Conscious Intentions

The Social Creation Myth

Elisabeth Pacherie

What are intentions for? Do they have a primary purpose or function? If so, what is this function? I start with a discussion of three existing approaches to these questions. One account, associated with Michael Bratman's planning theory of agency, emphasizes the pragmatic functions of intentions: having the capacity to form intentions allows us to place our actions more firmly under the control of deliberation and to coordinate our actions over time. A second account, inspired by Elizabeth Anscombe's theory of intentions, emphasizes their epistemic function and their contribution to self-knowledge. A third account, developed by David Velleman, suggests instead that the capacity for intentions may be an accident or a spandrel, that is, a byproduct of some more general and fundamental endowments of human nature. I argue that these accounts are at best partial and largely overlook two important dimensions of intention. I introduce and motivate a further pragmatic function of intentions, namely their role in the control and monitoring of ongoing action and argue that acknowledging the existence and importance of this function allows us to plug gaps in these accounts. I further argue that this pragmatic function of intentions plays a crucial role in contexts of joint action where agents must align their representations in order to coordinate their actions towards a joint goal. I speculate that a capacity for conscious control might have become established because of the role it served in solving inter-agent coordination problems in social contexts and because of the benefit conferred by the forms of cooperation it thus made possible.

Keywords
Action coordination | Conscious action control | Intention | Joint action | Planning | Representational alignment | Self-knowledge

Author

Elisabeth Pacherie
elisabeth.pacherie@ens.fr
Ecole Normale Supérieure
Paris, France

Commentator

Andrea R. Dreßing
andrea.dressing@uniklinik-freiburg.de
Klinik für Neurologie und Neurophysiologie
Universitätsklinikum Freiburg
Freiburg, Germany

1 Introduction

What are conscious intentions for? What do we gain from having a capacity for intentions as opposed to simply a capacity for desire-belief motivation? Do intentions have a function not just in the sense that they have a causal role but in the normative sense in which having this function confers benefits on intention-forming creatures that explain why these creatures have this capacity. In other words, do intentions have a teleofunction? Is there something they are for? And if so, what is this teleofunction?

Roughly, the notion of intention is that of a mental state that represents a goal (and means to that goal) and contributes through the guidance and control of behavior to the realization of what it represents. Thus, my intending to go to my office will control and guide my behavior (e.g., leaving my house, taking the bus, walking from the bus stop to my office), thus contributing to the realization of the goal represented by the intention. Many philosophers hold the view that if we do something intentionally, we must be aware of what we are doing.

Therefore, they consider that it is of the essence of intentions to be conscious. I have argued elsewhere (Pacherie 2008) in favor of a notion of motor intentions whose contents may not always be accessible to consciousness. On my view then, the phrase "conscious intentions" need not be pleonastic. Here, however, my focus will be on intentions qua conscious states and I will use "conscious intentions" and "intentions" interchangeably.

In his 2007 paper, "What good is a will?", David Velleman considers the question whether the human will, understood as the capacity for (conscious) intentions, has a purpose or teleofunction. He discusses two accounts that assume that the will has a purpose but disagree on what this purpose is. On one account, associated with Bratman's planning theory of agency, the primary function of intentions is pragmatic: having the capacity to form prior intentions is good because it allows us to place our actions more firmly under the control of deliberation and to coordinate our actions over time. On the other account, inspired by Anscombe's theory of intentions, the primary function of intentions is epistemic. Intentions are good because they provide self-knowledge: an intention on which one acts provides us with a special kind of knowledge of what one is doing.

David Velleman is himself skeptical that the attitude of intention has a teleofunction. Rather, he suspects that the human will is an accident or a spandrel, that is a byproduct of some more general and fundamental endowments of human nature. Velleman suggests, however, that our hypotheses about the origins of the will, including his own, must be closer to creation myths than to scientific theories. Talk of myths, of course, has both negative and positive connotations. On the negative side, myths are, if not downright false or unfounded, at least ultimately unverifiable. On the positive side, myths are dramatization devices that serve to highlight, and make sense of, the value or function of a practice, of an institution or, in the case at hand, of a cognitive capacity. Here, I will offer my own creation myth for intentions, a myth that emphasizes the social dimension and social function of conscious intentions. The main claim I will defend is that having conscious intentions is a good thing in large part because it facilitates coordination and cooperation with others and because cooperation is itself fitness enhancing. My aim in proposing this social creation myth is not to entirely displace other creation myths, but rather to complement them, to highlight an important facet of conscious intentions that traditional philosophy of action has tended to neglect and to plug some holes in the stories told in other myths.

Here's how I will proceed. In section 2, I will present the two creation myths considered and rejected by Velleman and discuss some difficulties they raise. In section 3, I will discuss Velleman's own creation myth. In section 4, I will introduce and motivate a pragmatic function of intention largely overlooked by these creation myths, namely their role in the control and monitoring of ongoing actions. In section 5, I will tell my own social creation myth. I'll argue that this pragmatic function of intentions plays a crucial role in contexts of joint action where agents have to align their representations in order to coordinate their actions towards a joint goal. I'll speculate that the main evolutionary benefit conferred by a capacity for conscious intentions is that it enables a considerable increase in the possibilities for joint action and cooperation.

2 Two teleological creation myths

Velleman (2007) points out a methodological assumption common in functionalist psychology, namely the assumption that our attitudes or cognitive faculties have a function not just in the sense that they have a causal role but in the sense that they have a purpose, something they are designed to do and thus ought to do. Functions in this latter sense are commonly called teleofunctions. This methodological assumption needs not entail a belief in some intelligent designer. Instead, it can be cashed out by appealing to evolutionary theory and to natural selection as a blind designer. Typically, the evolutionary story goes like this: a trait has the teleofunction of producing effect E just in case producing this effect conferred some benefit that

contributed to the reproductive success of organisms endowed with the trait and, thereby, to the propagation of the trait itself. This methodological assumption, when it guides our inquiry into intentions, leads us to take the question what intentions are for, i.e., what purpose are they meant to serve, as necessarily meaningful and demanding an answer.

Velleman discusses two teleological stories meant to answer this question. He links the first story to Bratman's theory of intentions (1987) and the second to Anscombe's theory (1963). I start with the story inspired by Bratman's theory.

We are, in Bratman's words, planning agents regularly making more or less complex plans for the future and guiding our later conduct by these plans. This planning ability appears to be if not unique to humans at least uniquely developed in the human species. People can, and frequently do, form intentions concerning actions not just in the near but also in the distant future. Why should we bother forming future-directed intentions? What purposes can it serve? What benefits does it bring us? What features of future-directed intentions allow them to serve these purposes?

Bratman offers two complementary answers to that challenge. The first stems from the fact that we are epistemically limited creatures: our cognitive resources for use in attending to problems, gathering information, deliberating about options and determining likely consequences are limited and these processes are time consuming. As a result, if our actions were influenced by deliberation only at the time of action, this influence would be minimal as time pressure isn't conducive to careful deliberation. Forming future-directed intentions makes advance planning possible, freeing us from that time pressure and allowing us to deploy the cognitive resources needed for successful deliberation. Second, intentions once formed commit us to future courses of action, thus making the future more predictable and making it possible for agents to coordinate their activities over time and to coordinate them with the activities of other agents. Making deliberation and coordination possible are thus the two main benefits that accrue from a capacity to form future-directed intentions.

