' responses to finite sets of examples and the determinate rules represented by such sets. This requirement, we may note, evidently rules out any reductive account in which rule-exemplification is not in the final analysis to be understood as a genuine relation involving determinate rules but merely as a matter of regular patterns of agents' responses. And indeed, Pettit wants to show that, thanks to our second-order inclinations to seek interpersonal and diachronic standardization, our responses to examples can enter into relationships with rules *qua* normative constraints across an infinite variety of cases. As agent-based, however, Pettit's theory relies on the idea that agents' inclinations to regard a set of examples as exemplary for a certain course of action can, in principle, be properly specified without reference to rules; were this not so, such inclinations could not be treated as prior in the order of explanation.[15] So we should ask whether the inclinations postulated by Pettit's theory can be explained or properly specified without reference to rules.

We can give initial expression to the worry that this question must be answered in the negative by reference to Pettit's second postulate in particular, viz. the second-order inclination to engage in practices of interpersonal and diachronic standardization. This postulate, we might naturally suppose, involves the idea of agents who are in a position to assess convergence on, and divergence from, particular ways of extrapolating from examples. If we consider what being in this position would involve, however, it is difficult to see how this could be properly specified, even in principle, without reference to rules. For it seems that, in assessing whether or not agents converge on one and the same course of action, Pettit's agents would be assessing whether or not each others' responses accord with one and the same rule: were this not so, the

agents would appear to be lacking any way of determining what counts as one and the same extrapolation, one and the same course of action. It seems these second-order dispositions would, in other words, already presuppose the representation of determinate rules by finite sets of responses. But that would plainly be incompatible with the supposition that these dispositions can be specified without reference to rules and so treated as basic in the order of explanation. Moreover, we may observe quite generally that any theory that is agent-based, in accordance with Claim 2, looks likely to be hard-pressed to specify a suitable relationship between the responses involved in the exemplification of a rule and the rule itself. It looks this way because it is difficult to see what could possibly count as two or more agents extrapolating from a given set of examples in the same, determinate way, i.e. their regarding the set as exemplary for the same, determinate course of action, were this not a matter of their being governed by the same, determinate rule.

Two replies to this line of objection might be offered on Pettit's behalf.[16] In the first place, one may observe that Pettit's account of the way rule-followers act on their second-order inclinations to seek interpersonal and diachronic standardization is precisely designed to show how disagreements about what a rule requires can be resolved without agents already knowing what the rule requires. In Pettit's account, that is, rule-followers act on their inclination to seek standardization by taking up a process of negotiation in which divergent responses are discounted if these can be put down to perturbing factors. But – and this is the first reply – rule-followers can surely take into account perturbing factors (such as tiredness, inebriation, contextual factors and the like) without already knowing what the rule requires in a given case. Disagreements about what a rule requires can thus be settled, not on the basis of an appeal to the rule itself, but rather by reference to factors that pertain to (what the theorist may describe as) favourable conditions.

For the sake of argument, let us grant that, once a disagreement has been identified in a given case about what a rule requires, rule-followers could resolve the disagreement in the way Pettit envisages, by appeal to perturbing factors, and this *sans* appeal to the rule in question. The issue raised by the objection we are considering, however, is what must be involved in the ability to regard something *as* a convergence or divergence of the appropriate kind in the first place, and so in the ability to identify something as a disagreement about what a rule requires. To illustrate this issue, we can recall Wittgenstein's famous awkward pupil, whose responses to the instruction to add 2 includes his writing '1004' after '1000'. As Wittgenstein observed, we might specify this pupil's divergence from our way of going on by saying that the pupil understands the instruction to add 2 in the way that *we* should understand the instruction, 'add 2 up to 1000, 4 up to 2000 … and so on' (see

Wittgenstein, 1967 [1953]: §185). Now, it appears that, in order for us to be able to regard the awkward pupil in this way – i.e. not merely as vaguely aberrant, as making noises and marks we don't make, but as following a different rule or as following the rule on a different interpretation – we must already be in a position to regard a finite set of examples as that which exemplifies a determinate (interpretation of a) rule, albeit a non-standard one; in this case, a set of examples of the pupil's responses to the instruction to add 2. If this is right, however, it follows that Pettit's second-order inclinations cannot be primary in the explanation how in general examples can represent rules since the latter is presupposed by an account of the former.

