

# Proleptic Self-knowledge

Jan Bransen, Radboud University Nijmegen

## Abstract

In this paper I discuss whether it would be a good strategy to improve one's self-knowledge by looking for predictions about one's own behaviour. I argue that the strategy is self-defeating. This has consequences for the philosophy of the behavioural sciences, because my argument undermines the claim that it would always be a good strategy to prefer predictions to expectations, epistemic reasons to normative reasons, and the quality of one's own epistemic capacities to the quality of other person's normative capacities. The argument does not only limit the scope of this kind of scientific approach to social action. There is a more serious drawback: we rely on predictions and on predictions alone whenever we are not in a position to switch back and forth between predictions, expectations and intentions. Always preferring predictions would seriously diminish the resources of our attitudinal and behavioural rationality. To improve our self-knowledge we should rather look for situations and scenarios in which we could switch back and forth between predictions, expectations and intentions. This is not an anti-scientific argument; it is rather a concern for situating the results of the behavioural sciences within the context of a broader conception of rationality.

**Keywords:** self-knowledge, prolepsis, practical reasoning, predictions, habits, agency, epistemic reasons, expectations, normative regularities, causal regularities, behavioural science (philosophy of).

## 1. Introduction

In this paper I shall address the following question: Could an agent have reliable beliefs about his own habits from which he could derive predictions of his own future behaviour that *as predictions* he could use to improve the success of his actions? I shall defend a negative answer, but not merely by arguing that using these predictions as input to practical reasoning would change them into expectations or intentions. My main argument shall be that by using predictions of one's own future behaviour as input to one's deliberations, an agent destroys the reliability of his beliefs about his own habits, thus obstructing the very possibility to make predictions of his own future behaviour.

This answer evokes two further questions: (1) Could an agent have reliable beliefs about his own habits? and (2) Are there striking dissimilarities between the role of predictions of one's own future behaviour and the role of predictions of the future behaviour of one's fellow agents in one's practical reasoning. I shall answer the first of these questions affirmatively: yes, we can have reliable beliefs about our own habits. The answer to the second question shall also be affirmative,

although I shall argue that these dissimilarities are located in very different features of using predictions in practical reasoning than commonly assumed.

But let me start by discussing some of the key terms I need to be able to address my main question.

## 2. Predictions and expectations

For successful social action an agent needs a lot of reliable beliefs about the future behaviour of other agents. Painting a house together, meeting down town, playing squash, or attending a workshop is impossible for an agent without any reliable beliefs about what his fellow agents shall do. Such beliefs cannot qualify as knowledge, because behaviour still to be performed is the wrong kind of intentional object for mental states to qualify as knowledge. To qualify as knowledge, mental states should be "fact tracking",<sup>1</sup> and that implies a temporal relation between future behaviour and belief that could never occur. But although they cannot qualify as knowledge, they should nevertheless be reliable: they should

---

<sup>1</sup> Nozick 1981.

turn out to be true for social action to be successfully possible at all. How could an agent get such beliefs — beliefs that are not true, but of which the agent knows that they will be true?

There are two different commonsensical strategies for this. Firstly, an agent can make predictions. That is, an agent might derive reliable beliefs about the future behaviour of other agents from other beliefs he has that do qualify as knowledge. In playing squash I might predict the direction of your return from my knowledge of your skills, together with my knowledge of how I hit the ball. And secondly, an agent can have expectations. That is, an agent might derive reliable beliefs about the future behaviour of other agents from normative constraints that apply to those agents. In meeting you down town I might expect that you will be at the place and the time we agreed on.

In using the terms 'prediction' and 'expectation' as labels for these two strategies, I take the liberty to stress some of their connotations, and thereby make them a bit of terms of art for the occasion.

There is no neat and unproblematic division between the two strategies, perhaps not even conceptually, but the basic idea is intuitively

simple enough. Reliable beliefs are trustworthy. They are trustworthy because the agent has good reasons for them. These reasons could be epistemic, or normative, and this difference highlights a distinction between two sources in which the trust in question is grounded. If a prediction turns out to be wrong, there must have been a mistake in the epistemic reasoning. The responsibility for such a mistake resides with the agent. He was not trustworthy in this case. His prediction was not reliable. If I fail to catch your return in playing squash because my prediction about its direction was wrong, it is me, and only me, who is to blame for loosing the game. But if an expectation turns out to be wrong this need not be a consequence of a mistake in the agent's normative reasoning. He might have reasoned well, and might have formed a *justified* expectation. The responsibility for failing to meet such a justified expectation resides with the other agent. This other agent must have made a mistake in his normative reasoning. He turns out not to be trustworthy in this case. If I expect to meet you down town at the time and the place we agreed upon, and you don't show up, it is you who is to blame.