What makes it possible for future-directed intentions to yield these benefits is, according to Bratman, the fact that they essentially involve commitments to action. Bratman distinguishes two dimensions of commitments: a volitional dimension and a reasoning-centered dimension. The volitional dimension concerns the relation of intention to action and can be characterized by saying that intentions are "conduct-controlling pro-attitudes" (Bratman 1987, p. 16). In other words, unless something unexpected happens that forces me to revise my intention, my intention today to go shopping tomorrow will control my conduct tomorrow. The reasoning-centered dimension of commitment is a commitment to norms of practical rationality and is most directly linked to planning. What is at stake here are the roles played by intentions in the period between their initial formation and their eventual execution. First, intentions have what Bratman calls a characteristic stability or inertia: once we have formed an intention to A, we will not normally continue to deliberate whether to A or not. In the absence of relevant new information, the intention is rationally required to resist reconsideration: we will see the matter as settled and continue to so intend until the time of action. Intentions are thus terminators of practical reasoning about ends or goals. Second, during this period between the formation of an intention and action, we will frequently reason from such an intention to further intentions. For instance, we reason from intended ends to intended means or to preliminary steps. When we first form an intention, our plans are typically only partial, but if they are to eventuate into action, they will need to be filled in. Thus intentions are also prompters of practical reasoning about means. Third, because intentions are commitments to action, our intentions should be jointly executable. Finally, taken together the volitional and the reasoning-centered dimensions of commitments help explain how intentions can promote coordination. They provide support for the expectation that agents will act as they intend to and these ex-

pectations are central in turn to both inter- and intra-personal coordination. In particular, this is what motivates the rational agglomerativity requirement on intentions, i.e., the requirement that my intentions be jointly executable.

The benefits that accrue from a capacity for intentions are, ultimately, pragmatic benefits. As Bratman puts it, future-directed intentions "enable us to avoid being merely time-slice-agents" (1987, p. 35). Instead of constantly starting from scratch in our deliberations and simply weighing current belief-desire reasons, intentions allow us to become temporally extended agents. They provide a background framework that allows us to expand the temporal horizon of our deliberation while at the same time narrowing its scope to a limited set of options. In so doing they contribute in the long run to our securing greater desire-satisfaction than simple desire-belief practical reasoning would.

Velleman (2007) sees three main problems with Bratman's pragmatic account of what intentions are for. The first problem concerns the status and role of present-directed intentions. On Bratman's account, a future-directed intention requires a present-directed intention to convey its motivational force and guide the action once the time to act is seen to have arrived. Bratman identifies no further role or function of present-directed intentions beyond conveying the motivational potential of future-directed intentions. At the same time, he insists that intentional actions, whether or not they are preceded by future-directed intentions, always involve present-directed intentions. This leaves us with a potentially large class of spontaneous intentional actions that involve present-directed intentions but are not preceded by future-directed intentions. These intentions do not incorporate the results of any prior deliberation, they don't set the stage for any further planning and they don't provide a basis for any coordination. The first worry raised by Velleman is thus that these intentions do not seem to serve any of the pragmatic purposes that, on Bratman's account, constitute the raison d'être of intentions. Second, Velleman points out that a similar worry arises for the intentions involved in various cases of planning. He illustrates his point with a voting example. He argues that while there may be good reasons for my starting to think about my vote in advance, such as giving me sufficient time to deliberate, there doesn't seem to be any good reason for settling in advance of my arrival in the voting booth whom I will vote for. On the contrary, settling in advance seems to carry potential costs, by making me resistant to reconsideration, without procuring any benefits, since the actual act of casting my ballot doesn't require any particular prior preparation. Thus, at least in these cases where no further planning is needed once one has settled on a course of action, it is unclear what purpose settling in advance could serve.

Velleman's third worry relates to Bratman's view that intention need not imply belief. Bratman indeed maintains that "there need be no irrationality in intending to A and yet still not believing one will", but that, in contrast, "there will normally be irrationality in intending to A and believing one will not A" (1987, p. 38). According to Velleman, this view of Bratman's leaves much of his functional account of intentions unmotivated. In particular, it becomes unclear why in intending to A, an agent should be rationally required to identify means of *A*-ing or to rationally constrain her subsequent practical reasoning by ruling out options inconsistent with her *A*-ing, if she is agnostic whether she will in fact carry out her intention. Similarly, it becomes unclear why we should impose an agglomerativity requirement on intentions. As Velleman points out, it is unclear why intentions should be jointly executable if the agent can be agnostic as to whether they will be executed.

In my view, Velleman's third worry is exaggerated. Firstly, while Bratman indeed maintains that an intention to *A* does not require belief that one will *A*, he insists at the same time that an intention to *A* normally supports the belief that one will *A*. Secondly, Bratman also makes the point that agnosticism about whether one will act as intended does not directly undermine coherent planning but makes it more complex, leading us to form conditional intentions and plans for both failure and success to act as intended. Of course,

the viability of such a move depends on agnosticism being the exception rather than the rule; otherwise, we would have an unmanageable proliferation of conditional branching in our plans.

Velleman's first and second worries run deeper. If the only purposes of intentions are the pragmatic functions Bratman identifies, then there appear to be many instances where intentions don't serve these purposes or where serving them is actually counterproductive. This may be taken to indicate that Bratman's account is incomplete and that he has overlooked some of the functions intentions serve. This line of thought can be pursued in two different directions. On the one hand, we may try to identify further pragmatic functions that intentions, including present-directed intentions, could serve; on the other hand, we may look for non-pragmatic functions that intentions could serve. As we will now see, Velleman explores the second option, turning to Anscombe's theory of intentions in search of an answer. In contrast, what I will do myself later in this paper is explore the first option, giving it a social twist.

Velleman argues that Bratman's account of intentions misses an important function of intention, a function that is a central theme in Anscombe's theory of intention. In her book Intention (1963), she argues that intentions provide us with a special kind of self-knowledge and claims that this knowledge is special in two ways. It is knowledge of our own intentional actions, i.e., knowledge not just of what one is attempting to do, but of what one is actually doing, and it is knowledge without observation. Much philosophical ink has been spilled on how exactly these two claims should be interpreted. Following Falvey (2000), Velleman favors a reliabilist interpretation of these claims. According to this interpretation, knowledge of one's own intentional actions is non-observational because it is given by the content of our intentions and intentions in turn normally constitute (practical) knowledge of our own intentional actions because they reliably cause the facts that make them true. Note also, that on this reliabilist reading, Anscombe's claim is not that the content of our intentions provides us with infallible knowledge of what we are doing. To say that there normally exists a reliable connection between our intentions and actions is not to say that there cannot be cases when this connection does not obtain. However, as Velleman emphasizes, on Anscombe's account, failures of reliability undermine not just the epistemic status of intentions, they also undermine the intentionality of actions. If my intending to A does not reliably cause my A-ing, then, on the one hand, my intending to A will not amount to knowledge that I am A-ing and, on the other hand, my A-ing when it happens will be an accident rather than an intentional action. According to Anscombe, intentional actions are those "to which the question 'Why?' is given application" (1963, p. 9) and having practical knowledge is knowing a description of what one is doing, has done or is proposing to do that answers the question "Why?". Thus, the basic epistemic function of intentions is to provide us with a form of self-knowledge and self-understanding qua intentional agents.

