Everyday Reason Talk: An Introduction

Katrien Schaubroeck[1]

We don't always know why we do the things we do. Some of the greatest stories in the history of mankind are built on that tragic fact of human life (think of Oedipus, Othello, Madame de Bovary…). A first important theoretical account of this fact was given by 'les maîtres du soupçon' at the end of the 19th century (Freud, Marx, Nietzsche). They argued that unconscious motives, social structures and cultural particularities guide our actions beyond our knowledge, *a fortiori* beyond our control. The last three decades have witnessed the rise of a new generation of masters of suspicion, this time with a background in experimental psychology. Many of these suspicion-inducing experiments will be discussed throughout the special issue. Let us just mention one paradigmatic study by the psychologists Nisbett and Wilson from 1977. One study will suffice to open up the full range of philosophical questions addressed in the contributions to follow.

In the Nisbett and Wilson study people are invited to choose a pair of stockings from a display table. When asked for their reasons, they point to a feature of the preferred pair (its softness, its superior knitting pattern, its elasticity…). In reality however the stockings are all identical, possessing the mentioned features to the same degree. The study shows that the actual cause of people's preference for a particular pair is its position on the table. In the wake of this seminal study, dozens of experiments have been taken to show that human beings are prone to confabulation, that is to making up reasons for a judgment or an action post hoc. From this it is inferred that often the reasons we cite are illusory since we are not aware of the 'real reasons' for our actions or choices.

First of all this experiment raises the question whether we should be suspicious of reason talk in general. In everyday life we work with an image of human beings as rational agents. Though it is common knowledge that sometimes motives remain curiously hidden for the agent himself, we constantly engage in the practice of giving and asking for reasons. We treat each other as reason-responsive beings, trusting that there is a connection between (conveyed) reasons and performed actions. Do the developments in experimental psychology urge us to revise this image, and to abandon the practice?

✉ Katrien Schaubroeck
Katrien.Schaubroeck@uantwerpen.be

[1] Philosophy Department, University of Antwerp, Grote Kauwenberg 18, 2000 Antwerpen, Belgium

Before this key-question can be answered, other questions need to be resolved. Concepts need to be clarified and theoretical assumptions scrutinized. For example, it is not clear what scientists mean by reasons when they say that agents don't know the reasons for their actions. Perhaps the reasons examined in their experiments and which are uncovered as inventions, are not the reasons that feature in everyday conversations. Or, if they are, perhaps the inventory nature of the reasons cited in everyday conversations does not prevent these reasons from doing their job, such as enabling mutual understanding and distributing agential responsibility. Taking up the challenge posed by psychology studies about confabulation and implicit motives, this special issue 'Everyday Reason Talk' aspires to find out whether the reasons that people talk about on an everyday basis are illusory, or whether the fact that people talk about them is enough for them to be real.

After all, reasons are not natural kinds, their nature is heavily debated in meta-ethical circles (for a recent exemplary discussion, see Scanlon 2014). In this issue the metaphysical question what reasons are will not be addressed directly but indirectly, for example via the question what it means to act for a reason (Sie, Sandis), or what it takes to explain an action (O'Brien), or what it means for an attitude to be reasons-responsive (Vierkant and Hardt). Despite their different strategies there is a consensus among the contributors that we should not interpret psychologists' claims about the omnipresence of confabulation and the effectiveness of unconscious motives as threatening the default assumption that human beings are rational and responsible agents. In other words, there are no defenses of a debunking strategy to be found in this special issue. The contributors either deny the relevance of the psychological data for questions about agency and rationality or they argue that instead of raising suspicion these data rather enhance our understanding of the complicated ways in which human agents are guided by reasons.