Given this distinction one might think that, whatever the plausibility of the distinction, the strategy of making predictions will always be a better way of obtaining reliable beliefs about the future behaviour of fellow agents than the strategy of having expectations. An expectation is reliable to the degree that one's fellow agents are trustworthy in their normative reasoning. But a prediction is reliable to the degree that the agent himself is trustworthy in his epistemic reasoning. Replacing expectations by predictions is therefore a good strategy for a social agent who wants to control the situation, who wants the guarantee that the quality of his epistemic and agential endeavours will pass over to the success of his actions. A social agent who can just rely on predictions is an agent who has eliminated the possible negative effects of the possible weaknesses of other agents' normative reasoning capacity on the quality of his actions.

Note that relying merely on predictions might serve two related ends. One is to make the success of one's social actions a function merely of *one's own* agency. And the other is to make the success of one's social actions a function merely of *epistemic* reasons.

An example might help to clarify the distinctions made. Suppose you are one of my students who attended a class in which I explicitly accepted to play the devil's advocate and argued for an eliminative materialist view of mind and morals. I have a tutorial with you now, and in virtue of my knowledge of the fundamental attribution bias<sup>2</sup> I predict that you will think of me as an eliminative materialist, despite the fact that you know I was just pretending in class. So, in order for the tutorial to be successful I should not expect you to have an open mind. I should not feel misunderstood if you mistake my remarks due to your disposition to think I'm an eliminative materialist. I should know better: in order to be the successful teacher I aim to be, I should on this occasion by-pass your competence as a rational, conversational interlocutor<sup>3</sup> by basing my actions on the prediction that you will be biased. My attempt might be to remove the effects of *your* contribution to the success of our social action: I watch over the quality of the tutorial by preparing for appropriate responses to the moves I predict you will make. My attempt

---

<sup>2</sup> Nisbett, R.E. and Ross, L. 1980. See the classic case of people's pro- or anti-attitudes towards Castro reported in Jones, E. E. and Harris, V. A. 1967.

<sup>3</sup> See Pettit & Smith 1996.

might also be to remove the effects of *normative* reasoning to the success of our social action: I watch over the quality of the tutorial by basing my moves on my knowledge of the regularities that govern your behaviour, and not on the normative constraints to which we are subject.

Preferring predictions to expectations, epistemic reasons to normative reasons, and the quality of one's own epistemic reasoning to the quality of other agents' normative reasoning, are striking features of an attitude that is in concordance with the dominance in modern societies of a scientific approach to social action. It is an attitude that can be defended on paternalistic grounds as contributing to the well-being of all agents involved.<sup>4</sup> But the attitude is based on assumptions that are in the end, I believe, totally unconvincing.

That is not what I shall argue for in this paper, though. My aim here is a more modest one. I aim to argue that the strategy of improving one's own self-knowledge by looking for predictions of one's own behaviour is self-defeating. The reasons supporting this argument will, I shall suggest,

---

<sup>4</sup> This is the kind of enlightenment reasoning developed in early eighteenth century by French scientific determinists such as La Mettrie and d'Holbach.

have some broader force as reasons against a scientific approach to social action.

### 3. Causal and rational regularities

In order for predictions to be reliable beliefs about the future, i.e. beliefs that will turn out to be true, they have to be based on other beliefs that are true, or at least reliable, and trustworthy so. These latter beliefs should not be themselves beliefs about the future. Deriving beliefs about the future from beliefs that are not about the future is possible because of the regularities over time that are part and parcel of reality. The idea is that when we know the state a system is in, and know about the regularities that govern the changes in the states of the system, we can derive reliable beliefs about future states of the system. Thus, I can predict that my alarm clock will ring at 7 AM if I know the alarm is on, the alarm time is set to 7 AM, and the clock is running right. I can predict this, because of the regularities that govern the clock's behaviour. Likewise, I can predict that my mom will knock me up at 7 AM

if she knows I have to get up at 7 AM to be in time for school. I can predict this because of the regularities that govern her behaviour.

But what regularities? The common story says there are two different kinds: causal and rational regularities. The clock is supposed to be a mechanism subject to causal regularities, and my mom is supposed to be an intentional agent subject to rational regularities. Causal regularities are cited in explanatory sense-making strategies, and rational regularities are cited in empathetic sense-making strategies.<sup>5</sup> Causal regularities are realised in physical infrastructures that display contingent law-like counterfactual dependencies between the various states they can be in.<sup>6</sup> And rational regularities, in contrast, are realised in scenarios that display normative counterfactual dependencies between attitudes, situations and behaviours. I shall be more precise about the details later on.

Obviously, causal regularities are most suited to making predictions about the behaviour of one's fellow agents, because causal regularities

---

<sup>5</sup> See for the distinction between explanatory and empathetic sense-making Bransen 2001.

<sup>6</sup> I rely here on Wim de Muijnck's integrated theory of causality. See De Muijnck 2003.

will hold whatever the quality of their normative reasoning. But it doesn't seem to be impossible to predict the future behaviour of a fellow agent on the basis of true beliefs about the rational regularities this agent is subject to. That is, after all, what we do in playing chess and in anticipating the best possible move of one's opponent. And it is what I do in predicting rather than expecting that my mom will knock me up at 7 AM.

