On Exploring Normative Constraints in New Situations

Jan Bransen, Utrecht University

In this paper I shall defend an adaptation of Philip Pettit's ethocentric account of rule-following[1] that emphasises the crucial role the account implicitly gives to intrinsically ambiguous assumptions about normality conditions. I shall first give a short exposition of Pettit's account as basically a story about our capacity to reason (section 1), and then show how and why assumptions about what is and what is not normal are rightly built into it (section 2). In section 3 I shall discuss how agents could reason their way out of the practical problems they typically will be confronted with in radically new situations. I shall suggest that an important ingredient of their capacity to determine the normative constraints that apply to the situations they find themselves in, is their capacity to articulate, specify and evaluate their implicit assumptions about what is and what is not normal. In the final section I shall discuss one consequence of my emphasis on the role of normality assumptions in the ethocentric account: the predominant role of *intrapersonal* divergence and the related, merely criteriological role of convergence with respect to determining the right rules to follow.

My basic interest in the topic derives from such a story as the following, one that is suggestive but according to me utterly implausible:

In the New World there were no rules. Everything was possible, and everything permitted. The New World changed only gradually into a world that displayed a normative structure, and this was so merely as a consequence of the fact that the immigrants from Europe succeeded in making agreements, in establishing conventions. Thus: whenever there are rules or normative constraints, these are constituted by or realised in agreements. Arguing that a radically new situation has of itself a normative structure waiting 'out there' to be detected, makes sense only as a reversal: it is not the situation that is new, but the newcomer who fails to share, at least initially, in the common knowledge about prevailing conventions.

I basically see this paper as providing at least the beginning of an argument that aims to show that this picture is false, because (1) novelty is a relational property connecting situations and agents, and (2) the ways in which situations and individual agents are connected is always normative in a way that exceeds mere habitual familiarity.

1. Pettit's ethocentric account of rule-following

Philip Pettit developed his ethocentric account of rule-following as part of a more comprehensive picture of the requirements intentional agents should meet in order to count as thinking subjects, i.e. as agents that have the competence to satisfy their own desire to be rational. Such agents, Pettit argues, (1) should be able to entertain beliefs and desires about the *content* of their beliefs and desires (this is the requirement of intentional ascent), and (2) should be able to treat some of these contents as normative constraints. This means that they should be able to change, adjust and/or revise their beliefs and desires, as well as the intentions and actions to which these beliefs and desires give rise, in the light of the import of these contents. This second requirement is that of rule-following, and it is met, according to Pettit's ethocentric account, by intentional agents whose *habits* of response and *practices* of negotiation are such that their actions could rightly be described as displaying a pattern of being sensitive to and guided by the normative import of the contents of their mental states. These contents function as normative constraints or rules, i.e. as entities that, as Pettit says, are normative over an indefinite variety of cases, which means that for each of these indefinite number of cases the entity or rule (i.e. the content) identifies which one of many possible beliefs and/or desires and/or intentions and/or actions the agent should entertain, form and/or perform.

There are many difficulties with spelling out the full details of such a general story about such a general capacity. From a strategic point of view it is therefore admirable that Pettit has succeeded in moulding Kripke's sceptical challenge[2] as an extremely useful stepping-stone for the articulation of his account. Pettit aims to offer a non-sceptical solution, and this means that he starts from the assumption that the situations intentional agents could find themselves in can be presented to or experienced by these agents in terms of normative constraints that inform these agents about what to believe, desire, intend or do in these situations.[3] This is a realist assumption of, in my view, a very unspecific and uncontroversial kind. Although it gives us the beginning of what could arguably be conceived of as a transcendental argument, it is, perhaps for that very reason, in itself a quite innocent assumption. This is so because it is compatible with the Kripkean/Wittgensteinian observation that every situation could be so presented in countless different ways, each of which would entail different normative constraints that could all sensibly

be said to apply to the situation in question. Yet I think the assumption is interesting because it gives us a picture of reality as normatively rich. That is, I take the assumption to entail the conceivability of realism, of the thesis that reality, the world out there as it is in and of itself, should play a role in determining the appropriateness of our beliefs, desires, intentions and actions.

With this assumption in place Pettit proceeds by asking whether finite creatures such as we could identify and understand[4] the normative import of the content in terms of which a situation would typically be presented to (or experienced by, as I would say) the kind of creatures we are. Pettit's claim that we can is based on his view that the relevant relation between rules and situations to which they apply is not a two-term relation of instantiation, but a three-term relation of exemplification. Situations *exemplify* a particular rule or normative constraint *for particular subjects*. The idea is that a situation is presented to or experienced by a particular kind of subject not just in terms of any conceivable contents, but in terms of a quite particular and specific content that exemplifies for this type of subject not just any normative constraint that could be instantiated by the features of the situation in question, but a quite specific rule that provides this subject with a content that is determinate in terms of its normative import.

Pettit argues that this is due to the limited capacity of agents such as we to develop dispositions, or extrapolative inclinations, in response to each *kind* of situation. Thus, if presented with an English letterbox, a ripe tomato, a fire engine and a heated metal, we human beings just develop the single extrapolative inclination to go on in terms of selecting further items by their red colour. Although the situation in question could be presented to the subjects in question (human beings, Pettit assumes) by means of many different contents, our nature interferes, so to speak, and this results in the fact that the situation in question is experienced by these particular subjects in terms of a content (redness) that exemplifies merely one determinate rule.