According to Velleman, acknowledging this epistemic function of intentions does much to alleviate the worries raised by Bratman's practical account. With respect to the first worry – that present-directed intentions serve no purpose – one can now argue that while they might serve no practical purpose they still serve an epistemic function. With respect to the second worry – that on many occasions making one's mind in advance serves no pragmatic purpose –, one can now reply that in matters that are important to one's self-conception, uncertainty about one's future behavior is both uncomfortable and undesirable and that forming an intention allows us to gain self-knowledge and avoid this mental discomfort. With respect to the third worry – that absent a strong enough connection between intention and belief, it is unclear why intentions should be subject to the practical rationality requirements emphasized by Bratman–, Anscombe's theory regarding the epistemic function of intentions lets us see how the epistemic role of intentions could support their pragmatic functions.

The story as told so far suggests that we should think of the epistemic and pragmatic

functions of intentions as complementary. However, as Velleman points out, it still leaves us with two possible hypotheses or creation myths about the origin and ultimate purpose of intentions. On the pragmatic creation myth, the ultimate purpose of intentions would be pragmatic and their epistemic function would be subservient to their pragmatic functions, but may occasionally exemplify re-purposing: "That is, intention might have been designed to embody self-knowledge for the sake of facilitation coordination, but it might then be used on occasion, for the sake of self-knowledge alone, when coordination isn't necessary" (Velleman 2007, p. 208). By contrast, on the epistemic creation myth, the ultimate purpose of intentions may be to embody self-knowledge, and the pragmatic functions of intentions might have emerged simply as a fortuitous by-product of self-knowledge.

While Velleman has more sympathy for the epistemic than for the pragmatic creation myth, he thinks both should ultimately be rejected. In the next section, I'll consider his reasons for rejecting them, discuss the alternative story he proposes, and advance my own reasons for being skeptical about this story.

3 Velleman's spandrel

Despite their differences, the epistemic and the pragmatic creation myths rest on the common assumption that intentions have a teleofunction, some ultimate purpose they are designed to serve. Velleman thinks it is more plausible that their existence is an accident, that is to say, that they are the byproduct of some more general endowments of human nature. In other words, Velleman is tempted to think of the human will as, in Gould & Lewontin's phrase (1979), a spandrel, a feature formed not by design but as an accidental byproduct of some other designed feature or features. This leads him to be skeptical about both teleological myths. In telling his own creation myth, Velleman pursues two aims. His first aim is to show that the assumption behind the two teleological myths can be dispensed with. His second aim is to show that the accident that led to the emergence of the human will more closely approximates the epistemic than the pragmatic creation myth.

Velleman's own account of intentions characterizes them as an agent's commitment to the truth of some act-description of his or her forthcoming behavior that reliably causes this act-description to come true. He argues that this account of intentions "posits nothing more than the predictable consequences of two motivational states whose utility in the design of a creature is far more general than that of the human will" (Velleman 2007, p. 211). In other words, the human will is a spandrel, a feature arising from the accidental confluence of two designed features. What are these two features? The first, according to Velleman, is curiosity, defined as the creature's drive to understand what goes on in its environment. The second is self-awareness, through which the creature realizes that it is part of its environment and that its own behavior is part of what goes on in this environment. Self-awareness thus allows a creature to acquire an objective conception of itself. A creature that is both curious and self-aware will in turn be driven to understand its own behavior, that is, to understand "how the egocentrically conceived world of doing things is connected to the objectively conceived world of things understood" (Velleman 2007, p. 211). In understanding this, it will have acquired the capacity for intentions.

We can now see why Velleman thinks his own creation myth has more affinities with the epistemic than with the pragmatic creation myth. Curiosity is an epistemic drive and self-awareness is an epistemic capacity. As their byproduct, the capacity for intentions inherits this essential epistemic dimension. We can also understand why he means his own myth as an antidote to the methodological assumption inherent to the idea that intentions serve a specific teleofunction. Curiosity and self-awareness are, Velleman claims, designed for far more general purposes than that of the human will.

I think, however, that this is also where the creation myth told by Velleman reaches its limits. Important questions are left unanswered: What are these more general purposes served by curiosity and self-awareness? What good is curiosity?

What good is self-awareness? Unless he is willing to consider the will as a spandrel of spandrels, Velleman owes us answers to these questions. From an evolutionary point of view, it is unclear what benefits knowledge of their environment and knowledge of themselves could confer on creatures endowed with curiosity and self-awareness unless this knowledge found some behavioral expression. It isn't too difficult to see how a better understanding of their environment can promote more effective behavior, enhance the satisfaction of desires and needs, and ultimately have a differential impact on reproductive success in creatures endowed with curiosity. One should note, however, that pushing Velleman's story one step further in his direction has the effect of undermining his claim that his own myth has strong affinities with the epistemic creation myth for it suggests that the epistemic function of curiosity is ancillary to its pragmatic function, rather than the reverse.

It is less obvious how we should answer the question what good is self-awareness, what purposes it is designed for. My aim in the next two sections will be to remove two obstacles that prevent us from looking in the right direction for an answer to this question. The first obstacle lies in the fact that philosophers have tended to neglect an important pragmatic function of intentions. Thus, Velleman notes, rightly in my view, that Bratman's account of the pragmatic functions of intentions leaves many present-directed intentions without a purpose. However, rather than looking for some further pragmatic purpose intentions may serve, beyond scheduling deliberation and enhancing action coordination over time, Velleman turns his attention to epistemic functions. I will argue in section 4 that they both neglect a further important pragmatic function of intentions, namely their role in the online monitoring and control of action. The second obstacle lies in the fact that one central feature that makes us human, our deep sociality, is either ignored or at best a peripheral concern in philosophical accounts of intentions. Of course, I am not denying the obvious: many philosophers, and Bratman prominently among them, have explored joint agency and collective intentionality. Typically, however, their focus has been on whether or not joint agency should be seen as continuous with individual agency and thus on whether or not the conceptual framework developed to account for individual intentions could be fruitfully extended to shared intentions.[1] Rarely if ever, however, do they consider the possibility that shared intentions may shed light on some of the features and functions of individual intentions. In section 5, I will argue that the control and monitoring function of intentions plays a crucial role in contexts of joint action. I will further argue that this function might indeed be the primary function of intentions and might have become established because of the role it serves in solving the coordination problems that arise in joint action and because of the benefit thus conferred on creatures capable of solving these coordination problems.