Note the difference here between predicting and expecting. If I predict that my mom will knock me up at 7 AM, for example in a bet with my brother that she will whether or not I want her to, there is no room for moral sentiments towards my mom, but if I expect her to knock me up there is. In the latter case I will blame her if she forgets it, as I might feel gratitude for each time she doesn't forget. This is similar in the case of chess, where expectations, in the term of art sense that I'm using here, do not seem to be in place at all, and where there is, consequently, no room for moral sentiments towards one's opponent. This opponent isn't, after all, striving to satisfy one's predictions, and his attempt to make them false is not to blame.

The argument so far suggests a second move for those under the spell of the advantages of a scientific approach to social issues. They will not merely prefer predictions to expectations, but will tend to extend this to preferring predictions based merely on causal rather than on rational regularities. For the latter ones will bring in the normativity back again, by requiring our true and trustworthy beliefs to be of scenarios displaying normative counterfactual dependencies, thus weakening the reliability of our predictions. This suggests that in each case we rely for our predictions on the rational regularities that govern an intentional agent's attitude and behaviour, we might wish for more reliable beliefs about this agent's future behaviour if we could find ways to replace beliefs about these rational regularities by functionally equivalent beliefs about causal regularities.

But can we replace our beliefs about rational regularities by beliefs about causal regularities? Or stated differently: can beliefs about causal regularities be functionally equivalent to beliefs about rational regularities? I shall discuss this question by focussing on an agent's habits.

## 4. Habits

Habits seem to involve the right kind of behavioural regularities to admit of both causal and rational redescriptions. A habit is a constant disposition to respond in a specific way to a specific stimulus in specific circumstances. As such a habit can be understood as very much a matter of (quasi-)automatic processes taking place on a sub-personal level. As Dretske would say: a habit is the result of the recruitment of internal states as causes of specifically required bodily movements. Such a habit shows itself in behaviour that displays a causal regularity. The habit of a smoker to smoke a cigarette each time he completed a specific task could successfully be redescribed in causal terms. This might involve changes in the smoker's neurophysiological make-up that would be caused by the occurrence of the smoker's awareness of the completion of the task, and that would, in its turn, cause the smoker to reach out for a cigarette.<sup>7</sup> In Dretske's terms, the neurophysiological state that functions as the

---

<sup>7</sup> Most research on addictions nowadays focusses on the neurobiology, in particular the role of the mesocorticolimbic dopamine system. See e.g. Lingford-Huges and D. Nutt. 2003.

triggering cause between the awareness of the completion of a task and the reaching for a cigarette is recruited in a learning process of reinforcements that gives this state as a structuring cause the meaning of a desire for a cigarette. Dretske's language is suggestive enough to provide room for a redescription of the smoker's habit in terms of a rational regularity. The constant disposition to respond in a specific way to a specific stimulus in specific circumstances is in such rationalising terms a matter of proceeding in the right direction if put in the right circumstances. A habit conceived of as displaying a rational regularity is a matter of rule-following, of rightly responding to a normative constraint. This constraint is present as the content of a particular mental state (in this case the state of desiring to smoke a cigarette after completing a task) that elicits an appropriate response, namely reaching out for a cigarette.

The similarity in the rational and the causal picture of the specific response to a specific stimulus in specific circumstances is the important part of the example. We should not be distracted by the unfamiliarity to describe a smoking habit as a matter of responding in the right way to a normative constraint. It works as well the other way around, as, for

example, in a case in which an agent has the habit to check whether his mobile phone is in his suitcase before leaving the house. This habit obviously seems to display a rational regularity: it is in rightly responding to a normative constraint ("Never leave the house without your mobile.") that the habit consists. But it might well be that we can correctly redescribe the habit as displaying a merely causal regularity. We might not know enough yet about the physical infrastructure that displays a contingent law-like counterfactual dependency between the state of intending to leave the house and the state of looking for a mobile, but the current orthodoxy seems to be that somehow such behavioural habits have to be realised in the neurophysiological workings of the human body.

The claim here is a non-reductive one, congenial to the kind of logical behaviourist claims put forward by Ryle. I am not defending a claim about the nature of habits nor about the nature of rational and/or causal regularities. The question I am discussing here is merely whether there could be a functional equivalence in an agent's practical reasoning between on the one hand a belief about the rational regularity displayed in an other agent's behaviour and on the other hand a belief about the causal regularity displayed in the same behaviour of this other agent.

## 5. Proleptic Self-Knowledge

In the previous sections I did not concentrate on the specificities of agential self-knowledge. Quite the contrary: I discussed some of the concepts involved in thinking clearly about the role of knowledge of other agents' behaviour in the deliberations of an agent who wants to perform a successful social action. I discussed these concepts in the context of social action because, I believe, a discussion of agential self-knowledge — the main topic of this paper — will be more promising against the background of social action. The parallel I shall build my case on will be fruitful, or so I argue.