To be clear, then, the objection we are considering does not allege that, in order to settle a disagreement about what a rule requires, Pettit's agents must already know what the rule requires. Rather, the objection alleges that, in order for us to conceive of these agents as assessing convergence and divergence in the following of rules, we must conceive of them as already able to follow rules on the basis of examples, and that for this reason we cannot coherently give explanatory priority to such assessments in our account of how examples can serve to represent rules. A second reply, however, is this. Perhaps Pettit should say that, in acting on their second-order inclinations, we need not suppose that rule-followers must be able to represent to themselves convergence and divergence in the following of rules; it is enough that they are in fact sensitive to such convergence and divergence. In other words, on this line of response, it is sufficient for Pettit's purposes that courses of action can more or less immediately strike rule-followers as converging or diverging and that members of the rule-following community have relatively similar patterns of response in this respect.[17]

Let us grant that when Pettit's agents act on their second-order inclinations they need not represent to themselves convergence and divergence in the following of rules. Again, however, we may observe that whatever exactly sensitivity to convergence and divergence is supposed to amount to here, it must amount to more than merely tracking a difference between those who make the noises and movements we make and those whose noises and movements strike us as alien; for this would surely fall short of anything properly called sensitivity to convergence and divergence in the following of rules or with respect to certain courses of action.[18] Moreover, as we have observed, it would be in violation of R2 to suppose that the exemplification of rules is, at bottom, merely a matter of regular patterns of response. The difficulty remains therefore how Pettit's agents are able to track convergence and divergence of the relevant kind if this is not on the basis of the representation of rules by examples, viz. examples of agents' responses. For again, if we suppose that this is indeed the basis on which Pettit's agents are able to

track convergence and divergence, we must infer that finite sets of examples are capable of representing rules to agents prior to and independently of the outcome of any process of negotiation and standardization among them.

Turning to R3 – which, to recall, rules out any account that relies on the supposition that the instantiation of a rule in a finite set of examples is sufficient to represent the rule – I think we also have reason to doubt that Pettit's theory can satisfactorily meet this requirement. In order to articulate the difficulty here, it will be useful to further specify what is ruled out by R3 by envisaging what we can call an *instantiation-plus* model of rule-exemplification. On this model, the exemplification of a rule is to be understood as a two-stage process: in the first stage, the agent identifies some among the indefinitely many rules instantiated by the examples; in the second stage, the agent singles out just one of these rules as the unique rule he or she intends to follow. So, to invoke Kripke's well-worn example, if a set of sums instantiates both PLUS and QUUS then, on the instantiation-plus model, the set exemplifies PLUS, and not QUUS, just in case the agent has identified the PLUS rule as one among the myriad of rules instantiated in the set and the agent intends to follow the PLUS rule. We can notice that, whatever its shortcomings, the instantiation-plus model does at least suggest a ready story about what it means for an agent to *follow* a rule on the basis of examples, roughly this: rule-followers first identify rules that are represented to them by examples in a way that is entirely independent of their responses, that is, just by dint of being instantiated therein; they can then try to bring their actions into conformity with one of the rules so identified. Plainly, however, this conception of what it is to follow a rule on the basis of examples is available only to a theory that contravenes R3.