4 Control: A further pragmatic function of intentions

Bratman (1987) considers future-directed intentions as the central case of intending to act and contrasts this approach to intention with an alternative approach that gives priority to immediate intentions or intentions in action. He notes that this second approach naturally leads to the idea that intentions in action reduce to complexes of beliefs and desires, i.e., that what makes it the case that an agent acts with a certain intentions are simply facts about the relation between the agent's actions and his beliefs and desires, and that this in turn tempts us into thinking that the same reductive strategy can be extended to future-directed intentions.[2] Focusing instead on future-directed intentions as the central case of intending allows us to identify functions of intentions that cannot easily be accommodated within a belief-desire model and thus makes the reductive strategy much less appealing. This would account for Bratman's emphasis on the deliberative and coordination functions of intentions. The flip side of the coin, however, is that present-directed intentions are then seen as little more than transmission belts needed to convey the motivational force of future-directed intentions. As noted by

1 See e.g., Bratman (2014) for a positive answer to these questions and Gilbert (1992, 2009) for a negative answer.

2 See for instance Davidson (1980, Essay 1) and Goldman (1970) for belief-desire reductive models of intentions.

Velleman, this leaves us with a potentially large class of actions where present-directed intentions appear to have no role to play, namely all these actions that are intentional yet not preceded by future-directed intentions. What belief-desire reductive approaches, Bratman's account and Velleman's account all seem to overlook is a specific pragmatic function of intentions in action or present-directed intentions, namely their role in the guidance, control and monitoring of action execution.

Harry Frankfurt (1978) was one of the first philosophers to criticize this oversight and insist on the importance of this pragmatic function of intentions. He emphasized that "a person must be in some particular relation to the movements of his body during the period of time in which he is presumed to be performing an action" (Frankfurt 1978, p. 157) and characterized this relation as one of guidance. Other philosophers have since shared his insight. For instance, Brand (1984), Bishop (1989) and Mele (1992) all insist that an adequate account of intentions should incorporate the guiding and monitoring roles of intentions in order to properly capture the close and continuous connection between intention and ongoing action.

The main reason why this connection between intention and ongoing action is needed is that human agents are neither infallible nor omniscient. Their expectations about the circumstances in which the action is to take place may not always be correct and they may fail to anticipate some of the relevant aspects of the situation of action. In other words, their situational beliefs may be incorrect or incomplete. The same goes for their instrumental beliefs. Suppose, for instance, that I intend to visit a colleague in her office. I may be wrong in thinking that this is the door to her office (incorrect situational belief) or unsure which door is her office door (lack of relevant situational belief). Similarly, I may also be wrong in thinking that I should pull the door to open it (incorrect instrumental belief) or unsure whether to push or pull (lack of relevant instrumental belief). If intentions are to reliably produce behavior matching their representational content (e.g., visiting my colleague in her office), they should have some flexibility and incorporate monitoring processes to detect deviations that jeopardize the success of the action and correction processes to trigger compensatory activity.

This emphasis on control finds a strong echo in the literature on motor cognition (see, e.g., Jeannerod 1997, 2006). Indeed, it is in this literature that we can find the most precise characterization of the monitoring and control functions of intentions and of the mechanisms that support them. According to the very influential internal model theory of motor control, motor control strategies are based on the coupling of two types of internal models: inverse models and forward models (Frith et al. 2000; Jordan & Wolpert 1999; Wolpert 1997). Inverse models compute the motor commands needed for achieving a desired state given the current state of the system and of the environment. An efference copy of these commands is fed to forward models, whose role is to make predictions about the consequences of the execution of these commands. The control of action is thought to depend on the coupling of inverse and forward models through a series of comparators: error signals arising from the comparison of desired, predicted, and actual states (monitoring) are used for various kinds of regulation (control). In particular, they can be used to correct and adjust the ongoing action in the face of perturbations, as well as to update both inverse and forward models to improve their future functioning.

Recent experimental work in motor cognition also suggests, however, that much of action control is automatic and proceeds independently of conscious awareness. For instance, in an experiment (Castiello et al. 1991) participants were asked to reach for and grasp a target as quickly as possible and their hand trajectories were recorded. On some trials, though, the target shifted position after the movement had started. When this happened, participants were instructed to correct their movement in order to reach accurately for the target and to signal the time at which they became aware of its displacement by shouting "Tah!". The experiment showed that the participants started correcting

their movements more than 300ms before they signaled awareness of the target displacement. A subsequent study (Pisella et al. 2000) was especially instructive. In a first experiment they used a similar paradigm but introduced a condition where participants were requested to interrupt their movement when the target changed location. Despite the instruction, the participants could not prevent themselves from correcting their movements instead of stopping for a good 200 ms. In contrast, however, in a second experiment green and red targets were presented simultaneously in the two positions and the participants' task was to point at the green one. On some trials, the color of the two targets could be unexpectedly interchanged at movement onset. When this happened, one group of participants was instructed to interrupt their ongoing movement and the other group to correct it. In contrast to what happened in the first experiment, no automatic corrective movements were observed in the group instructed to interrupt their movement and in the other group corrections involved a significant increase in movement time. Thus, these results suggest that while corrections made in response to spatial perturbations are under automatic control, corrections in response to chromatic perturbations require intentional control.

On the one hand, the mere fact that some or much of action control can be automatic is not a sufficient reason to deny a control function to intentions. The experimental studies presented in the previous paragraph suggest that action control can indeed operate automatically and outside of conscious awareness and that when there is a conflict between automatic and intentional control, automatic control may take precedence over intentional control. Yet, they also provide evidence that some corrections cannot be carried out automatically but depend on intentional control. On the other hand, the mere fact that intentional control seems needed to compensate for chromatic perturbations may not provide sufficient ground for considering that the intentional control of action execution is a central function of intentions. One would want a more systematic account of the respective roles of automatic and intentional control. Recent developments of the internal model approach to motor control may constitute a useful guide.

While the internal model approach to motor control was initially introduced to account for fine-grained aspects of motor control, more recent versions of this approach emphasize the hierarchical nature of motor control (Hamilton & Grafton 2007; Jeannerod 1997; Kilner et al. 2007). They propose that internal inverse and forward models are arranged in a hierarchy and that error signals generated at one level of the hierarchy can propagate to the next level when correction mechanisms at this level are not able to make the necessary compensations. I have suggested elsewhere (Pacherie 2008) that one can distinguish three broad levels in an action specification hierarchy. At the highest level, action representations represent the whole action as a unit, in terms of its overarching goal and of the sequence of steps or subgoals needed to achieve that goal. At this level, the action may still be represented in a rather abstract format. The second level is concerned with the implementation of each step in the action plan and involves selecting an appropriate motor program given the immediate goal and contextual information about the current state of the agent and the current state of its environment. In other words, processes at this level are in charge of anchoring the successive steps of the action plan in the current situation and of selecting appropriate motor programs. Finally, once a motor program has been selected, the exact values of its parameters must still be set. This is done at the third level, where incoming sensory information about external constraints is used to specify these values.

Acknowledging the existence of different levels of action control corresponding to these different levels in the action specification hierarchy may allow us to accommodate both automatic and intentional action control processes. As long as error signals can be reduced by automatic corrections made at lower levels in the hierarchy, there is no need for the intervention of intentional control. However, there are two classes of cases where automatic corrections may not be sufficient to put an action back on

track. First, important external perturbations can lead to discrepancies that are too large to be automatically compensated. In such a case, error signals would propagate upwards, we would become aware of them and shift to a conscious, intentional compensation strategy. Second, in some instances there may also be discrepancies in the ways the action is or can be specified at different levels of the action representation hierarchy (inter-level representational misalignment). Thus, the study by Pisella and colleagues (Pisella et al. 2000) suggests that action specification at the sensorimotor level does not encode chromatic information and uses spatial information as a proxy for it. When chromatic information and spatial information vary independently, as they do in one of the conditions of the experiment, representations at different levels of the action representation hierarchy become misaligned and the intervention of conscious control becomes necessary to realign them.