Pettit introduces the relevance of *human nature* to the determination of the kind of rules that exist for us by arguing that these rules (which intentional agents of our kind *should* follow) have a special and very intimate relation to our dispositions. It is in terms of the disposition that I develop spontaneously when confronted with certain exemplars, that I should identify the normative import of the content I experience when confronted with these exemplars. This identification, Pettit argues, is independent of the particular exemplars that prompt me to develop the disposition. I might well have learned to select items by their red colour without ever having seen an English letterbox or a fire engine; in fact each of us (excluding the colour-blind) develops the general disposition to respond appropriately to red items through being confronted with a literally different and extremely small set of red colour samples.

Apart from their role in the identification of the rules to follow, Pettit maintains that our dispositions also serve to guarantee that we can read these rules directly. My disposition is, after

all, nothing over and above the inclination to proceed in one way and not in another way. My disposition is, in other words, nothing but an inner drive to form a particular belief, desire or intention, or to perform a particular action, and it is such a drive precisely because of the content it has.

The very intimate relation between rules and dispositions defended by Pettit is likely to support the suspicion that the ethocentric account is vulnerable to the Kripkean/Wittgensteinian thesis that an account in terms of dispositions entails that subjects cannot be wrong about how to proceed. But Pettit has an answer to this threat: the direct readability does not exclude the fallible readability, because the relation between rules and dispositions is only so intimate as to be a priori in favourable circumstances. And because subjects can never know in advance (i.e. at the time of being inclined to respond in a particular normatively constrained way) that the circumstances are favourable, it is always a real possibility that they are mistaken about the normative import of the content they experience.

This fallible readability leads to the conceivable possibility that subjects that seem to be similar in all relevant ways and that are in the same situation (and thus will seem to experience the same contents) will nevertheless be inclined to respond in diverging ways, e.g. by forming different beliefs, desires or intentions, or performing different actions. The actuality of such diverging extrapolative inclinations calls for the second part of Pettit's ethocentric account: our practices of negotiation. Should we notice a divergence in our spontaneous responses (and such divergence can be an intrapersonal or interpersonal affair), we will step back and assume that something has gone amiss. Divergence in the extrapolative inclinations of relevantly similar, equally co-authoritative subjects in identical circumstances calls for an explanation that could help us determine (and, consequentially, neutralise) the perturbing and/or limiting factors. Divergence requires that we see ourselves as less than decisive authorities on the rules to follow, as fallible employers of the rule and as negotiating parties with an interest in restoring convergence. The rules that it is right for us to follow in a particular situation, are, according to Pettit, those that correspond to the inclinations that would survive negotiation over discrepancies.

It is because of the crucial role of the *habits* of response and the *practices* of negotiation that are characteristic of the way in which we succeed in determining the normative import of the contents we experience, that the account outlined above is called *ethocentric*.

2. What is 'normal'?

Pettit's ethocentric account of rule-following implies a so-called global response-dependence account of concepts and properties.[5] The contents we are able to experience — according to the

ethocentric account — in a way that allows us to determine their normative import, are contents determined (as to their normative import) *in virtue of* the regularity displayed by our converging habits of response. The fact that we are disposed to subsume a particular experienced object under the concept 'box', and its colour under the concept 'red' (i.e. the fact that it *seems to be* a box and that it *seems to be* red), is a fact not merely about our habits of response, but also, *and by the same token*, about reality, about the experienced object's *really being* a red box. It is of course crucial to immediately add two qualifications that belong to the core of the theory.[6] Not just any disposition is in an *a priori* way connected to reality: it is only the dispositions of (1) *normal* subjects in (2) *normal* circumstances that provide such an *a priori* path to reality.

Let me elaborate on what it means for subjects and circumstances to be *normal*. As has been noted before, the phrase 'normal' is ambiguous in an important and intriguing way.[7]

On the one hand, 'normal' just means 'usual': normal subjects and normal circumstances are those we are used to, those that are identified by our *ethos*, i.e. our habits of response and practices of negotiation. On this reading my own idiosyncratic, implicit assumptions about which subjects under which circumstances are authoritative on a particular matter, X, are essential to the meaning of X. Granted Pettit's social holism, these assumptions are not merely private; they do not reside in the subjective particularities of a solitary thinker. Yet, on this reading of 'normal' it is merely because of the dispositions of familiar subjects in familiar circumstances that a particular set of exemplars exemplifies the specific rule these subjects associate with the content presented by these exemplars in those circumstances. In this sense I might, for instance, have learned to use the term provo for people who resemble my uncle Ben, relying implicitly on the habits of response of the members of my family who, if confronted with people that, as far as I could discern, had something in common with uncle Ben, noticed that these were provos, like Ben.

On the other hand, however, 'normal' means 'ideally suited' in the sense of 'the absence of any distorting factors'. Normal subjects are those that are well-equipped to detect the normative import of the contents exemplified by a particular set of exemplars, and normal circumstances are those that are favourable to the detection of the normative import in question. On this reading the appropriately equipped subjects and the appropriately favourable circumstances are, as it were, identified in terms of the contents 'out there' to be detected. It is not so much the case that the dispositions of familiar subjects in familiar circumstances are decisive; it is rather that those circumstances should be sought-after that facilitate the detection of the contents 'out there' to be detected, and those dispositions should be decisive that are 'attracted by'[8] these contents. This means that, in a way, it is merely because of the specific rule that applies to a particular case that certain exemplars exemplify this rule to qualified subjects in ideally suited circumstances. In this sense we could think of some of our extremely powerful measuring equipment as normal in the

sense of 'ideally suited' — say a camera-operated stopwatch that tells us that Carl Lewis needed 9.88 seconds to run 100 metres.