First a few words about agential self-knowledge. Recently there has been quite a lot of philosophical interest in what intuitively seem to be two striking features of self-knowledge: privileged access and authoritativeness.<sup>8</sup> Somehow, it seems, we are in a special intimate way aware of the contents of our own mental states (we don't need to infer these contents from gathered evidence), and in virtue of this privileged

---

<sup>8</sup> See e.g. Wright, Smith & Macdonald (eds.). 1998.

access we are, it seems, generally in a better position than others to know what we are thinking. It is important to note here that self-knowledge implies a reflexivity best understood as involving a first- and a second-order state of mind. In a state of self-knowledge I know what state of mind I am in. If I know I am happy, say, I am in a state of happiness, and on top of that, as it were, also in a cognitive state of being aware of the fact that I am happy. It is possible just to be happy, without the self-reflexive awareness. That is, I might merely be in a first-order mental state. But I should be in *some* first-order mental state in order to have the second-order cognitive state that constitutes the kind of self-reflexive awareness characteristic of self-knowledge. And the intuition is that when I know I am happy (note the two levels), I just know it, in a special intimate way that doesn't require that I infer my own happiness from evidence, and consequentially also in a special authoritative way that does imply that, generally speaking, my knowledge of my own happiness is more trustworthy than possible doubts others might have about it.

There is, however, an epistemological problem with this intuition, due to the plausibility of externalism in the philosophy of mind, i.e. the view that the content of mental states is constituted by concepts that have an

intrinsically external, i.e. public nature<sup>9</sup>. A simple example might illustrate the problem: suppose I'm twelve years old and I think I am a responsible, mature person. Suppose I explicitly think to myself, in a very self-conscious and considerate way, that I really think I am a responsible, mature person.<sup>10</sup> Epistemologically speaking this means that I ascribe a mental state with a particular content to myself, and that I know, in the sense of having a trustworthy fact-tracking belief, that I correctly ascribe this mental state with this particular content to myself. The intuition is strong enough: after all, I *know* what I think.

But wait! What particular content? We should appreciate that in the example I'm twelve years old, which would give me, quite likely, a biased conception of what it means to be responsible and mature. This should not mislead us, however, to think of my self-reflexive *Cogito*-state as involving a conceptual content biased by my twelve years old perspective. For I'm obviously not thinking that I believe myself to be responsible and mature according to the standards of a twelve years old child. But

---

<sup>9</sup> Tyler Burge's work on externalist accounts of conceptual content has motivated much of the recent interest in the issue. See e.g. Burge 1988.

<sup>10</sup> Note the double "I think that I think" to mark the self-reflexivity of these *Cogito*-states.

alternatively, we should also not conclude too easily that I falsely ascribe to myself the belief that I am a responsible and mature person according to adult standards, nor that I correctly ascribe to myself another belief, e.g. a belief with an undeterminate or blurred content. But then, what remains of the authoritativeness of self-knowledge?

There are several epistemologically respectable ways to try to solve this problem.<sup>11</sup> I shall here, however, attempt to solve the problem by taking it to be basically a practical problem, a problem of what might be called *agential self-knowledge*.<sup>12</sup> On this approach correctly ascribing to oneself a mental state with a particular content is not so much a matter of having a trustworthy *fact-tracking* belief about what one believes, but rather a matter of mental agency, of authoritatively committing oneself to a trustworthy *proleptic* belief about what one believes.

Let me hasten to make a few comments. The first thing to emphasize is that such an approach does not imply that one only believes what one

---

<sup>11</sup> Accepting the sceptical conclusion that self-knowledge has no authoritativeness whatsoever is one such epistemologically respectable ways. But there are also a number of ways to save this authoritativeness. See the collection of Wright, Smith & Macdonald.

<sup>12</sup> My approach is congenial to, but not similar to the approach defended by Richard Moran in his 2001.

explicitly endorses, or that one only has beliefs as a consequence of a preceding conscious mental act, or that introspective self-consciousness gives one the freedom to believe whatever one wishes to believe. None of this is implied. One might find oneself committed to beliefs one just happens to have. The approach I suggest attempts to save the authoritativeness accompanying one's privileged access from externalist worries. And the crucial move is to give an agential reading of one's privileged access, which leads to an authoritativeness which is not epistemic in character but normative.

An agential reading of one's privileged access to the contents of one's mind is a reading that tries to make most of the similarity between the intimacy with our actions we experience in acting and the intimacy with our beliefs we experience in thinking. Perhaps the best way to make this intuitively clear is by observing the difference between the way I am related to events, to the actions of others and to my own actions. I can, for example, wait for the printer to print out my paper, as I can wait for my student to give me her paper. But I cannot in a similar way wait for myself to write a paper. I have to *do* it. I just have to do it myself! In waiting for the printer I can profit from my predictions about the

printer's causal regularities — for example in determining the moment I could get hold of the completely printed paper. In waiting for my student I can profit from my expectations about my student's behaviour — for example in determining the moment I should remind her of the fact that I'm waiting for her paper. But I cannot profit in a similar vein from my intentions towards my own actions — they won't do it unless *I* do it. My privileged access to my own behaviour makes me the author of my actions. I cannot relate to them as events that will happen; I have to make them happen, as actions, *my* actions.