Now, it seems clear that Pettit's theory is not supposed to be of the instantiation-plus variety; on the contrary, as we have emphasized, he maintains that the very notion of the representation of a rule by examples must be understood in unitary way, in terms of a three-place relationship of exemplification rather than in terms of a two-place relation of instantiation. Nonetheless, many passages appear to express Pettit's commitment to the idea that following a rule on the basis of examples constitutively involves trying to bring one's responses into conformity with antecedently identified rules. According to his prominent general gloss of what it is to follow a rule, for example, '[t]o follow a rule is to conform to it intentionally: to conform as a result of trying to conform' (Pettit, 2002: p. 3). More:

... [A rule] should be determinable or identifiable by a finite sub-
ject independently of any particular application: the rule-follower
should be in a position to identify the rule in such a manner that
he can sensibly try to be faithful to it in application. If the rule
were identified by reference in part to how the subject responded
in a given case, then the subject could not see the rule as some-
thing to which he should try to be faithful in that case. He could
not see it as a normative constraint for him to try to respect there.
(Pettit, 1990a: p. 3)[19]

The question arises however whether Pettit can make good on these
claims without relying on the supposition that rules are represented by
examples in a way that is independent of any particular response on the
part of the agent. To be sure, Pettit is careful to emphasize that it is not
part of his view that, in order to follow a rule on the basis of a set of
examples, one must reflect on how one is inclined to regard the exam-
ples. On the contrary, at the psychological level, such inclinations are
said merely to play a 'certain background, causal role' in making salient
a certain way of going on (Pettit, 1990b: p. 438). And Pettit evidently
thinks that to say that rule-followers must be able to recognize what a
rule requires 'independently of any particular application' is not to say
that rule-followers must be able to recognize what a rule requires with-
out having to rely on their inclinations to extrapolate in a certain way
when faced with a finite set of examples.[20] So, despite the impression
that might be given by the passages just cited, it may be a mistake to
ascribe to Pettit the view that following a rule on the basis of examples
constitutively involves a prior mental act of identifying the rule to which
one subsequently tries to conform. Still, we may express our worry here
in the form of a dilemma: if Pettit is committed to the view that follow-
ing a rule on the basis of examples requires such prior acts of identifica-
tion, then it is hard to see how he can make good this commitment
without violating R3; but if not, then it is hard to see how his theory can
cash out the idea that 'to follow a rule is ... to conform by trying to con-
form'. The issue here, then, is an apparent tension between the
implications of R3 and the general conception of what it is to follow a
rule that Pettit's ethocentric theory is expressly designed to underwrite.[21]

In the face of these difficulties regarding R2 and R3, one might be
tempted to conclude that the prospects for Pettit's approach to the rep-
resentation problem are not bright. Certainly, I think his proposed solu-
tion is unacceptable as it stands. As I also hope to have made clear,
however, Pettit's general approach to the problem, via Claim 1, can be
considered apart from the particular way in which he develops this
approach, via Claim 2. We should therefore ask whether his appeal to

the exemplification/instantiation distinction can be defended, and his general approach to the representation problem advanced, on the basis of some alternative account. In my view, the general approach remains a promising one, provided Pettit's agent-based theory can be replaced with a different sort of account to which I now turn.

IV

Consider then the following alternative to Claim 2:

> *Claim 3*. In order to exemplify a determinate rule, a finite set of examples depends on being regarded as exemplary for a certain course of action, where what it is for agents to regard examples in this way *cannot* be explained without reference to rules.

What is called for by Pettit's general approach to the representation problem, I submit, is a defence of Claim 3. We may envisage such a defence as advancing a 'no-priority' account since such an account would treat as primary neither the notion of that which determinately represents a determinate rule nor, on the other hand, the notion of that which agents regard as exemplary for a certain course of action.

In what remains of this paper, I shall confine our discussion of Claim 3 to two basic questions: First, is this claim, and the idea of a 'no-priority' account, even intelligible? And second, how appropriate is such an account to advance Pettit's general approach to the representation problem? My aim is to indicate how a no-priority account can provide us with a way to recover Pettit's important insight that rule-exemplification is not rule-instantiation, and so to advance his proposed solution to the representation problem, but without conceding to agents and their inclinations the pride of place they have in the ethocentric theory.