It is important to stress the distorting one-sidedness of each interpretation. In their extremes both readings provide, one could say, limiting concepts of normality. Neither is it the case, as I shall put it, that the three-term relation of exemplification could be replaced by a two-term relation of instantiation (as if a particular content instantiates no more and no less than exactly one normative constraint for every thinking subject that could entertain a thought with that content), nor by a two-term relation of projection (as if a particular subject endows some content with a normative import that is completely alien to this content).

Thus, although it might be the best we have, someone in possession of a suitably installed and properly working camera-operated stopwatch is not by definition an ideally suited observer in favourable circumstances such that it is a priori true that the length of a period *is* the way it *seems to be* to this observer. After all, two stretches of time that according to this stopwatch each took 9.88 seconds need not be equally long to every subject capable of thinking about the length of these periods (not, for instance, to subjects capable of measuring time by three decimal places, let alone in *nanoseconds*). On Pettit's response-dependent account, the length of a period is a three-term relation between the experienced period (i.e. the exemplar), the objective measure of time (i.e. the rule exemplified) and the normal subjects in normal circumstances (whether or not they are assisted by measuring instruments that radically improve the quality of their responses). The account entails that we cannot eliminate the role of these normal subjects in normal circumstances and still talk about the normative import of a particular content.

With respect to the other limit it will no doubt be less of a surprise to accept that a person who is immediately recognised by anyone of my family as unquestionably identical to uncle Ben in all relevant respects, need not really be a provo, even if we were to assume for the sake of argument that uncle Ben himself *is* indeed a provo, and even if this other person definitely seems to be a provo to familiar subjects in familiar circumstances. The point is that, assuming it is a property, the property of 'denominably being a provo' is, on the response-dependent account, a three-term relation between a particular person (i.e. the exemplar), the property of 'noumenally being a provo' (i.e. the rule exemplified) and the normal subjects in normal circumstances.[9] Just as we cannot eliminate the role of normal subjects in normal circumstances, as we saw above, we cannot eliminate the role of the world and the properties it has in and of itself (i.e. the role of the rule exemplified).

There is, however, a problem related to the inevitability of the role of normality conditions given the intrinsic ambiguity of the concept of normality. For although we cannot talk about contents and their normative imports independent of the world and the properties it has in and of itself, the

inevitable mediating function of normal subjects in normal circumstances eliminates the unproblematic employability of this world and its properties in determining which subjects and which circumstances are normal in the sense of being ideally suited to the detection of any particular response-dependent property. That is, because response-dependent properties *are* in part how they *look*, it would be naive and also viciously question-begging to turn to the world and its noumenal properties in order to claim that favourable circumstances and appropriately equipped subjects are those that generate experiences the contents of which would make these properties look as they are. The fact is that on a global response-dependency account one cannot turn to the world and the way it is in and of itself prior to and independent of determining which subjects and which circumstances are normal.[10] This is to say that as soon as we come across diverging responses and have to enter a practice of negotiation, there will already be in play a number of implicit assumptions about who are the normal subjects and which are the normal circumstances. And the important thing to notice is that, because of the intrinsic ambiguity of the term 'normal', articulating the import of these assumptions will of necessity lead us in different directions. For, making explicit my assumptions about what is normal, will on the one hand lead me to specify which subjects and which circumstances I am familiar with, but will on the other hand lead me to specify which subjects and which circumstances are ideally suited to the detection of the rule I'm thinking of.

I think it is a mistake to maintain, or to take for granted, that these different ways of specifying normality conditions (one in terms of what we are used to, and one in terms of what we are after) can be reduced to one another. And I think Philip Pettit makes this mistake.[11] But I will not press that point here. What I want to do instead is emphasise the crucial role of *normal subjects* in the account, a role neglected or at least underestimated by Pettit.[12] Normal subjects (and the ambiguity resulting from not distinguishing between subjects we are used to and subjects who are ideally suited) have a crucial role, because they are the ones responsible for the implicit assumptions about normality conditions that underlie the very possibility to speak of normative constraints at all. These assumptions, as I shall try to show in the next section, are themselves a major perturbing and/or limiting factor. That is why one of the main tasks to be accomplished in our practices of negotiation is to explore and articulate these assumptions. I therefore consider it, notwithstanding the fact that the role of these assumptions is underestimated, one of the merits of the ethocentric account of our capacity to reason that it gives a prominent role to the specification of normality conditions. I shall attempt to clarify this claim by paying attention to practical problems that arise because of the fact that we find ourselves in *new* situations.

3. How to be rational in new situations

Let me recapitulate. First, I sketched Pettit's ethocentric account of rule-following, starting from the assumption that it is basically a story about our capacity to reason, i.e. our capacity to determine the normative constraints that apply to the situations we find ourselves in. Next, I argued that implicit and intrinsically ambiguous assumptions about normality conditions have an important role to play in Pettit's account. And what I suggested in the final paragraph of the previous section is that it is (perhaps contrary to expectations) an advantage of Pettit's account, as an account of our capacity to reason, that it can accommodate a prominent role for the need to articulate and specify assumed normality conditions. I take this to be a major advantage of the basic structure provided by the account because I think our capacity to reason is to a large extent nothing over and above our capacity to articulate, specify and evaluate our implicit assumptions about what is normal and what not.

I am not, however, in a position to argue for this at length here. In the remainder of this paper I shall therefore do two things. In this section I shall discuss a couple of cases in which an agent has to make a move in a radically new situation, and in the final section I shall discuss one corollary of my adaptation of the ethocentric account: the predominant role of *intrapersonal* divergence and the related, merely criteriological role of convergence with respect to determining the right rules to follow.