The agential intimacy implies a very different kind of authoritativeness. If I ascribe to myself the belief that I am a responsible, mature person, and if I do this correctly so, then, if someone challenges me, I should on this agential reading not think of this challenge as an *epistemic* threat to my authoritativeness. That is, in responding to the challenge so as to display my authority I should not produce epistemic evidence for my belief. I should not think of the case as of someone challenging my entitlement to a fact-tracking belief by requiring that I produce the evidence on which to base the belief (e.g. pointing to features I possess from which I could rightly derive that I do indeed

believe myself to be responsible and mature). I should rather respond to the challenge as a normative challenge so as to display my authority by producing normative evidence for my belief. I should think of the case as of someone challenging my entitlement to a proleptic belief by requiring that I produce the agential evidence that would support the belief (e.g. acting in ways that would lead the other to conclude that I do indeed believe myself to be responsible and mature).

It is in this context important to emphasize that in a case of proleptic self-knowledge the second-order state of ascribing to oneself a particular state of mind is a *belief*, a cognitive state — a belief that the agent takes to be reliable and true, and trustworthy so. It is not an intention, or a desire. It is a mental state that, like any other cognitive state, has a mind-to-world direction of fit.<sup>13</sup> That is, in caring for a fit between mind and world, a subject will take care with respect to his beliefs that his mind will serve to satisfy the world. If I believe it rains, I should take care that if evidence tells me otherwise, I should revise my belief. Conative states, in contrast, have an opposite world-to-mind direction of fit. In caring for a fit between mind and world, a subject will

---

<sup>13</sup> Cf. Searle 1983.

take care with respect to his desires that the world should serve to satisfy his mind. If I want to walk in the rain, I should take care to be out and walk if it rains, i.e. to change the world (which includes me and my behaviour) rather than revise my desires.

The self-reflexive awareness, the state of agential self-knowledge, is a matter of belief. Knowing that I am happy differs from wishing I were happy or intending to be happy, precisely in the sense that as a state of self-knowledge it is a cognitive state with a mind-to-world direction of fit. This implies that my care for a fit between mind and world will require me to revise my belief should it turn out to be the case that I'm not happy. This cognitive status of self-knowledge is not jeopardized by it also being an agential state involving prolepsis.

Let me explain this by saying a bit more about prolepsis. The term "prolepsis" has its origin in the philosophy of Epicurus. It became a key term in Hellenistic epistemology, meaning the preconception or initial conception of something that naturally comes to the human mind without special mental attention when confronted with a first instance of such a thing. Such preconceptions were considered to be indispensable starting points in philosophy. Being given (in a contingent, but still systematically

similar way to all human beings), they are what makes the search for and discovery of new knowledge possible.<sup>14</sup> In characterising agential self-knowledge as involving a *proleptic* belief about one's own mental states, I aim to connect my discussion to these ancient ideas because of the following three considerations: just as in the Hellenistic idea (1) the proleptic belief about one's own mental state is what one might call a 'flashforward' of the full content of one's own mental state that naturally comes to mind by being in this state; (2) the proleptic belief has a generic content that anticipates, provokes and precedes further mental states as being exemplars with a similar content; and (3) the proleptic belief is an indispensable starting point for the development of a mind capable of being in states with such a content.

Consider an example in which I ascribe to myself the belief that Kwibus is a dog. The ascription is a proleptic belief and as a belief it is epistemically servile to the world: the correctness of the ascription is a function of my responsibility as a cognitive subject to revise this belief, i.e. to withdraw the ascription, should it turn out to be the case that I do

---

<sup>14</sup> See Sandbach 1971, and Long & Sedley 1987 pp. 88-90 and pp 249-253. See also Bransen 1998b.

not believe that Kwibus is a dog. That the ascription is a *proleptic* belief rather than an ordinary fact-tracking belief now means that whether or not it is a fact that I believe that Kwibus is a dog (i.e. whether or not it is a fact that I am in a mental state the content of which is indeed to be individuated by the concept DOG and a token referring to "Kwibus"), is a function of the success of the proleptic state as a 'flashforward' of the full, generic content of the concept DOG and the reference to "Kwibus". But what happens now is that if the proleptic belief is a successful 'flashforward' of the full meaning of the reference to "Kwibus" and the concept DOG (and if Kwibus is indeed a dog), this 'flashforward' will anticipate, and precede, but also provoke further instances in my mind of the belief that Kwibus is a dog. It will do so, because the prolepsis (i.e. the self-ascription) presents *normative constraints* to my mind by making me familiar with the content of a particular concept. And the content of a concept just is a normative constraint to a mind. Being acquainted with Kwibus and the meaning of DOG requires of me that I believe that Kwibus is a dog.

Thus, in proleptically ascribing this belief to myself, I become acquainted with normative evidence for it, thereby displaying my

normative authoritativeness in knowing the content of my own mental states by committing myself to the belief that Kwibus is a dog. I know my own mental states in an epistemically servile way, accepting to revise my beliefs about my own mental states should they have different contents. But the self-knowledge it constitutes is at the same time an agential affair, in which my privileged access involves an authoritativeness that commits me to believe what my proleptic self-ascription entitles me to believe.