The question about intelligibility arises immediately because of the impression of vicious circularity in Claim 3. That is, it seems an account of the sort envisaged would, in circular fashion, try to explain the exemplification of rules by appeal to agents' inclinations to regard finite sets of examples as exemplary in certain ways, but then to try to explain this latter phenomenon, in turn, by appeal to the exemplification of rules by finite sets of examples. But it is of course a natural and orthodox view that explanation is an asymmetrical relation. We may think here, for example, of Socrates' rejection of Euthyphro's proposed definition of piety, on the grounds (according to one reading) that the most that could be established by demonstrating a reciprocal relationship between *being pious* and *being regarded as pious by the gods* is that these terms are necessarily co-extensive, not which explains which (cf. *Eutrhyphro* 11a–c: Cooper, 1997: p. 11).

Now, in considering the idea of a no-priority account, it is first of all important to recognize that such an account must be conceived as different in kind to Pettit's agent-based theory. For we will surely be unable to make any sense of this idea if we suppose that the aim of such an account is to furnish an explanation of the sort Pettit's theory purports to provide, viz. an explanation of how rule-following 'gets going' on the basis of our natural responses to examples and our propensities for standardization. To be clear, then, to advance a no-priority account would be to give up that part of Pettit's project that involves the idea of a genealogy or aetiology of rule-following. Perhaps this will be regarded as a high price to pay. But we may note that the idea of a genealogy of rule-following was in any case no part of our original account of the general requirements on the appropriate defence of Claim 1 for the purpose of a providing a non-sceptical resolution to the representation problem.

How, then, are we to understand the idea of a no-priority account, if not as a genealogical explanation? We can make good sense of this idea, I want to suggest, if we can make sense of the claim that *being that which exemplifies a determinate rule* stands in a relationship of reciprocal dependence with *being regarded as exemplary for a certain course of action*. But of course our misgivings regarding the intelligibility of this formulation may remain. For essential dependence and explanation are often regarded as closely related if not identical notions, and both as distinctively asymmetrical.[22]

In response to these misgivings – and I consider them here only insofar as they bring into question the very intelligibility of the idea of essential co-dependence – we may begin by observing that claims to such interdependence are sometimes made in other contexts. Whether or not we agree with them, we seem to be able to make sense of such claims: consider, for instance, the claim that *being of a certain hue* is co-dependent with *being of a certain saturation*.[23] But there is a more direct response to the worry. For it is not in fact clear that relationships of essential dependence must be conceived as asymmetrical or acyclic. Kit Fine, for example, has forcibly argued to the contrary that the notion of essential dependence can be cogently introduced in such a way as to admit of reciprocity (see Fine, 1994: pp. 282–4). This relation will be cycle-tolerant, Fine observes, under roughly the following definition: x essentially depends on y 'if either x is simultaneously defined with the help of y or if there is a definition of x in terms of y' (1994: p. 283). All that is needed in order to make good sense of the idea of two-way essential dependence, Fine goes on to argue, is the idea that two or more entities may enter into an irreducible essential relationship in virtue of being simultaneously defined. Certainly, there is no obvious reason to dismiss this notion of simultaneous essential definition as unintelligible; on the contrary, there is a natural analogy here with the familiar idea of

simultaneous *nominal* definition, i.e. the notion of mutually defined terms. And it may be that the idea of a no-priority account of rule-exemplification is best understood in these terms, that is, as offering a simultaneous definition of *being that which exemplifies a determinate rule* and *being regarded as exemplary for a certain course of action*.