I shall concentrate on new situations for four related reasons. Firstly, new situations are particularly challenging for an account of rule-following that gives a major role to habits. For, although habits are such that they will always suggest which option to take, they are also commonsensically speaking not straightforwardly attuned to new situations, but marks of being familiar with the familiar. Secondly, given the likely friction between habit and novelty, new situations are an interesting playground for exploring the need to articulate and evaluate one's implicit assumptions about what is normal. Thirdly, rule-following as a matter of extrapolation, of 'knowing how to go on in the right direction', is in the Wittgensteinian tradition commonly connected to questions about the next (i.e. the new, as yet unknown) number to be added to a given sequence. And finally, new situations are particularly challenging for an account of rule-following that is non-sceptical, because new situations are likely to be either such that the novice does not know which rules apply because she has not been informed about prevailing conventions, or such that they are entirely devoid of normative constraints.

Consider the following cases:

(a)   Anna moves to a new neighbourhood. It takes a while before she understands how to go about greeting her neighbours.

(b)   Brenda and her colleagues go out for the weekend. The informal context creates many misunderstandings.

(c)   Carla lives in a city that has undergone many drastic changes due to the arrival of large numbers of newcomers from all over the world. She is often insecure about what to do.

(d)   Doris is a member of a committee installed to explore and determine the stance of her political party towards genetic engineering. She finds it extremely difficult to make up her mind.

(e)   Edith is confused by an artificially intelligent robot which teases her and cries that she should not turn it off.

These cases differ in significant ways. Anna enters a new situation that is very familiar to everyone she meets. She is the only newcomer, as a consequence of which everyone, except Anna, seems to know quite well how to behave. Anna might for instance think that greeting your neighbour means greeting everyone you meet in the street, whereas her neighbours tend to neglect her at the beginning because they only greet people they've seen before, or on a regular basis, in their street.

Brenda and her colleagues enter a situation that is new to all of them, even though they are familiar with one another in different circumstances. Brenda knows quite well how to cope with her colleagues, but not in these circumstances. For instance, it is not customary to discuss private matters with colleagues at the office, but what marks the difference here — the colleagues or the office? After all, they will be back in the office on Monday morning.

Carla's case is quite different: although she didn't move, her social and cultural circumstances changed so as to be radically new to her and all her fellow citizens (for many of whom the situation might seem to be like that of Anna or Brenda). Carla's problem is less evidently a problem to be understood in terms of her personal perspective. The novelty she faces is to be characterised in terms of a much more general perspective. Although some of the old structures with which Carla is familiar still exist, they seem to have lost their normal validity. As a result, many of the normative constraints that should apply to the situation Carla finds herself in have to be determined anew.

Doris faces still other problems. Although she must obviously be well-trained in independently and authoritatively articulating the impact or normative import of the basic principles that give her political party its identity, she faces a situation that is new to humanity as such. We could think of her problem as a relative one, i.e. as mainly about how to exploit the

differences between her party and its rivals, but there is more to it than just this rhetorical question. Doris is pioneering: she's exploring the normative constraints that should apply, according to a particular ideology, in a situation with respect to which the ideology was not originally designed.

Edith's problem is different in yet another way. She too faces a situation that is new to humanity as such. The situation makes it unlikely that she will fail to recognise this, but if she does, her situation might seem to be quite similar to that of Anna or Carla: how should she behave with respect to a stranger? But actually Edith's problem is different. The stranger she meets is not a human being; it might not even be a stranger at all (if, for instance, only persons can be strangers, and if robots [fail to be = ? = are not counted as persons).

What should these women do if they want to use their capacity to reason, i.e. if they think they should identify, understand and follow the rules that apply to the situations they find themselves in? The ethocentric account provides a two-step procedure[13], of which the first step is best understood not as a deliberate step in a procedure, i.e. not as a move a person can make in order to live up to some good advice or in an attempt to consciously perform a task. After all, this first step says that the women should act in accordance with the extrapolative inclinations that spontaneously arise in the situation they find themselves in. As a procedure that it makes sense to consult, the ethocentric account consists basically in just one step: the second one. It says that we have to enter a practice of negotiation in case the first step doesn't lead to a plausible result.

This does not mean that the first step could be skipped. As we've seen above, the role of our habits of response is crucial to the capacity to follow a rule, according to the ethocentric account, because it is in virtue of these habits that we can identify and understand the rules we should follow in situations we find ourselves in. But in new situations — and perhaps such situations are a proper subset of situations in which we face practical problems (i.e. situations that can be correctly described in terms of not knowing the rule that should be followed) — the first step does not result in knowing what would be best to do.

At first sight it might seem as if this could be either a matter of facing such situations without spontaneously developing any specific extrapolative inclination at all, or a matter of responding to such situations in a number of apparently incompatible ways, i.e. by spontaneously finding oneself moved by a number of incompatible, or in any case seriously diverging dispositions. The first possibility, that the situation does not strike a chord at all, is quite unlikely to be a real possibility for mature human beings of whom it can sensibly be said that they have developed their capacity to reason. For there seems to be much truth in the claim that the development of the capacity for rational agency is fundamentally a matter of developing spontaneous habits of response.[14]