The latter strategy is effectively applied by Maureen Sie in her contribution "Moral Hypocrisy and Acting for Reasons: How Moralizing Can Invite Self-Deception". Sie contests a wide-spread reading of a series of experiments by social psychologist Daniel Batson in which subjects evaluate their action as wrong and therefore try to conceal their action from others. Batson and many others interpret this as an affirmation of human hypocrisy: we advocate high moral principles but act against them while hiding our ulterior, egoistic motives. According to Sie, however, moral hypocrisy or the desire *to appear* to act morally does not undermine 'true moral agency' (as Batson has it). Rather, what the findings suggest is that we care so much about morality that when some course of action is called morally required (even if it isn't and even if we do not treat it as required in our everyday practices) this influences people's moral evaluation of their actions, even to the extent that some deceive themselves about the nature of their motivation. In support of her reading of the data Sie challenges Batson's understanding of 'true moral agency'. People often do the right thing without being able to articulate the reasons for which they do these things or the reasons for which these things are right. This is relevant for the interpretation of the Batson studies because it allows for the possibility that the subjects in fact do the right or permissible thing according to moral principles they silently accept (so they do act as true moral agents), while at the same time believing that their action violates the moral requirement presented to them by the experimenters (so they do take morality seriously, even to the extent that they want to *appear* moral above all).

Likewise, Frank Hindriks' article "How Does Reasoning (Fail to) Contribute to Moral Judgment? Dumbfounding and Disengagement" warns against a simple reading of psychological experiments as providing evidence for the non-rational picture of human agency. On

second thought the concepts in play (rationality, agency, reasoning) are too complicated and the results are too ambiguous to allow for a simple debunking of rationality. In particular Hindriks wants to rehabilitate reason against the attack by Jonathan Haidt's dumbfounding-experiments. What is special about Hindriks' endeavor is that he combats the reason-skeptics with their own weapons, building his case on evidence from psychology. He confronts Haidt's claim that moral reasoning is biased and ineffective with Albert Bandura's studies of moral disengagement which show that moral reasoning often does contribute effectively to the formation of moral judgments. Combining these results with data from Moral Identity Research Hindriks reaches the optimistic conclusion that effective moral reasoning need not be self-serving. It is not always emotion, or as Haidt says intuition, which carries the day. Reason can be powerful and unbiased since people are demonstrably capable of reasoning themselves into doing what they hold to be right.

As an activity moral *reasoning* obviously needs to be explicit and conscious. One could object against Sie that also *reasons* must be explicit in order to count as reasons. Following up on this objection it would not make sense to assess the unconscious motives upon which someone acts as more or less rational. Yet it happens that we blame someone for his immoral or irrational conduct regardless of whether the action was done for an explicit reason or an implicit motive. Tillman Vierkant and Rosa Hardt validate and underpin this practice theoretically in "Explicit Reasons, Implicit Stereotypes and the Effortful Control of the Mind". They provide a detailed explanation of what makes a state implicit as well as what makes an attitude reason-based, so that they can defend the claim that acting on an implicit bias is a form of acting for reasons, and moreover acting that we are responsible for. They develop their view in discussion with Neil Levy, according to whom our explicitly held attitudes are in a special sense relevant for moral responsibility because the assertibility of explicit beliefs entails that they have a unique way of unifying the agent, namely by being available for syntactical operations. It is interesting though not really surprising that it cannot be consciousness which separates explicit from implicit attitudes. After all, we can be conscious of our implicit biases, which makes them no less implicit. But if it is not consciousness, then what could make explicit beliefs and intentions more relevant for moral responsibility than implicit attitudes? As Hardt and Vierkant show, the fact that beliefs and intentions can be sincerely asserted and integrated in a web of reasoning does not suffice as an answer, since arguably implicit biases are also integrated in an agent's outlook even if they cannot and will not be asserted by the agent. Hardt and Vierkant argue that implicit biases are responsive to reasons in the sense that they contribute to the coherent set of views and values an agent uses to make sense of the world. It follows from their argument that the harbouring of implicit biases falls within the scope of an agent's responsibility.

As should be clear from these first three articles, psychologists who conclude from their confabulation-experiments that agents do not know for which reasons they act, and philosophers who argue that we cannot be responsible for the implicit attitudes that we hold, have a lot of conceptual explanation to do. Yet there are not only various conceptual assumptions, but also diverging theoretical commitments underneath the different positions in the debate. For example a prima facie simple claim like "confabulation entails that agents do not know the reasons for which they act" presupposes 1. a particular theory of human agency, 2. a particular understanding of successful action explanation, 3. a distinctive model of adequate self-interpretation. These three premises underpinning the argumentation of the masters of suspicion are scrutinized and criticized in the next three contributions.