This proposal allows us to conceive of a no-priority account in terms of two concomitant claims. It would claim, firstly, that for a given set of examples to exemplify for an agent a determinate rule there must be a certain way of extrapolating from the examples the agent finds natural. What of cases in which the same examples strike members of a rule-following community in divergent ways? A no-priority account of rule-exemplification is as such silent on this issue, though we may note that the idea of such an account is not obviously incompatible with the idea, central to Pettit's ethocentric theory, that it is natural for humans to seek to resolve such disagreements through practices of negotiation and standardization.[24] Secondly, however, a no-priority account would further claim that, conversely, for it to be, precisely, a certain way of extrapolating from a given set of examples that an agent finds natural, the examples must serve to exemplify for the agent a determinate rule. The key phrase here is, 'a certain way': the claim is that an agent is properly said to extrapolate from examples *in a certain way* only on the supposition that the examples function to exemplify for the agent a determinate rule. This claim can be understood as motivated by the thought, which has played an important role in our criticism of Pettit, that agents are properly said to converge on or diverge from a certain course of action only if their responses count as rule-governed.

Understood in this way, how appropriate is such an account to advance Pettit's general approach to the representation problem? As in Pettit's theory, and in accordance with R1, it should be clear that, so construed, a no-priority account makes ineliminable reference to ways in which agents respond to finite sets of examples. In common cause with the ethocentric theory, such an account would have to defend the view that, in order to exemplify a determinate rule, a finite set of examples depends on being regarded by agents as exemplary for a certain course of action. The account would therefore support Pettit's general gloss on the notion of rule-exemplification as a three-place relation, constituted not only by the rule itself and by a set of examples but also by the responses of the agent. And we may add that there is no obvious reason why the account could not make use of Pettit's postulates, suitably reconceived, in order to articulate and defend the relevant contrast between exemplification and instantiation.

But the crucial point of comparison with Pettit's agent-based theory is the contrasting conception, in a no-priority view, of what it is for an agent to regard a set of examples as exemplary for a certain course of

action, i.e. the claim that one cannot properly specify a given response to examples as falling under this description without reference to rules. Whilst a no-priority account would affirm the constitutive role of agents' responses in the exemplification of rules, such an account would therefore rescind from any problematic attempt to specify the former without reference to the latter. More positively, the account would be well-placed to satisfy R2 by suitably describing the relationship, in cases of rule-exemplification, between agents' responses to finite sets of examples and the determinate rules represented by such sets: that is, in terms of the relation of *being governed by*. For a no-priority account would have to maintain that agents' inclinations to regard finite sets of examples as exemplary for certain courses of action are already rule-governed – and that this holds, so to speak, all the way down.

Would a no-priority account be appropriate to meet R2 in a way that also satisfies R3? Plausibly, this depends on whether we can make sense of the idea of a way of responding to examples that is, on the one hand, rule-governed, but which on the other hand does not involve prior acts of reflection on representations of rules. For whereas R3 rules out any appeal to the instantiation of a rule as constituting the appropriate material for such acts of reflection, it follows from a no-priority account that the relevant responses must nonetheless be conceived as rule-governed. In other words, if it is to advance Pettit's general approach to the representation problem, a no-priority account must make it clear how agents' *spontaneous* (i.e. non-reflective, non-discursive) responses to examples can nonetheless count as rule-governed.

No doubt, this is a difficult challenge. After all, Pettit is among many theorists to have supposed that the idea of one's being governed by a rule, or of one's following a rule – typically in contradistinction with the notion of one's merely *conforming* to a rule – must be understood precisely in terms of one's being in a position to identify a rule in advance of any particular application. And it ought in general to be clear that properly spelling out and defending a no-priority account is a large and further task. But we may at any rate recall in this connection Kant's saying about examples being the go-cart of judgement: for, as we have noted, according to the thought expressed in this saying it must indeed be possible to grasp a rule in and through one's spontaneous responses to examples, i.e. on the basis of responses that, although rule-governed, are not mediated by prior acts of reflection on representations of rules.[25]

These concluding remarks are admittedly tentative and programmatic, and the main upshot of this paper is that further work is needed if Pettit's proposed approach to the representation problem is to be advanced in a convincing way. To the extent that we find the proposal basically attractive and well-motivated, however – and I hope to have lent some support for finding it so – we shall want to take up the task of

developing a suitable account of rule-exemplification. To that extent, and inasmuch as it promises to meet the requirements set up by Pettit's approach, spelling out and defending what I have called a no-priority account seems a task worth taking up.[26]