Importantly, on this conception of intentional control and as Frankfurt had already noted, what is essential for actions to be intentionally controlled is not that intentional control processes actually affect their course, but that these control mechanisms would have intervened to adjust the action had the need arisen. In other words, an action may be intentionally controlled even though automatic rather than voluntary control mechanisms intervene to compensate for deviations, provided these voluntary control mechanisms would have kicked in, had automatic corrections proved insufficient.

Even more importantly, if action control is an essential function of intentions, then we should stop thinking of intentions as simply mental representations of goals somehow triggering motor processes that, if everything goes well, will yield the desired outcome. Rather, we should think of monitoring and control processes as intrinsic to intentions, that is, of intentions as encompassing not just representations of goals but also a specific set of monitoring and control processes organizing and structuring the motor processes that themselves generate movements.

In this section, I argued for the idea that the control of action execution is an important pragmatic function of intentions. Acknowledging the existence and importance of this function allows us to plug gaps in the creation myths considered earlier. First, it allows us to attribute a specific pragmatic function to present-directed intentions rather than considering them as mere transmission belts in charge of conveying the motivational force of future-directed intentions. We can thus assuage one of the main worries raised by Velleman against Bratman's pragmatic account of intentions and the pragmatic creation myth derived from it. Second, Anscombe's and Velleman's accounts of intentions both assume that intentions reliably cause behavior that matches their representational content. Human agents, however, are neither infallible nor omniscient. Their situational and instrumental beliefs can be incorrect or they can lack situational and instrumental beliefs that are relevant to the successful execution of their intentions. Thus, the reliability demanded by Anscombe's and Velleman's accounts largely depends on our having powerful and flexible control processes allowing us to put our actions back on track when perturbations deviate their course.

One may agree that the conscious control of individual action is a function of intention in the sense that intentions have this causal role, but still be skeptical that this is the role intentions are designed for, or to put it in other words, that it is a teleofunction of intentions. Thus, one could argue that very large external perturbations are rare and that inter-level representational misalignment is the exception rather than the rule. If so, most of action control would be automatic anyway and intentional action control would play at best a marginal role. It would therefore be unlikely to confer on intention-forming creatures benefits important enough to warrant the claim that intentions are designed for action control. As I have tried to argue in this section, the benefits conferred by online conscious control over actions are not as negligible as this deflationary view implies. In addition, I think we can build a very strong case that conscious action control confers im-

portant benefits if we consider joint activities rather than just individual actions. Acting jointly demands that we solve coordination problems that do not arise (or arise only in a very attenuated form) in individual action. In what follows, I will argue that online conscious control plays a crucial role in solving these coordination problems. I will further speculate that conscious online control over actions might indeed have become established as the primary function of intentions because of the role it served in solving these coordination problems and because of the benefit this conferred on creatures capable of solving these coordination problems and thus of acting jointly in an efficient and flexible way.

5 The social creation myth

Humans have been characterized as the ultra-cooperative species (Tomasello 2009, 2011). This ultra-cooperativeness has made us one of the most successful species on earth, spreading all over the planet, creating and developing cultural artifacts and practices that are themselves culturally transmitted and accumulate over time, thus giving us a further competitive edge over other species. According to Tomasello, underlying humans' ultra-cooperativeness are a set of species-unique skills and motivations for shared intentionality, involving "such things as the ability and motivation to form shared goals and intentions with others in collaborative activities, and the ability and motivation to share experience with others via joint attention, cooperative communication, and teaching" (2011, p. 6).

The gist of the social creation myth I am proposing in this section is that the main benefits associated with intentions and with the kind of control over actions they make possible arise in social cooperative contexts where agents have to coordinate their actions to achieve a shared goal. I start with an examination of the special demands for coordination acting jointly with others creates. I then explain how the capacity to form conscious intentions is a crucial component of our ability to meet these demands.

Successful joint action depends on the efficient coordination of participant agents' goals, intentions, plans, and actions. As I argued elsewhere (Pacherie 2012), it is not enough that agents control their own actions, i.e., correctly predict their effects, monitor their execution and make adjustments if needed. They must also coordinate their actions with those of their co-agents so as to achieve their joint goal. For that they must monitor their partner's intentions and actions, predict their expected consequences and use these predictions to adjust what they are doing to what their partners are doing. The implication of these processes, however, is not unique to joint action nor enough to promote their success. In competitive contexts they also play an important role. For instance, in a fight being able to anticipate your opponent's moves and to act accordingly is also crucial. What is furthermore required in the case of joint action is that co-agents share a goal and understand the combined impact of their respective intentions and actions on their joint goal and adjust them accordingly. In competitive contexts, an agent should typically aim at predicting his opponents' moves, while at the same time endeavoring to make his own moves unpredictable to his opponents. In contrast, in cooperative contexts mutual predictability must be achieved for efficient coordination towards a shared goal to be possible. Agents should be able to align their representations of what themselves and their partners are doing and of how these actions together contribute to the shared goal.

Various forms of uncertainty can undermine mutual predictability, the alignment of representations and hence coordination. They can be organized into three broad categories. The first category involves motivational uncertainty: we can be unsure how convergent a potential partner's interests are with our own interests and thus unsure whether there are goals we share and can promote together. The second category involves instrumental uncertainty: even assuming that we share a goal, we can be unsure what we should do to achieve that goal, or, if we have a plan, unsure how roles should be distributed among us, or, yet, unsure when and where we should act. The third category involves common ground uncertainty: we can be

unsure how much of what is relevant to our deciding on a joint goal, planning for that goal and executing our plan is common ground or mutually manifest to us.

Philosophical accounts of joint agency, including Bratman's (2009, 2014) do not ignore these challenges but they are essentially concerned with high-level processes involved in making decisions about whether or not to act together and in advance planning. Their focus is on the coordination of agent's intentions prior to acting and they pay little heed to the processes enabling people to coordinate during action execution. In contrast, in the last decade, cognitive scientists have investigated joint action by focusing on lower-level online coordination processes in relatively simple joint tasks and on the factors that affect these coordination processes. In what follows, I will argue that there are important limitations to what these advance and online coordination processes can achieve and that high-level online intentional control is crucial to overcoming these limitations. First, however, let us consider the main characteristics of the two sets of coordination processes philosophers and psychologists typically focus on.

Bratman's account of shared intentions is a good illustration of the way philosophical accounts approach coordination issues in joint action. In addition, its explicitness makes it possible to see clearly what advance coordination involves and how it is achieved.