The second possibility, i.e. that the situation triggers a number of seriously diverging dispositions, is much more likely to be commonly the case in situations in which we face practical problems or — to be more careful and to restrict myself to the present discussion — to be the case in the examples presented above. Anna might feel the inclination to greet everyone she meets in the street, but she might also have the inclination to wait and see what others do (assuming all others to be more familiar with the particular customs that prevail in her new neighbourhood), or she might feel frustrated by the fact that others tend to ignore her which gives rise to the inclination to distrust her usual greeting habits. Something similar seems to be the case with Brenda and Carla: there seem to be too many plausible alternative options suggesting themselves to these women. Edith's case is a bit different, in that it is perhaps not true that many different plausible alternative courses of action suggest themselves, but the case is similar in that Edith seems to be under the spell of a number of strong and diverging reactive attitudes (as Strawson would call them[15]), attitudes that share motivational force with the dispositions or habits of response that are at play in Pettit's ethocentric account. Doris's case is perhaps one about which we could have doubts with respect to whether it is best characterised as a case in which an agent is troubled by too many diverging habits of response. But this might well be a consequence of the fact that the case suffers from a stance that is almost *too* reflective, too disengaged — a stance characterised by putting in brackets one's own inclinations to see where principles could lead one.[16]

If it is indeed characteristic of agents in new situations that they are troubled by too many diverging extrapolative inclinations, we can see, in terms of the account, why these agents should take the second step in what I've presented as a procedure. For the women in my examples are confronted with *intrapersonal* divergence. And as we've seen above, such divergence calls for an explanation in terms of perturbing and/or limiting factors. It requires, as I put it in section 1, "that we see ourselves as less than decisive authorities on the rules to follow, as fallible employers of the rule, and as negotiating parties with an interest in restoring convergence."[17] And this means, according to my reasoning in section 2, that we see ourselves forced to articulate and evaluate our implicit assumptions about normal subjects and normal circumstances, because the intrinsic ambiguity of these assumptions is a major source of possibly perturbing and/or limiting factors.

Let me introduce some convenient abbreviations in order to facilitate the identification of the four different types of assumptions that, I argue, will be at stake. These are:

FS: assumptions about who are familiar subjects
IS: assumptions about which subjects are ideally suited to detect the right normative constraint

FC: assumptions about the familiar circumstances in which the agent is used to experience exemplars of the kind in question

IC: assumptions about which circumstances are ideally suited to the detection of the right normative constraint.

Let me use Anna's case to illustrate the point I am trying to make. If Anna tries to reason her way out, she should, according to the extension of the ethocentric account I am suggesting here, try to articulate the assumptions about normality conditions that underlie her various and diverging extrapolative inclinations. Thus, she should articulate which subjects she is used to in terms of greeting habits (FS), which subjects she assumes to be ideally suited to detect the greeting rule that applies to the situation she is in (IS), which circumstances she is used to in terms of greeting behaviour (FC), and which circumstances would be ideally suited to the detection of how one should greet in the situation she is in (FS). The following stipulative extension of Anna's example does not seem to me to be implausible.[18] Anna realises that she is used to people (herself included) who always return the greeting unless they are seriously angry with the person they meet (Anna's FS); that she thinks that the strangers she meets in her new street are ideally suited to detect the greeting rule that applies to greeting behaviour in her street (Anna's IS); that she is used to overt, verbal greetings in standard terms between any people that approach one another from opposite directions up to a distance of less than 10 metres (Anna's FC)[19], and that she thinks it should be possible to determine how best to greet if one could clearly and in a non-distorted manner determine the relative values of care, concern, freedom and courtesy (Anna's IC). The obvious tension that I've made up between Anna's IS and her IC is deliberate. A plausible alternative to her IC that could be in tune with her IS would be something like (Anna's IC*): it should be possible to determine how best to greet if one knew the greeting rule that applies to people that meet in her street. But notice that the required knowledge could not be a matter of tacitly knowing how to behave (it should be knowledge *that* a particular rule is the rule that applies to the situation).

We could of course make up the peculiarities of Anna's or any other case in countless different ways, but it seems plausible to draw some general conclusions from just this single imagined case. One conclusion is that likely misunderstandings are easily produced by not distinguishing between Anna's FS and her IS. These are two distinct assumptions that could both, however, be presented in an indiscriminate manner as Anna's assumption about which subjects are normal. Confusing Anna's FS and her IS leads to the probably mistaken view that the strangers in her street are angry with her, and also to the equally implausible view that all the people in her street, except Anna herself, know how to go about greeting. A second conclusion is that similar misunderstandings could easily follow from not distinguishing between Anna's FC

and her IC (or IC*). These are clearly distinct assumptions that could both, however, be presented in an indiscriminate manner as Anna's assumption about which circumstances are normal. There is, however, nothing simply uncontroversial about implementing such abstract values as freedom, care and courtesy in concrete behaviour that involves among other things the ability to estimate a distance of less than 10 metres, to interpret the use of non-standard language and bodily gestures, to judge the measure of the neighbourhood, etc. Such implementations are always controversial, even though we use them unthinkingly on countless occasions in which we simply trust our own spontaneous habits of response. A third conclusion is that neither Anna's FS and FC nor her IS and IC need be thought of as packages that count for one. That is, there need be no logical or causal continuity between assumptions about subjects and circumstances that share familiarity nor between those that share ideal suitability. The obvious tension between Anna's IS and her IC does not make their co-occurrence in one single thinking subject logically or metaphysically impossible.