## 6. Making up one's mind

Agential self-knowledge involving proleptic beliefs about one's own mental states and characterised by privileged access and authoritativeness might, for all what I've said about it so far, seem to be a rather bizarre and outmoded idea that should not have survived the various attacks on the prominence of self-consciousness in human self-understanding. The attacks have been various: there is the Wittgensteinian strand of arguing against the possibility of a private language, there is the post-structuralist strand of arguing against the conscious subject as an autonomous author of the meaning of experience, and there is the naturalistic strand of arguing against the irreducibility of the personal level of psychological explanations. Against this background it may seem as if I try to make far too much of a competence we might seem to possess, phenomenologically speaking, but of which we should not think it plays a major role in producing and explaining action.

I appreciate the general thrust of these attacks on the 'Cartesian Theatre', but I also think the tide is turning. It seems we're up to a new level in the dialectic of human self-understanding: despite the prominence

of shared meanings, the importance of interpersonal recognition and the interdependency of personal and sub-personal psychological explanations, it seems there is some space and some need for proleptic self-knowledge, including the agential privileged access and authoritativeness it features. I am not claiming proleptic self-knowledge is a major factor in human self-understanding, but in situations of mental indeterminacy, in situations in which we have to make up our minds, it seems we cannot make progress without it. We can do without proleptic self-knowledge in most of our daily routines. I do not have to ascribe to myself the belief that I should check whether my mobile phone is in my suitcase before leaving the house. My habit will suffice. And I do not have to ascribe to myself the belief that it will rain today, nor the desire to stay dry, in order to take my umbrella. My belief and my desire themselves will take care of that. But when I don't know what to do, when I am in doubt, when I face a practical dilemma, especially one that forces me to choose between two different alternatives *of myself*<sup>15</sup>, I cannot proceed without proleptic self-knowledge.

---

<sup>15</sup> See Bransen 1996, 2000 for an analysis of the idea of an alternative *of oneself*.

I shall not argue the case decisively here. Let me just give an example that will make the claim intuitively clear enough. Suppose I am offered a job. It is a great job, it is just what I want to do, professionally speaking, but it is too far from home to commute. I shall have to move house, leave the town where most of my friends and most of my family lives. I shall see them less often; I shall have to use the phone or the internet to communicate; I shall have to get along with them in ways never conceived; to my parents, for instance, I will become a different kind of son. What should I do? I cannot wait for someone else to decide — not choosing is choosing as well in this case. And I cannot wait for myself to decide, unless, of course, I do decide (in which case I didn't wait for myself to decide after all). In such a situation I have to make up my mind, I have to determine what I believe about the job offer and what I value most. That is, I have to reach a state of determined self-knowledge, I have to ascribe to myself an attitude towards the job offer. This will be a proleptic belief about myself, about my own motivational profile, a belief that will anticipate, precede and provoke the attitude in question towards the job offer. But it will do so *qua* belief, with epistemic servility, with a mind-to-world direction of fit, with a concern to revise the belief in case

I don't happen to have the attitude towards the job my proleptic state of self-knowledge asserts I have. In reaching proleptic self-knowledge I don't just *wish* for some attitude towards the job nor do I just *intend* (all the way down to action) to have some attitude towards the job.

The example shows nicely that the authoritativeness of the state of self-knowledge, if eventually arrived, will not be a matter of being in charge of empirical evidence on which to base the claim to know one's own motivational profile, but will rather be a matter of being in charge of normative evidence that will allow me to support my commitment to a specific course of action as indicative of my own motivational profile. In this sense my proleptic self-knowledge will be self-fulfilling, being an agentially authoritative second-order state of mind that determines my first-order state of mind towards, in this case, the job offer by articulating, as a 'flashforward', the normative import of the anticipated and provoked content of this first-order mental state. Making up my mind is a matter of determining my attitude towards some object or state of affairs, and this is as much a matter of decision as a matter of discovery. It is a sequence of mental life that might, in its explicit vividness, be

rather uncommon, but it will occur in the lives of most of us. And when it occurs it will be a matter of proleptic self-knowledge.

## 7. Knowing what would be best to do

I am now in a position to connect my discussion of proleptic self-knowledge to my preceding discussion of the role of predictions and habits in the practical reasoning of an agent who tries to determine what would be best to do. My leading question here shall be whether the role of an instance of proleptic self-knowledge in sequences of practical reasoning could be taken care of by a prediction of one's own future behaviour. This question takes up the discussion of the first part of this paper (1) by acknowledging the proclaimed advantages of a scientifically inspired strategy of replacing expectations and normative reasons by predictions and epistemic reasons, and (2) by investigating the prospects for a functional equivalence in an agent's practical reasoning between a state of proleptic self-knowledge (i.e. a state that stresses an agential concern for normative constraints) and a state of predicting one's own future behaviour.