Bratman (2009) proposes that shared intention involves the following conditions as its main building blocks:

1. Intentions on the part of each in favor of the joint activity.
2. Interlocking intentions: each intends that the joint activity go in part by way of the relevant intentions of each of the participants.
3. Intentions in favor of meshing subplans: each intends that the joint activity proceed by way of subplans of the participants that are co-realizable and can be consistently agglomerated.
4. Disposition to help if needed: given that the contribution of the other participants to the joint activity is part of what each intends, and given the demands of means-end coherence and of consistency that apply to intentions, each is under rational pressure to help others fulfill their role if needed.
5. Interdependence in the persistence of each participant's relevant intention: each believes that the persistence of the other participants' intention in favor of the joint activity depends on the persistence of his own and vice-versa.
6. Joint-action-tracking mutual responsiveness: each is responsive to each in relevant subsidiary intentions and in relevant actions in a way that tracks the joint action.
7. Common knowledge among all participants of all these conditions.

Let me offer some comments on these conditions. First, Bratman offers these conditions as a set of sufficient conditions for a shared intention, leaving it open that shared intentions may be realized in other ways, in particular in cases of joint activities involving institutions. Second, conditions (1), (2) and (5) are meant to deal with motivational uncertainty. Bratman points out that the concept of a joint activity that figures in the contents of the intentions in (1) should be understood in a way that is neutral with respect to shared intentionality. So condition (1) only insures that agents share goals in a weak sense of the notion. Rather it is condition (2) that is in charge of insuring that the motivational states of the agents align in the way required for joint cooperative activity: it is the fact that for each participant, the content of their intention refers to the role of the intentions of other participants that, for Bratman, captures the intentional jointness of their actions. Condition (5) in turn specifies how these motivations stay aligned. Third, conditions (3), (4) and (6) relate to means-end uncertainty and are meant to reduce it. According to Bratman, they can be derived from condition (2) taken together with the norms of practical rationality that already govern individual planning and acting. Bratman's key idea is that the interlocking and interdependent intentions of individual participants, in responding to the norms of

practical rationality governing individual planning agency, will also respond to the norms of social agglomeration and consistency, social coherence and social stability shared intentions are subject to. This would involve, in Bratman's terms, commitments to mutual compatibility of relevant sub-plans, commitments to mutual support, and joint-action tracking mutual responsiveness. Finally, the function of condition (7) is, rather obviously, to reduce common ground uncertainty.

Bratman's basic idea is thus that this structure of interlocking and interdependent intentions, when it functions properly, frames relevant bargaining and shared deliberation and thus supports and guides coordinated planning and action in pursuit of the intended shared activity. Unsurprisingly, since Bratman's theory of joint agency is continuous with his planning theory of individual intentions, it is in virtue of the pragmatic functions intentions already serve in the individual action case that the interlocking and interdependent intentions of individual participants can also support coordination in the joint action case.

While Bratman, in his condition (6), stipulates that agents should be mutually responsive not just in their relevant intentions and subsidiary intentions but also in relevant actions in a way that tracks the joint action, his account doesn't tell us by what means mutual responsiveness in action is achieved. To know more about this, we have to turn our attention to recent psychological work on joint agency. In contrast to philosophical approaches, cognitive psychology studies of joint action typically focus on the perceptual, cognitive, and motor processes that enable individuals to coordinate their actions with others online.

Knoblich and colleagues (Knoblich et al. 2011) distinguish between two broad categories of coordination processes, emergent and planned. In emergent coordination, coordinated behavior occurs due to perception-action couplings that make multiple individuals act in similar ways. One source of emergent coordination is entrainment, the process of synchronizing two or more actors' rhythmic behaviors with respect to phase (e.g., Richardson et al. 2007). A second source of emergent coordination is perception-action matching, whereby observed actions are matched onto the observer's own action repertoire and can induce the same action tendencies in different agents who observe one another's actions (Jeannerod 1999; Prinz 1997; Rizzolatti & Sinigaglia 2010; Knoblich & Sebanz 2008). Importantly, emergent forms of coordination are independent of any joint plans or common knowledge, which may be altogether absent. They support basic forms of motor and representational alignment that can facilitate mutual responsiveness in action, but they do not ensure that the agent's actions track a joint goal. Indeed, the successful performance of some joint actions may require that these automatic coordination processes be inhibited. For instance, the performance of composer Steve Reich's famous piece, *Drumming*, based on the technique of phasing, requires the musicians to play the same rhythmic pattern out of sync.

In planned coordination, agents plan their own actions in relation to the joint goal and also to some extent to their partners' actions. As emphasized by Knoblich et al. (2011), shared task representations play an important role in planned coordination. Shared task representations do not only specify in advance what the respective tasks of each of the co-agents are, they also provide control structures that allow agents to monitor and predict what their partners are doing, thus enabling interpersonal coordination in real time. Empirical evidence shows that having shared task representations influences perceptual information processing, action monitoring, control and prediction during the ensuing interaction (Heed et al. 2010; Schuch & Tipper 2007; Sebanz et al. 2006; Tsai et al. 2006). Furthermore, several studies (Sebanz et al. 2005; Sebanz et al. 2006) have shown that actors may form shared representations of tasks quasi-automatically, even when it is more effective to ignore one another.

Several researchers have also suggested that joint attention provides a basic mechanism for sharing representations of objects and events and thus for creating a perceptual common ground in joint action (Tomasello & Carpenter 2007; Tollefsen 2005). To act jointly, it is often

necessary not only that the co-agents identify the objects to be acted upon, their location as well as the location of possible obstacles, but also be mutually aware that they do. Joint attention may thus play an important role in ensuring that co-agents track the same objects and features of the situation and be mutually aware that they do. In a recent study, Böckler et al. (2011) showed that attending to objects together from opposite perspectives makes people adopt an allocentric rather than the default egocentric frame of reference. These authors suggest that taking an allocentric reference may support the efficiency of joint actions from different spatial orientations. Independently of mutual manifestness, being able to assess what others are perceiving, or can or cannot perceive at a given moment in time may also facilitate coordination. For instance, a study by Brennan and colleagues (Brennan et al. 2007) demonstrated that co-agents in joint visual search space were able to distribute a common space between them by directing their attention depending on where the other was looking and that their joint search performance was thus much more efficient than their performance in an individual version of the search task.

There are, however, important limitations to what these emergent and planned online coordination processes can achieve. First, to the extent that they exploit perceptual information, they can be of no help unless a certain amount of common perceptual information is indeed available to co-agents. Second, even when common perceptual information is available, there are limits to our processing capacities. An agent may be able to simultaneously track what a small number of other agents are currently doing or attending to, but when the number of agents increases, this capacity soon finds its limits. Our capacity to co-represent the actions, goals, and intentions of other agents we observe acting encounters similar limitations. Understanding of actions through motor resonance and mirroring works only to the extent that the observed actions are part of the action repertoire of the observer. Similarly, when actions are relatively novel, agents may not yet have formed sufficiently detailed shared task representations. Finally, unexpected effects of action execution or failures of coordination may reveal various forms of misalignment between partners' representations or indicate that their representations, though aligned, were inaccurate.

When pre-alignment is insufficient or breakdowns occur due to misalignment in the action execution phase, the deliberate and conscious production of social signals aimed at aligning or realigning relevant representations becomes crucial. Agents cannot count on alignment arising spontaneously. They have to make it happen. Intentional communication, whether verbal or not, is then needed to make it happen.