In "Verbal Reports and 'Real' Reasons: Confabulation and Conflation" Constantine Sandis argues that much of the psychological literature on confabulation rests on a conflation of the

reasons for which people act with the facts within which these agential reasons are nested—or more subtly: a conflation between the reasons for which agents act and reasons why their actions occur. Applied to the Nisbett and Wilson experiment, Sandis' argument entails that agents might very well be correct when they say that they chose the pair of stockings on the right end of the display because of its superior knitting pattern. The right-hand bias is deposed by Sandis as only 'a nesting reason': an explanation of why the knitting of the stockings on the right appeared to the agent as a reason to choose them. As long as we keep the distinction between agential reasons and nesting reasons in mind we will see that uncovering a nesting reason does not reveal the 'real' reason for an action. Thus Sandis emphasizes that the position of a pair of stockings on a table is *not* the reason for which the agent chose them over another pair. Nor is the agent's own verbal reports about her reasons invalidated by psychological explanations of why the agent takes something to be a reason.

Disagreeing with Vierkant and Hardt about the following example, Sandis maintains that it is misleading to say that an implicit bias is the reason for which a committee chose the male candidate for the job. According to Sandis the real reason for choosing that particular candidate is that he is considered to be best qualified. The implicit bias plays a role, but on a different level. It does not guide one's action like reasons do. It rather disposes an agent to view the world in a certain way, in the example at hand it explains why the candidate appeared best qualified. But it is the features as we see them that guide our actions *qua* agential reasons. Sandis warns that it is dangerous to conflate the agential reasons upon which an agent acts with the implicit mechanisms that affect his conduct, because the agential reasons determine the description under which an action is intentional and an agent responsible for this action. Thus Sandis is also of a different opinion than Vierkant and Hardt with regard to the question whether one is responsible for actions affected by implicit bias.

One way to avoid the difficult task of identifying the reasons for which an agent acts is dropping the concept of reasons altogether. In "Beyond Psychologism and Anti-Psychologism" Lilian O'Brien maintains that action explanation is *not* reasons explanation (the explanation of action by appeal to a practical reason). For her, explaining an action requires attention for the agent's *reasoning*, and more specifically, it requires an explainer's simulating that reasoning. In the context of this special issue the idea that we do not need reasons in order to explain actions can seem a bit ironic. But perhaps O'Brien is right that when we are after an explanation for someone's action in everyday situations, we are not that much interested in revealing the reasons for which this person as a matter of fact acted, but rather we want to come to see the action as intelligible or, as O'Brien says, reasonable.

In opposing both Psychologism (which holds that reasons are psychological states) and Anti-Psychologism (which holds that reasons are normative states of affairs) O'Brien gives up on the realist picture of action explanation. On the alternative view that she proposes (and that she calls Proceduralism) action explanation is not a matter of capturing the existing relations between reasons, the agent's mind and the action, but rather of an explainer engaging in the process of trying to see what the agent saw in his action. The empathy put at the center of action explanation not only requires the reconstruction of the agent's deliberation but also acceptance of the reasoning as sound.

Interestingly, O'Brien takes practical deliberation (and its simulation in the context of action explanation) to involve the influence of personal sensibilities and values and other elements of one's psychology. Deliberating about which course of action to take is more than assessing the validity of an inference. Also contingent aspects of the agent's and explainer's psychology are part of the agent's reasoning and explainer's simulated reasoning. That is why for example

someone might fail at explaining an agent A's action at one point in his life while succeeding at a later point.