It is obvious that the second step of the procedure suggested by the ethocentric account is likely to generate a lot of material that could profitably be used to overcome the divergence that led to taking this second step. It will be possible to realise a quite substantial convergence by simply becoming aware of the ambiguity implicit in our assumptions about normality conditions. *Articulating the variety* of independent assumptions that cannot easily be combined in one single view of what is normal, might in itself be a major contribution to the evaluation of the conditions one considers to be truly decisive about what is to count as normal and what is not. But this is just an assessment. I don't think it is possible to say something of a general nature about how to reach convergence in particular situations one enters by means of holding incompatible, or in any case seriously diverging extrapolative inclinations. However, one thing strikes me as significant about new situations: they could give rise to practical problems that are best interpreted in terms of *intrapersonal* diverging habits of response. That is, the problems rational thinking agents face in new situations are not primarily a matter of *interpersonal* divergence, i.e. a matter of social conflict, but primarily a matter of intrapersonal divergence. Anna might seem to have a problem with her new neighbours that fail to greet her in the way she expects, but this is actually a problem that could well (and better) be described in terms of her own incompatible, or at least diverging assumptions about normality conditions. And the same is true of Brenda, Carla, Doris and Edith. Although their problem might be triggered by the actions of other people, it does not present itself as primarily a problem of how to cope with these other people. Of course, this will be a problem these women will have to solve too. But it seems necessary for them first to reason *their* way out of their predicament in order to be able to know what they have to agree about with their interlocutors. This leads me to the last topic I should like to say something about: the question of whether the ethocentric account with its emphasis on our practices of negotiation

entails the view that interpersonal convergence is constitutive of the fact that particular rules apply to particular situations. In the next section I shall conclude that it does not.

4. Intrapersonal convergence to a rule

One of the main results of the previous section is the plausibility of the idea that agents who are troubled by diverging extrapolative inclinations because of finding themselves in a radically new situation, are facing an *intrapersonal* divergence that could be elucidated (and overcome) by articulating (and evaluating) their implicit assumptions about normality conditions. Emphasising the pervading presence of these intrinsically ambiguous assumptions makes it clear that what might present itself as an interpersonal problem to be solved in a practice of negotiation involving all relevant agents, could perhaps better be recast in terms of the accompanying (mirroring) *intrapersonal* problem generated by these assumptions. If this is right, and if agents in new situations should reason *their* way out, the question arises how they should be able to do that according to Pettit's ethocentric account of rule-following. The question could be phrased like this: in order to reason their way out, agents in new situations have to engage, according to the account, in practices of negotiation. Such practices are typically of an interpersonal nature: they typically involve a number of interlocutors that could come to agree with one another about the right rule to follow. And the question now is what is entailed by such an interpersonal convergence used to solve an intrapersonal divergence.

Let us look at Doris's case, as it is most likely the case in which the convergence eventually reached will be of an interpersonal nature. Let us assume for reasons of simplicity that the committee Doris is a member of has merely one other member: Fred. And let's suppose, for the sake of argument and contrary to what I remarked above[20], that Doris enters the case without any serious extrapolative inclinations. She has suspended her habits of response in order to think clearly and without bias, or so she thinks, about the radically new problem of judging the acceptability of genetic engineering. How will they proceed? Well, obviously one of them has to begin by introducing a first opinion, not merely to get the discussion going, or to make the divergence visible, but also to make an eventual convergence possible at all. They cannot both continue to suppress their ideas or their habits of response. Well, suppose Fred opens the discussion by saying that genetic engineering is unacceptable because of Q.[21] How many options has Doris? She could agree, or disagree, or be unsure about the precise meaning and/or consequences of Q, and/or of her (dis)agreeing with Q, but whatever option she chooses, she won't be able to do so unless she proceeds by going along with any one of her habits of response.

This, at least, is what she cannot but do according to the ethocentric account of rule-following. Thinking presupposes the identification and understanding of rules of thought, and for this a thinking subject needs her habits of response. Thus, Doris will have to follow her spontaneous extrapolative inclinations in her reaction to Fred's proposal Q.

Now, what happens if she does? In following her habits of response Doris has to rely on her assumptions about what is normal. This, again, is what she cannot but do according to the ethocentric account of rule-following that, as I've discussed in section 2, implies a global response-dependence account of concepts and properties. This is not a very controversial claim: it is just a consequence of the fact that when we say that something is P, this implies (granted sincerity obtains) that the thing seems to be P to us, and that we assume we are normal subjects in normal circumstances (and are therefore justified in believing that the thing *is* P).

Suppose Doris disagrees with Fred in believing that the truth of Q is no reason for thinking genetic engineering is unacceptable, and let's suppose that Fred holds on to his conviction notwithstanding the fact that he understands Doris's decided rejection of his view. They clearly have diverging dispositions as members of the committee, and it is indeed part of the example that the result of this committee will set the rule for all members of Doris's political party. So there is a clear sense in which the convergence to be reached by Doris and Fred will be constitutive of the party line, and that means that their agreement will be constitutive of particular normative constraints. In that sense, the convergence to be reached by Doris and Fred will have a social role similar to the function of a constitutive rule, such as a rule of chess.[22] But what about Fred and Doris? Will they too be able to use the party's singular and coherent view as the overriding reason for their convergence? As Pettit has argued in a slightly different context, the answer should be no.[23] Why? Let me give the argument in terms of my discussion of the various assumptions about what is normal that figure predominantly in intrapersonal instances of divergence.[24] Why *intrapersonal*? After all, the case involves, if any case does, an obvious *interpersonal* divergence. It is Fred and Doris who disagree and need to come to terms with one another!