As emphasized by Herbert Clark (2006), joint activities can typically be partitioned into two sets of actions: a basic joint activity and coordinating joint actions. The basic joint activity comprises all the actions essential to achieving the basic joint goal, while the coordinating joint actions consists in the set of communicative acts about the basic activities that insure relevant representational alignment. To study this partitioning of joint activities, Clark and his co-workers ushered two people in a small room, gave them the parts of a kit for a TV stand and asked them to assemble the stand, videotaping them and recording their verbal exchanges while putting they did it. Here's a short extract of their exchanges, taken from Clark (2006, p. 128):

> Ann Should we put this in, this, this little like kinda cross bar, like the T? like the I bar?
>
> Burton Yeah ((we can do that))
>
> Ann So, you wanna stick the ((screws in)). Or wait is, is, are these these things, or?
>
> Burton That's these things I bet. Because there's no screws.
>
> Ann Yeah, you're right. Yeah, probably. If they'll stay in.
>
> Burton I don't know how they'll stay in ((but))

Ann Right there.

Burton Is this one big enough?

Ann Oh ((xxx)) I guess cause like there's no other side for it to come out.

Burton M-hm.

[8.15 sec]

Burton ((Now let's do this one))

Ann Okay

First, it should be noted that, as often happens in daily life, this joint activity was not planned in advance. Instead, Ann and Burton discover that they have to assemble a TV stand and work out together what they should do as they go along. Second, Clark points out that Ann and Burton's coordinating joint actions are structured in what he calls projective pairs, comprising a proposal and an uptake (i.e., full acceptance, altered acceptance or rejection of proposal). Third, the exchanges can be gestural as well as verbal. For instance, instead of, or concomitantly with, asking verbally whether Burton is ready to fasten the screws, Ann may present him with the screwdriver and his taking it count as acceptance. Fourth, the contents of these exchanges show that they are aimed at reducing instrumental uncertainty. Typically, they are about what should be done and how, who should do what, and when and where it should be done. When the task presents difficulties, they may also serve to reduce motivational uncertainty. For instance, Burton might ask whether they should give it a last try and Ann either acquiesce or reject the proposal. Finally, the structure of the projective pairs shows that at the same time they aim at reducing common ground uncertainty. Proposals are about potential alignments and full acceptance confirms alignment and common ground. Tellingly, with altered acceptance uptakes, projective pairs evolve into projective triads, the third element of the exchange being the proposer's uptake on the alteration.

Importantly, to negotiate and achieve alignment in this way, we must be aware of our own and others' intentions and beliefs and this at two levels, corresponding to the two sides of the partitioning characterized by Clark. On the one hand, it is essential to the fulfillment of communicative intentions that they be recognized as such by the addressee (Grice 1957; Recanati 1986; Sperber & Wilson 1986). On the other hand, what agents communicate in these contexts is information about their beliefs and intentions regarding the joint action. This suggests that the development of self-consciousness and consciousness of other minds, of intentional communication, and of increasingly complex forms of coordinated joint action go hand in hand.

The success of both individual and joint action depends on representational alignment. In the case of individual action, representation alignment takes two main forms. First, at a given level of action specification, a match should be achieved between representations of desired, predicted and actual states. We can call this first form of alignment intra-level alignment. Second, inter-level alignment is also necessary; that is, despite differences in representational format and resources, action specifications at different levels of the action representation hierarchy should be kept aligned. Conscious online control may be needed to restore alignment when severe intra- or inter-level discrepancies occur. However, it may be argued that in the individual case alignments are taking place within a single cognitive system and that this system is normally sufficiently integrated or unified that serious misalignments are rare and thus that the need for online conscious control is limited.

The main difference between individual and joint actions lies in the coordination demands essential to joint action. Thus, a third form of representational alignment becomes crucial in joint action. In addition to individual intra-and inter-level representational alignment, inter-agent representational alignment is necessary to meet coordination demands. Inter-agent alignment may be achieved in part through advanced planning, as proposed by Bratman. It

can also be achieved in part through online emergent and planned coordination processes of the types explored and described in the recent psychological literature. However, there are important limitations to what these coordination processes can achieve. Advance planning, when it takes place, may help define a shared background framework for the joint action, but at this stage it is typically impossible to anticipate all the coordination demands that will arise at the execution stage. Some of these demands may be met by the kinds of online coordination processes reviewed earlier in this section, but, as I pointed out, there are also important limitations to what they can achieve. In many instances, the progress of a joint action is hindered or the action breaks down due to various forms of misalignment between agents' representations. In such instances, individual corrections do not suffice to put the joint action back on track. Rather, to overcome these failures, agents need to align or realign their representations. This process calls for what Clark calls coordinating joint actions, that is, communicative acts about the basic joint activity. These communicative acts in turn are intentional and aim at communicating information about the agents' intentions and beliefs with a view to achieve alignment. But one can only communicate intentionally about one's beliefs and intentions if one is aware of them. Conversely, one can only understand the communicative acts of other agents if one realizes that these agents have a capacity for intentions. Finally and crucially, as already emphasized by Velleman (2007) in his discussion of Bratman's account, intentions could not serve their pragmatic functions unless they also had an epistemic role. In other words, if my having the intention to *A* didn't count as a form of practical self-knowledge and didn't give me grounds to believe that I would act as intended, my communicating (sincerely) about my intention to *A* would not license other agents to form beliefs about my future actions and thus would not yield the kind of inter-agent representational alignment needed to achieve coordination.

To recap, joint actions create more comprehensive demands for representational alignment than individual actions, since their success depends not just on individual intra- and inter-level representational alignment but also on inter-agent representational alignment. New resources are needed to meet these demands. On the social creation myth proposed here, a capacity for conscious intentions is crucial to inter-agent representational alignment. Having conscious intentions allows us to communicate about them and engage in coordinating joint actions that create common ground and promote the success of basic joint activity. The answer this myth offers to the question what is the purpose of conscious intentions is then that it is to enable more efficient inter-personal coordination in joint action and thus reap the benefits that come with increasingly complex and flexible forms of coordinated actions. The social creation myth doesn't deny intentions an epistemic role. On the contrary, it acknowledges that intentions couldn't serve their inter-personal coordination function if they did not at the same time provide us with a form of self-knowledge. However, it views their epistemic function as subservient to their coordination function. The social creation myth does not deny either that conscious intentions play a role in the online control of individual action. Rather, it proposes that conscious control of individual action may be a by-product of a capacity for conscious control that became established in social contexts because of the role it served in solving inter-agent coordination problems and because of the benefit conferred by the forms of cooperation it made possible.

6 Conclusion: Relating creation myths

The Bratmanian creation myth is pragmatic but also diachronic and individualist. Intentions have a purpose or teleofunction. This function is pragmatic insofar as the main benefit attached to intentions is to allow us to secure greater desire satisfaction. The way intentions secure this benefit is by allowing us to organize and coordinate our actions diachronically, in other words to become planning agents. As noted by Velleman, this emphasis of diachronicity and future-directed intentions leaves present-directed

intentions without a clear function. Finally, this myth is to a large extent individualist. While planning agency also enables inter-individual coordination, the social dimension of intentions remains secondary in Bratman's account and again his main concern is with diachronically organized joint actions.