It is at first sight a bit odd to take seriously the idea of using predictions of one's own future behaviour in one's deliberations about what would be best to do. But on the other hand, instances of proleptic self-knowledge are self-fulfilling in a way that might invite us to think of them as some kind of prediction. So, imagine a case where a couple wanders down the street, looking for a place to eat. Their conversation is a matter of deliberation, of practical reasoning, of making up their minds. Now suppose this couple has wandered down this street already dozens of times, and each time they ended up in the small Thai restaurant next to the popular pizzeria. Would it make sense for the man to suggest to his wife to skip the tedious deliberation and to jump to the conclusion, by informing her about his prediction that they're bound to end up in the Thai restaurant anyway? I know this is a rhetorical question. I know it wouldn't make sense as a contribution to their conversation, at least not as a prediction. It would change the subject of their discussion; it could spoil the fun of their qualms in front of the pizzeria; it would alienate them of what they were doing, or thought they were doing, by removing their habit from their background of interaction displacing it in the center of attention as a topic to discuss. But suppose the man has been

thinking about their regular proceedings towards the Thai restaurant, and has formed a reliable belief about their habit. Could he use it to derive a prediction about their future behaviour that he could use fruitfully as input in his practical reasoning about what would be best to do?

Note that this question allows me to concentrate on the possible functional equivalence in an agent's practical reasoning between proleptic self-knowledge and predictions of one's own future behaviour. Focusing on this issue we should distinguish between the man's using a prediction about his wife's future behaviour and about his own future behaviour. Only the latter use could be functionally equivalent to a state of proleptic self-knowledge. The former case might be functionally equivalent, and preferable (given a strategy of preferring a scientific approach to social action), to an expectation of the wife's behaviour. But could the man use his knowledge of their habit always to end up in the Thai restaurant as a premise in a deliberation aimed at determining the best thing for him to do. Perhaps it could... but not as a prediction. Using the prediction that they always end up in the Thai restaurant in his practical reasoning would change the prediction into a state with normative import. It might become an expectation, an attitude giving rise to moral sentiments, of

blame, or praise, or confidence — an attitude that would no longer be based on the *belief* that it was their habit always to end up in the Thai restaurant. Or it might become an intention, a conduct-controlling pro-attitude.<sup>16</sup> Either way the prediction would not be a prediction after all. It would not be a state whose contribution to the decision making procedure was based on the fact that its content was derived from the reliable belief that it was their habit always to end up in the Thai restaurant. The content of this mental state would have a different function: it would express a normative constraint the man was expected to, or intended to respect.

But something worse would happen on top of that. If it would be possible for the man to use the prediction that they always end up in the Thai restaurant as input to his deliberations, he would destroy the reliability of his beliefs about their habit. For if it would be possible to use this prediction *once*, it would be possible to use it *always*, on each instance the couple would be looking for a place to eat. But if that would be possible, the habit would fail to be a habit that displays a causal regularity. If the couple would always end up in the Thai restaurant, and

---

<sup>16</sup> See Bratman 1987, pp. 16f.

if that could always be the result of a practical reasoning that contained as a premise the man's expectation or intention to respect a normative constraint that would lead them to go to the Thai restaurant, the main reason for their behaviour would not be a habit but an evaluative judgement, such as "The Thai restaurant is the best in town". But if that were the man's main reason, he could not develop a reliable belief about their habit always to visit the Thai restaurant without basing this belief on the evaluative judgement. The habit would be a different one: always to visit the best restaurant, which would lead them always to end up in the Thai restaurant only conditionally; namely on the condition that they continued to consider the Thai restaurant the best in town.

The consequence of this argument is fatal to the claim that predictions of one's own behaviour could be functionally equivalent to states of proleptic self-knowledge. After all, the claim proves to be self-defeating. Using predictions of one's own behaviour in one's deliberations instead of states of proleptic self-knowledge, changes the source of these predictions, thus frustrating the very possibility to make predictions of one's own future behaviour. That is, an agent can only make

predictions of his own future behaviour, if he avoids *using* them *as predictions* in determining his own future behaviour.

Does my argument imply that one cannot have reliable beliefs about one's own habits? No, it does not. I presume we all have plenty of reliable beliefs about our own habits. All I have argued for is that we should not use these beliefs to derive predictions about our own behaviour that we intend to use in deliberations we make prior to deciding on whether or not to perform the behaviour predicted. And I argued only instrumentally for this claim: my aim has been to argue against the functional equivalence in practical reasoning between predictions of one's own future behaviour and states of proleptic self-knowledge.

Does my argument overlook the dissimilarities between using in one's practical reasoning predictions of one's own future behaviour and predictions of one's fellow agents' future behaviour? No, it does not. There are obvious dissimilarities between the role these different sorts of predictions could play in one's deliberations. I have argued that we should not use predictions of one's own future behaviour *as predictions* in

one's own practical reasoning. My argument does not affect the use of knowledge of one's own habits in one's deliberations. Such knowledge could play a role of relevance in informing one's expectations and intentions. And neither does my argument affect the use of predictions of one's fellow agents' future behaviour. Such predictions might be very useful, as I argued in sections 2 and 3, as input to one's deliberations about what would be best to do.