University of Essex, United Kingdom

Notes

1 This approach is summarily laid out in Pettit, 1990a; reprinted in Pettit, 2002: pp. 26–48.
2 In this paper, I shall consider Pettit's approach to the representation problem in abstraction from his wider response to 'the sceptical challenge about rules' which he characterizes more generally in terms of a basic tension between certain 'objective' and 'subjective' conditions of rule-following (see Pettit, 2002: p. 32). This means that I shall be restricting attention to what is, strictly speaking, only part of Pettit's account, viz. that part which engages directly with (what I am calling) the representation problem. This restriction of focus will appear especially judicious if we share Paul Boghossian's view that '[i]t's hard to see … exactly what problem about rule following Pettit has identified' in his general idea of a tension between objective and subjective conditions (Boghossian, 2005: p. 191).
3 Especially because of the 'any' in (1), it might seem this claim is implausibly strong in an obvious way, since it might seem to rule out any possibility of there arising genuine doubts, for a given set of examples, regarding which rule is represented by the set. But this is a mistake, since the claim makes reference only to those cases in which it is given that a rule *may* (i.e. may appropriately) be applied.
4 This line of argument is of course usually associated with Wittgenstein and, especially, Kripke (see, e.g. Wittgenstein, 1958: p. 15; Kripke, 1982: pp. 8–9, 43; cf. Pettit, 2002: pp. 33–4). But see also Fogelin, 1976: Ch. X for an earlier reading of Wittgenstein along the lines Kripke takes.
5 For Pettit's 'stipulative account' of rules in general see Pettit, 2002: pp. 27–31. Pettit offers no substantive defence of his claim that it is of the nature of rules to range determinately across an infinite variety of cases. But the idea might reasonably be thought to get support from other quarters; in the case of grammatical rules, for example, from Chomsky's Poverty of the Stimulus argument (for a forceful defence see Laurence and Margolis, 2001).
6 'It is only when we have thus produced synthetic unity in the manifold of intuition that we are in a position to say we know the object. But this unity is impossible if the intuition cannot be generated without a rule. Thus we think a triangle as an object, in that we are conscious of three straight lines according to a rule by which such an intuition can always be represented … a concept is always, as regards its form, something universal which serves as a rule' (Kant [1787] A 105–6, p. 135).
7 Nothing in what follows, however, turns on the plausibility of Pettit's general theory of concepts (see Pettit, 2002: pp. 59–60). The restriction of focus adopted here (see note 2) has the advantage of allowing us to assess Pettit's account independently of these, no doubt controversial, commitments.
8 For a forceful analysis and defence of this line of argument see Bell, 1987a and 1987b. Bell shows that this argument is a serious problem for Frege's

theory of judgement and that it is crucial both for Kant's doctrine of the schematism and for many of Wittgenstein's remarks on rule-following.

9 It is moot how exactly we are to understand the 'for' in this 'for whom'. I shall argue (in Section III below) that Pettit's developed theory of rule-following fails to furnish a satisfactory account of the way in which agents' responses enter into exemplification relations.

10 One might object to R3 by observing that, quite generally, exemplification implies instantiation and inferring that for this reason R3 cannot possibly be met. Thus, in an influential treatment, Nelson Goodman characterizes exemplification as 'possession plus reference' (1976: p. 53); and Catherine Elgin writes that '[a]n item that at once refers to and instantiates a feature may be said to *exemplify* that feature' and that 'exemplification requires instantiation' (1999: pp. 171–2). But it would be fallacious to infer on these grounds that R3 cannot be met. It is of course true that exemplification implies instantiation, if by 'x instantiates F' we mean simply 'x is F'; for it is of course true that if x exemplifies F then x is F. But it does not follow that R3 cannot be met, since it does not follow from the claim that x is F that x's being F is sufficient to explain how x can represent F. What is ruled out by R3 is not the claim that examples instantiate rules but rather a particular account of what it is for examples to represent rules, namely one which appeals to just these two *relata*, examples and rule, and the relationship between them.