Despite this obvious fact, I believe it makes sense to argue that underlying the interpersonal divergence there are two *intrapersonal* divergences (one among the various extrapolative inclinations of Fred, and one among those of Doris) that need to be overcome first, or the resolution of which will be required for a rational agreement between Doris and Fred.[25] Let's concentrate on Doris. She believes that Q is no reason for thinking genetic engineering is unacceptable, despite Fred's persistence to the contrary. In such a situation the ethocentric account tells her to look for perturbing and/or limiting factors that might help explain the divergence. I've added to this that her (as well as Fred's) implicit and ambiguous assumptions about what is normal might be a major perturbing and/or limiting factor. As I've suggested above, the

second step of the ethocentric procedure entails that she should articulate these assumptions in their variety and evaluate them comparatively. Thus, instead of trying to reach an agreement with Fred head-on, and instead of merely trying to clarify her IC — that is, her assumption about which circumstances are ideally suited to the detection of the right normative constraint[26] — Doris should try to articulate her full set of possibly perturbing assumptions about what is normal: her FS, IS, FC and IC. She should do so in order to be able to explain the interpersonal divergence between her and Fred in terms of what she could wholeheartedly come to recognise as normal (in both senses) in the new case she is in. And it is quite likely that she will use her discussion with Fred not merely, nor primarily in order to reach an interpersonal agreement, but to *understand* the interpersonal divergence between her and Fred in terms of their mutual assumptions about what is normal and what is not. Thus, she will use Fred as a sounding board, as a useful intermediary for the articulation of the various assumptions she has taken for granted in the first part of her discussion with Fred, and that she might then use to make up her mind about genetic engineering, and to explain the interpersonal divergence between her and Fred's initial extrapolative inclinations.

One might object at this point that it is one thing to show that intrapersonal divergences probably underlie interpersonal divergences, but that, even if this is granted, it is quite another thing to maintain that convergence (of whatever nature) is merely criteriological to and not constitutive of the normative constraints that apply to a situation. For suppose — and this seems a plausible assumption in the radical kind of new situation Doris and Fred are in — that Doris thinks there are no rules whatsoever that tell anyone what to think about such an unheard of new possibility as genetic engineering. Suppose she thinks that questions about the moral value of genetic engineering are as empty as questions about personal identity in Parfit's branching teletransporter thought experiment.[27] Suppose Doris thinks the common moral qualifications we're accustomed to are of no avail at all in evaluating the radically new technique of genetic engineering. Suppose she thinks it's up to all of us to decide on the (un)acceptability of the technique. And suppose Fred thinks so too. They are moral anti-realists — at least if pressed, and if anthropocentrism is unavailable (as, they think, it is, or would be, in such cases as encounters with extraterrestrials, genetic engineering and artificially intelligent 'persons').

This is a big issue that does not stand a chance of being satisfactorily dealt with in the final paragraphs of this paper. But I should like to suggest the line of argument that seems most compelling to me, i.e. that based on Taylor's well-known criticism of the feasibility of the radical choice theory.[28] And the basic move in the present context would be to argue that it is, despite the professed anti-realism, highly unlikely that Fred and Doris will enter their new situation without any extrapolative inclinations at all. That is, the problem they will have in deciding on what to think about genetic engineering is most unlikely to be the mere formalistic problem that

a political party cannot afford to have no ideas whatsoever about genetic engineering. Put differently: if Doris understands the problem, and understands it *as a problem* (i.e. as a phenomenon she's incapable of thinking about in a coherent way), this reveals, Taylor likes to argue[29], that she experiences the situation as one that triggers a number of seriously diverging dispositions, and not as a situation that strikes no chord at all.

A second move would be to claim that, however biased these dispositions are in terms of being the mere result of a contingent upbringing, Doris cannot begin to make progress in lumping these dispositions together, unless she finds reason (in the articulation of her various assumptions about what is normal and what not) to conclude that some of the things that *seem to be* X to her *are indeed* X. And the general point of the Euthyphro argument applies here: "We don't know whether it was possible for the gods to love something spontaneously, without any inkling of the properties they started to love, but we know that it is impossible for us to have an extrapolative inclination without any inkling of the rule exemplified in the thing which arouses this inclination."[30] That is, we cannot sincerely believe that something is X *merely because it seems to be* X *to us*. We cannot but accept that something is X in virtue of its seeming to be X, unless we believe that the noumenal property that makes it denominably X *causes* it to seem to be X to us. This means, in one final formulation, that an agent can only take her extrapolative inclinations seriously (and that is what she has to do on pain of not being a thinking subject at all) if she holds herself justified to maintain that at least one of her assumptions about being a normal subject is a justified assumption about being an ideally suited subject with respect to the rule (the noumenal rule) this assumption is about. And this means that she cannot think of her contribution to reaching a convergence (in the sense of a compelling conclusion) as really being constitutive of the normative import of the convergence, as opposed to being an indication that she and her interlocutors have succeeded in determining (in the sense of discovering) the normative constraint that applies to the situation.

The point of this proposed line of argument is not to defend a queer realism. I am not saying that Doris and Fred will eventually be able to *discover* a noumenal rule waiting 'out there' to be detected, a rule that articulates what members of their political party ought to think of genetic engineering, and that has been waiting 'out there' since eternity for their party to be founded, genetic engineering to be invented, their committee to be installed and their discovery to be made. No. What I am saying is that the phenomenology of making up one's mind is such that it involves experiences of contents that are normatively rich. These contents are not, of course, straightforward intuitions of the moral properties of genetic engineering. They are the contents of our implicit and ambiguous assumptions about what is and what is not normal with respect to our stance towards genetic engineering. These assumptions could, and should, be sorted out, articulated and evaluated, and the normatively significant contents that will emerge

from this process will determine[31] the appropriateness of our beliefs, desires, intentions and actions in the situations to which they apply.

References

Blackburn, S. 1998. *Ruling Passions*. Oxford: Oxford University Press.

Bransen, J. 1994. 'Anthropocentrism in Favourable Circumstances.' *Inquiry* 37, 339-347.

Damasio, A. 1994. *Descartes' Error*. New York: Putnam.