My argument does, however, affect the ways in which we should understand, and judge, the dissimilarities between the role in one's practical reasoning of predictions of one's own future behaviour and predictions of one's fellow agents' future behaviour. The bias to prefer predictions to expectations, epistemic reasons to normative reasons, and the quality of one's own epistemic reasoning to the quality of other agents' normative reasoning, so typical of the scientific approach to social action, might seem to lead to some kind of strategic regret about the limited role of predictions in situations involving agents who try to make up their minds in conversation. The view might be that unfortunately the privileged access of the agential perspective on one's actions frustrates

the objective and disengaged perspective of science that would allow one to view one's actions as so much causal regularity displaying events.

My argument, however, supports a different view; namely, that the strategic regret is totally misplaced. Of course it is important and often possible to use predictions about future behaviour as input in one's practical reasoning. But there does not seem to be a good reason why we should always prefer predictions to intentions and expectations. Quite the contrary. There are good reasons to switch back and forth between predictions, expectations and intentions. That is what we do, and what we need to do, in making up our minds, in using our attitudinal and behavioural reasoning skills. Making up one's mind requires an attunement to the normative constraints governing the scenario one is in. Such an attunement, or so I have argued, requires proleptic self-knowledge; it requires a 'flashforward' of the full meaning of the concepts in terms of which I attune to these constraints. And in the social world of people trying to get along with one another there are more constraints that govern the scenario than just the epistemic constraints that govern the formation of reliable, trustworthy beliefs about our habits. To acquire and maintain habits we precisely need a prior appreciation of the value, or

quality, of the normative constraints we are presented with by the concepts in terms of which we attune ourselves to the world.<sup>17</sup>

Making up my mind and developing habits requires, if my analysis is right, that we can move back and forth between predictions, expectations and intentions. It is this moving back and forth that helps an agent to be successful in reaching proleptic self-knowledge. If this is right it also means that the general strategy always to prefer predictions does not lead to an improvement of an agent's practical reasoning skills, but rather to a deterioration of the agent's capacity to reason about attitudes and behaviour. The upshot of these observations is that we shouldn't as a general strategy be looking for situations and scenarios in which we have to rely on predictions, but should rather be looking for situations and scenarios in which we could situate our predictions in a comprehensive, conversational context. It might be unnecessary to note that science itself is, and fosters, such a comprehensive, conversational context.

---

<sup>17</sup> The topic has received attention in Bransen 1998a, 1998b, 2001.

## References

- Bransen, Jan (1996), "Identification and the Idea of an Alternative of Oneself", *European Journal of Philosophy*, Vol. 4(1), 1996, pp. 1-16.
- (1998a), 'True to Ourselves', *International Journal of Philosophical Studies*, Vol. 6(1), 1998, pp. 67-85.
- (1998b), 'Making X Happen. *Prolepsis* and the Problem of Mental Determination', in Bransen & Cuypers (eds.) *Human Action, Deliberation and Causation*, Dordrecht: Kluwer, pp. 131-153.
- (2000), 'Alternatives of Oneself. Recasting Some of Our Practical Problems', *Philosophy and Phenomenological Research*, Vol 60(2), 2000, pp. 381-400.
- (2001), 'Verstehen and erklären; the philosophy of', *International Encyclopedia of the Social and Behavioral Sciences*, Neil J. Smelser & Paul B. Baltes (eds.), Oxford: Elsevier Science Ltd., 2001. 16165-16170.
- Bratman, Michael (1987), *Intentions, Plans and Practical Reason*, Cambridge (Mass.): Harvard university Press.
- Burge, Tyler (1988), "Individualism and Self-Knowledge", *Journal of Philosophy*.
- Jones, E. E. and Harris, V. A. (1967). 'The attribution of attitudes', *Journal of Experimental Social Psychology*, 3, 1-24.
- Lingford-Huges and D. Nutt (2003). 'Neurobiology of addiction and implications for treatment'. *British Journal of Psychiatry*, 97-100.
- Long, A.A. and Sedley, D.N. (1987), *The Hellenistic philosophers*. Cambridge: Cambridge University Press.
- Moran, Richard (2001), *Authority and Estrangement: an Essay on Self-Knowledge*, Princeton: Princeton University Press.
- Muijnck, Wim de (2003). *Dependencies, connections and other relations. A theory of mental causation*. Dordrecht: Kluwer, 2003.
- Nisbett, R.E. and Ross, L. (1980). *Human Inference: Strategies and shortcomings of social judgment*, New York: Prentice Hall.
- Nozick, Robert (1981), *Philosophical Explanations*, Cambridge (Mass.): Harvard University Press.
- Pettit, Philip & Michael Smith (1996), 'Freedom in Belief and Desire', *Journal of Philosophy*, 1996.
- Sandbach, F.H. ([1930] 1971), "Ennoia and Prolepsis in the Stoic Theory of Knowledge", in A.A. Long (ed.) *Problems in Stoicism*. London: The Athlone Press, 1971, pp. 22-37.
- Searle, John R. (1983), *Intentionality. an Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Wright, C, B. Smith & C. Macdonald (eds.) (1998). *Knowing Our Own Minds*, Oxford: Oxford University Press, 1998.

Published in:

Jón Ólafsson & Juha Räikkä, *Rationality in Global and Local Contexts*, Turku: University of Turku Press, 2007, 5-28.