11 'Ethocentric' because the theory 'gives centre stage to habits of response and practices of self-correction, and both notions are captured in the Greek word *ethos*.' (Pettit, 2002: p. 66). The ethocentric theory is spelled out in Pettit, 2002 at pp. 26–48.

12 Pettit most fully develops his distinctive, functionalist account of 'favourable conditions' in 'A Theory of Normal and Ideal Conditions' (Pettit, 2002: pp. 136–56). For critical discussion which focuses on this account see Hindriks, 2004 and Haukioja, 2006.

13 '[The ethocentric theory] tells a story about how rule-following might get going; it offers a genealogy of rule-following on a par with Hume's genealogy of causal talk or, more notoriously, Nietzsche's genealogy of morals.' (Pettit, 2002: p. 42; and see pp. 8–9, 66).

14 In Pettit's own preferred terms, his theory is a 'response-dependence' account. In a number of papers, Pettit has been concerned to clarify what this sort of account of concept possession amounts to in general (see Pettit, 2002: Part I). He has been especially concerned to show that whilst the account does commit us to a certain 'relativity to our species, perhaps even our culture' (p. 44), such that our responses to examples grant us privileged access to the entities to which our concepts apply, it does not comprise a commitment to an appropriate degree of realism about these entities. I should therefore emphasize that what is at issue, in the present context, is the priority given by Pettit's account to agents' responses *vis-à-vis* exemplified rules, not *vis-à-vis* the entities to which these rules may give us access.

15 Pettit expressly acknowledges this constraint when he notes that he takes it for granted 'that routines of extrapolation and revision need not involve the following of rules' (2002: p. 8).

16 I am grateful to two anonymous referees of this journal for pressing me to develop my line of objection in relation to these possible replies.

17 I take it that Pettit means to forestall the sort of objection I have raised when he writes: 'I take it for granted that a species like ours might have initiated rule-following on the basis of extrapolative and revisionary routines

that did not themselves presuppose rule-following, even if at a more advanced level those routines generally do' (2002: p. 8). This response will seem inadequate, however, to the extent that anything that could satisfy the description of a 'revisionary routine' involves assessing convergence on ways of extrapolating from examples. For the worry is that this sort of assessment *must* involve rule-following.

18   As James Conant remarks, *apropos* Frege's arguments against psychologism, creatures who moo and eat grass do not thereby exhibit an alien form of thought; see Conant, 1991: p. 147.

19   See also Pettit, 2002: p. 29; 2005: pp. 233–6.

20   Thanks to an anonymous referee of this journal for drawing my attention to this point.

21   Paul Boghossian has forcibly argued on general grounds that we should reject Pettit's view that following a rule constitutively involves intending to conform one's behaviour to a rule; a view, as Boghossian notes, that he himself once shared. See Boghossian, 2005: p. 194ff.

22   Thus Jonathan Lowe observes (albeit pending qualification) that 'the intuitive notion of ontological dependence that is in play here is the notion of a distinctively *asymmetrical* relation … because it expresses an explanatory relationship and explanation is asymmetrical' (2005).

23   The example is Husserl's (see e.g. 1997 [1907]: p. 143).

24   It is worth emphasizing that the idea of a no-priority account, as it is envisaged here, involves separating out two issues that are not clearly separated in the ethocentric theory: on the one hand, the issue how a set of examples can serve to represent a determinate rule (the representation problem); on the other hand, the issue how there can be correctness conditions for the application of rules (sometimes called the normativity problem). The no-priority account is to be thought of as specifically addressed to the representation problem.