Habermas, J. 1981. *Theorie des kommunikativen Handelns*. Frankfurt/M: Suhrkamp.

Kripke, S.A. 1982. *Wittgenstein on Rules and Private Language*. Oxford: Basil Blackwell.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Pettit, P. 1990. 'The Reality of Rule-Following.' *Mind* 99, 1-21.

Pettit, P. 1991. 'Realism and Response-Dependence.' *Mind* 100, 587-626.

Pettit, P. 1993. *The Common Mind. An Essay on Psychology, Society, and Politics*. Oxford: Oxford University Press.

Pettit, P. 1998. 'Noumenalism and Response-Dependence.' *The Monist* 81, 112-132.

Pettit, P. 1999. 'A Theory of Normal and Ideal Conditions.' *Philosophical Studies* 96, 21-44.

Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Stocker, M. 1976. 'The Schizophrenia of Modern Ethical Theories.' *The Journal of Philosophy* 73, 453-466.

Strawson, P. 1962. 'Freedom and Resentment.' *Proceedings of the British Academy*, 1-25.

Taylor, C. 1977. 'Self-Interpreting Animals.' in Taylor 1985, 45-76.

Taylor, C. 1985. *Human Agency and Language. Philosophical Papers I*. Cambridge: Cambridge University Press.

Wallace, R. J. 1994. *Responsibility and the Moral Sentiments*, Cambridge/Mass.: Harvard University Press.

---

[1] Pettit (1993), pp. 54-108.

[2] As developed in Kripke (1982).

[3] I take this to be the point of his assumption "that if a subject follows a rule, it is put in touch with an entity [that satisfies the objective condition of being normative over an indefinite variety of cases]" (p. 86).

[4] Pettit discusses three so-called subjective requirements rules should meet in order to allow that subjects can follow them. These requirements are (1) independent identifiability, (2) direct readability and (3) fallible readability. I use 'to understand' to capture the latter two requirements in one phrase.

[5] See Pettit (1991), (1998).

[6] In Pettit (1998, p. 114) a third qualification is added, a qualification that marks a sensible difference between cruder versions of the theory and Pettit's more sophisticated one: the

connection does not hold between 'seeming to be X' and 'being X, period', but merely between 'seeming to be X' and 'being denominably X'. This qualification allows for a world being a certain way in and of itself even if there are no appropriately equipped subjects in favourable circumstances in this world to whom it can seem to be that particular way. As a consequence, Pettit's version of global response-dependency does not entail anthropocentrism, despite his remarks in Chapter 4 of *The Common Mind*.

[7] Cf. Bransen (1994).

[8] This phrase is used by Pettit himself on at least one occasion: Pettit (1999), p. 38.

[9] I add the qualifications 'denominably' and 'noumenally' in accordance with Pettit (1998). Cf. footnote 5 above.

[10] Cf. Pettit (1993), p. 92.

[11] Cf. Bransen (1994).

[12] Pettit was brave in attempting to develop a theory of normal conditions (Pettit, 1999), but his theory pays no attention at all to the fact that the mental make-up (the intentional and non-intentional background) of the subjects that do their thinking under these conditions should be recognised as a major feature of these conditions. We need a theory not only of normal circumstances, but also of normal subjects.

[13] Pettit's ethocentric account of rule-following is not designed as a theory of deliberation. But as an analytical account of how rule-following is possible for deliberating agents, it seems an account that could easily be redesigned as a procedure agents could execute in deliberation, in making up their mind about what would be the best thing to do.

[14] Cf. Taylor (1977), Wallace (1994). Cf. also Blackburn (1998) and Damasio (1994).

[15] Strawson (1962).

[16] And precisely this attitude might be fatal to reasoning one's way out of a practical problem. Cf. Stocker (1976).

[17] Cf. Pettit (1993), p. 9, for an almost similar phrase.

[18] Ironically, if this appears to be implausible to the reader, the scheme presented might be used by the reader to make *my* point.

[19] Something such as the following is a likely additional specification: depending on the kind of relationship that exists between any two people that meet in the specified way, there are differences with respect to the time that has passed since their last meeting, and the time that has passed since their penultimate meeting. It is unlikely that Anna would continue for hours to greet her neighbour from next door every five minutes.

[20] In Section 3, pp. 19-20.

[21] For present purposes it is unnecessary, and could even be distracting, to provide realistic considerations as though I were reporting a discussion that actually took place. I shall therefore use abstract symbols.

[22] Cf. Searle (1969).

[23] Pettit (1993), pp. 299-302.

[24] This argument is congenial but not similar to the one developed by Pettit.

[25] I assume that in this case a rational agreement is not of the economic but of the properly political variety. I shall for that reason also assume (no doubt counterfactually — Cf. Habermas, 1981) that Fred and Doris have a discussion under ideal speech conditions. That is, Fred and Doris are not at all engaged in making impressions on one another. They are, or so I assume, sincerely interested in making up their minds. They are looking for a compelling conclusion, not

an attractive bargain. (Cf. Pettit, 1993, p. 291) Political parties cannot, of course, defend their party views (which should be distinguished from the compromises they might defend as part of a coalition) in terms of the best compromise that could be reached among the members of the party.

[26] According to Pettit, this is mainly what will happen once interlocutors start talking about the right-making, non-hypothetical property of the preferred option. Pettit (1993) pp. 300-302.

[27] Parfit (1984), pp. 119-217.

[28] Taylor (1985), pp. 27-35.

[29] Cf. Taylor (1985) pp. 29-30.

[30] Bransen (1994), p. 345.

[31] This is another, intriguingly ambiguous term, comprising both 'deciding' and 'discovering'.