This broad conception of practical reasoning is also behind Strijbos and de Bruin's ambition to develop a model of self-interpretation that integrates subliminal influences and unconscious motives as objects of self-management. In "Self-interpretation as First-person Mindshaping: Implications for Confabulation Research" they take issue with another purported consequence of confabulation studies, namely that we are bad at knowing ourselves. Whereas Hindriks deplores the fact that confabulation gives reasoning a bad name, Strijbos and de Bruin go further and wonder why confabulation itself should have a bad name. They show that confabulation only undermines the authority of self-attribution of mental states if this self-attribution is understood on a mindreading model, namely as representing a present or past state of mind. However, on a mindshaping model the self-attribution of mental states plays a role also in regulating *future* behaviour, a role whose valuable function is not undermined by confabulation. On their model the adequacy of self-interpretation does not depend on whether the self-interpretation represents a true state of affairs, but on whether the interpretation is *made* true. Now imagine one particular subject, who indeed erroneously took her preference for the far right pair of stockings to be the result of its superior knitting pattern. Strijbos and de Bruin suggest that she can restore her first-personal-authority when she *starts to live up to* her confabulation by buying only panty-hoses with this knitting pattern, recommending it to friends, etc. In this sense Strijbos and de Bruin's model of mindshaping gives an important role to future-oriented aspects of self-interpretation. Secondly, it creates room for self-regulation in cases where intentional direct control is ineffective. As the authors point out in agreement with Victoria McGeer, an agent can live up to his own commitments not only through acting on his deliberative conclusions, but also through recognizing and managing factors that influence his behaviour beyond the scope of his intentional and reason-mediated control. Through specific techniques (such as reorganizing one's environment, silencing certain stimuli, modifying a decision procedure) we can, as it were, 'tame' the unconscious influences on our behaviour so that they no longer prevent us from acting in accordance with our explicitly avowed commitments. Importantly Strijbos and de Bruin argue that self-regulation or self-management is just as well an expression of our autonomy as is guiding our actions through explicit deliberative avowal. So in a way they agree with Sandis that implicit biases and other psychological influences should not be accorded the status of 'the real reasons for our actions'. Yet in agreement with Vierkand and Hardt, they argue that people are responsible for letting implicit biases affect their behaviour in undesired ways. People should try to change these biases or at least diminish their influence through non-deliberative self-regulation.

Also Jan Bransen's contribution "Self-Knowledge and Self-Love" espouses the idea that we should take responsibility for the subliminal influences on our behaviour. Yet he finds this responsibility discharged not so much in the adoption of self-regulating techniques to protect ourselves against their unwanted interference, but first and foremost in the simple acknowledgment of these elements of our psychology as part of ourselves. When agents try to make their actions and choices understood, the self-attribution of reasons should be seen as a display of self-knowledge. Self-knowledge for Bransen is not a purely epistemic ideal of discovering which mental states one has. Rather self-knowledge is about relating to oneself in a truthful and caring manner, appropriating thoughts, feelings and actions as one's own, even if one recognizes their irrational (or non-rational) nature. It seems obviously true that self-knowledge requires truthful acceptance of how our minds happen to work. But the originality of Bransen's

position consists in the additional claim that being truthful, or in other words facing up to the limits and flaws of one's agency, requires self-love. A self-hating agent will be too harsh and merciless in resisting the thought that human agency is necessarily flawed and that her rational capacities are inevitably finite. A self-indulgent agent on the other hand will be too kind on herself, externalizing her rational failures as if they were imposed by the environment. Only a self-loving agent will strike the balance and will achieve an understanding of her agency which is worth calling self-knowledge.

If Bransen is right, we have every reason not to fear the results of psychological research on confabulation and implicit motives. A self-loving agent does not shy away from facing her own limits and shortcomings as a rational agent. To the contrary, she wants to know and to acknowledge her psychological set-up with its limitations and imperfections as this enables her to know herself and to determine the scope of her responsible agency. In the end all the contributors to this issue unite in their conviction that a deeper insight into the workings of the human mind does not call for suspicion but for a refinement of our philosophical models and an enrichment of our everyday reason talk.

References

Nisbett R, Wilson T (1977) Telling more than we can know. Psychol Rev 84:231–295
Scanlon TM (2014) Being realistic about reasons. Oxford University Press, Oxford