25   The idea that non-discursive responses are intelligibly conceived in terms of *following* a rule (rather than merely conforming or adhering to a rule) is, I believe, a central though often neglected theme in Wittgenstein. Compare: 'One follows a rule *mechanically*. Hence one compares it with a mechanism. "Mechanical" – that means: without thinking. But *entirely* without thinking? Without *reflecting*' (1978 [1956] VII-60: p. 422); and more famously: 'there is a way of grasping a rule which is *not* an *interpretation*' (1967 [1953]: §201).

26   I would like most especially to thank Vasilis Politis and David McNeill for their generous help with this paper. Thanks also to Phillip Pettit for his magnanimous responses to an earlier version; and to John Callanan, Béatrice Han-Pile, Andrew Howatt, Andrew Jorgensen, James Levine, Blain Neufeld and Komarine Romdenh-Romluc for their astute and extensive critical comments. A postdoctoral fellowship from the Irish Research Council of the Humanities and Social Sciences enabled me to pursue this work.

References

Bell, D. (1987a) 'Thoughts', *Notre Dame Journal of Formal Logic* 28(1): 36–50.

Bell, D. (1987b) 'The Art of Judgement', *Mind* 96: 214–44.

Boghossian, P. (2005) 'Rules, Meaning and Intention – Discussion', *Philosophical Studies* 124: 185–97.

Conant, J. (1991) 'The Search for Logically Alien Thought: Descartes, Kant, Frege, and the *Tractatus*', *Philosophical Topics* 20(1): 115–80.

Cooper, J. M. (ed.) (1997) *Plato: Complete Works*, Indianapolis: Hackett Publishing Company.

Elgin, C. Z. (1999) *Considered Judgement*, Princeton: Princeton University Press.

Fine, K. (1994) 'Ontological Dependence', *Proceedings of the Aristotelian Society* 95: 269–90.

Fogelin, R. J. (1976) *Wittgenstein*, London: Routledge.

Goodman, N. (1976) *Languages of Art*, Indianapolis: Hackett Publishing Company.

Hindriks, F. (2004) 'A Modest Solution to the Problem of Rule-Following', *Philosophical Studies* 121: 65–98.

Haukioja, J. (2006) 'Hindriks on Rule-Following', *Philosophical Studies* 126: 219–39.

Husserl, E. (1997) [1907] *Thing and Space: Lectures of 1907*, trans. R. Rojcewicz, Dordrecht: Kluwer Academic Publishers.

Kant, I. (2003) [1787] *Critique of Pure Reason*, trans. N. Kemp Smith, Hampshire: Palgrave Macmillan.

Kripke, S. (1982) *Wittgenstein on Rules and Private Language*, Oxford: Blackwell.

Laurence, S. and E. Margolis (2001) 'The Poverty of the Stimulus Argument', *The British Society for the Philosophy of Science* 52(2): 217–76.

Lowe, J. (2005) 'Ontological Dependence' *The Stanford Encyclopaedia of Philosophy* (Spring 2010 Edition), ed. Edward N. Zalta, available from: http://plato.stanford.edu/archives/spr2010/entries/dependence-ontological/ [accessed 14 November 2011].

Pettit, P. (1990a) 'The Reality of Rule-following', *Mind* 99: 1–21.

Pettit, P. (1990b) 'Affirming the Reality of Rule-following', *Mind* 99: 433–9.

Pettit, P. (2002) *Rules, Reasons and Norms*, Oxford: Oxford University Press.

Pettit, P. (2005) 'On Rule-Following, Folk Psychology, and the Economy of Esteem: A Reply to Bohossian, Dreier and Smith', *Philosophical Studies* 24: 233–59.

Wittgenstein, L. (1967) [1953] *Philosophical Investigations*, trans. G. E. M. Anscombe, Oxford: Blackwell.

Wittgenstein, L. (1978) [1956] *Remarks on the Foundations of Mathematics*, eds. G. H. von Wright, R. Rhees and G. E. M. Anscombe, trans. G. E. M. Anscombe, Oxford: Blackwell.