While the social creation myth also sees intentions as having a pragmatic purpose, in contrast to the Bratmanian myth, it emphasizes the social and synchronous dimension of intentions. Instead of self-coordination over time, it emphasizes cooperation and flexibly coordinated joint action as the main route to greater desire satisfaction. It thus reverses the Bratmanian perspective in proposing that intentions are designed to enable a more efficient online coordination of joint action and in considering future-directed individual or joint planning as derivative or secondary functions of intentions.

Because its main emphasis is on synchronicity rather than diachronicity, the social creation myth has no problem attributing a pragmatic control function to present-directed intentions. It is thus impervious to one of the attractions of the Anscombian creation myth. We need feel no temptation to attribute an epistemic function to present-directed intentions for lack of any other plausible option. The social creation myth, however, does not dispense with epistemic functions altogether, quite the reverse. Not only is the fact that intentions embody a form of self-knowledge essential to their role in the coordination of joint actions, but in addition the way intentions play their coordinative role is by contributing to the alignment of representations with co-agents and thus to the production of shared knowledge. Thus, on the social creation myth, the epistemic function of intentions is not just to provide us with self-knowledge about our intentions and actions, it is also to contribute to the formation of shared knowledge. However, the social creation myth remains closer to the pragmatic than to the epistemic creation myth in considering that the epistemic function of intentions is ancillary to its pragmatic purpose.

Finally, is the social creation myth a teleological myth or, like Velleman's myth, the story of a spandrel? I must admit that I am not sure what the answer to this question is or should be. Indeed, this was one of the reasons why I chose to call my story a creation myth. One thing is sure though, if it is a story about a spandrel, this spandrel is not the same as Velleman's. His spandrel is a by-product of curiosity and self-awareness. This spandrel, if it is one, would involve a third element, sociality or cooperativeness. Social theories of consciousness (Frith 2010; Graziano & Kastner 2011) propose that consciousness has evolved to facilitate social interactions and enhance social cooperation. On the one hand, a capacity for consciousness is of course a much more general capacity than a capacity for conscious intentions and this may suggest that the latter, as a by-product of this more general capacity, is itself merely a spandrel. On the other hand, if the ultimate purpose of consciousness is to enhance social cooperation, then conscious intentions are a key element in making this possible and calling our capacity for intention a spandrel would fail to do justice to their role.

References

Anscombe, G. E. M. (1963). *Intention.* Oxford, UK: Blackwell.

Bishop, J. C. (1989). *Natural agency: An essay on the causal theory of action.* Cambridge, UK: Cambridge University Press.

Brand, M. (1984). *Intending and acting: Toward a naturalized action theory.* Cambridge, MA: MIT Press.

Bratman, M. (1987). *Intention, plans, and practical reason.* Cambridge, MA: Harvard University Press.

——— (2009). Modest sociality and the distinctiveness of intention. *Philosophical Studies*, *144* (1), 149-165. 10.1007/s11098-009-9375-9

——— (2014). *Shared agency.* Oxford, UK: Oxford University Press.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M. & Zelinsky, G. (2007). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106* (3), 1465-1477. 10.1016/j.cognition.2007.05.012

Böckler, A., Knoblich, G. & Sebanz, N. (2011). Giving a helping hand: Effects of joint attention on mental rotation of body parts. *Experimental Brain Research*, *211* (3-4), 531-545. 10.1007/s00221-011-2625-z

Castiello, U., Paulignan, Y. & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness a study in normal subjects. *Brain*, *114* (6), 2639-2655. 10.1093/brain/114.6.2639

Clark, H. H. (2006). Social actions, social commitments. In N. J. Enfield & S. C. Levinson (Eds.) *Roots of human sociality: Culture, cognition, and interaction* (pp. 126-150). Oxford, UK: Berg.

Davidson, D. (1980). *Essays on Actions and Events.* Oxford, UK: Oxford University Press.

Falvey, K. (2000). Knowledge in intention. *Philosophical Studies*, *99* (1), 21-44. 10.1023/a:1018775307559

Frankfurt, H. (1978). The problem of action. *American Philosophical Quarterly*, *15* (2), 157-162.

Frith, C. (2010). What is consciousness for? *Pragmatics & Cognition*, *18* (3), 497-551. 10.1075/pc.18.3.03fri

Frith, C. D., Blakemore, S.-J. & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London*, *355* (1404), 1771-1788. 10.1098/rstb.2000.0734

Gilbert, M. (1992). *On social facts.* Princeton, NJ: Princeton University Press.

——— (2009). Shared intention and personal intentions. *Philosophical Studies*, *144* (1), 167-187. 10.1007/s11098-009-9372-z

Goldman, A. (1970). *A theory of human action.* Englewood Cliffs, NJ: Prentice-Hall.

Gould, S. J. & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society B*, *205* (1161), 581-598. 10.1098/rspb.1979.0086

Graziano, M. S. & Kastner, S. (2011). Human consciousness and its relationship to social neuroscience: A novel hypothesis. *Cognitive Neuroscience*, *2* (2), 98-113. 10.1080/17588928.2011.565121

Grice, H. P. (1957). Meaning. *Philosophical Review*, *66* (3), 377-388. 10.2307/2182440

Hamilton, A. F. & Grafton, S. T. (2007). The motor hierarchy: From kinematics to goals and intentions. In P. Haggard, Y. Rossetti & M. Kawato (Eds.) *Sensorimotor foundations of higher cognition* (pp. 381-408). Oxford, UK: Oxford University Press.

Heed, T., Habets, B., Sebanz, N. & Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Current Biology*, *20* (15), 1345-1349. 10.1016/j.cub.2010.05.068

Jeannerod, M. (1997). *The cognitive neuroscience of action.* Oxford, UK: Blackwell.

——— (1999). The 25th Bartlett Lecture. To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology*, *52* (3), 1-29. 10.1080/713755803

——— (2006). *Motor cognition.* Oxford, UK: Oxford University Press.

Jordan, M. I. & Wolpert, D. M. (1999). Computational motor control. *The cognitive neurosciences* (pp. 485-493). Cambridge, MA: MIT Press.

Kilner, J. M., Friston, K. J. & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8* (3), 159-166. 10.1007/s10339-007-0170-2

Knoblich, G., Butterfill, S. & Sebanz, N. (2011). Psychological research on joint action: Theory and data. *Psychology of Learning and Motivation - Advances in Research and Theory*, *54*, 59-101. 10.1016/B978-0-12-385527-5.00003-6

Knoblich, G. & Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of the Royal Society B*, *363* (1499), 2021-2031. 10.1098/rstb.2008.0006

Mele, A. R. (1992). *Springs of action: Understanding intentional behavior.* Oxford, UK: Oxford University Press.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107* (1), 179-217. 10.1016/j.cognition.2007.09.003

——— (2012). The phenomenology of joint action: Self-agency vs. joint-agency. In A. Seemann (Ed.) *Joint attention: New developments* (pp. 343-389). Cambridge MA: MIT Press.

Pisella, L., Grea, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G. & Rossetti, Y. (2000). An 'automatic pilot' for the hand in human posterior parietal cortex: Toward reinterpreting optic ataxia. *Nature Neuroscience*, *3* (7), 729-736. 10.3389/fnhum.2013.00336

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9* (2), 129-154. 10.1080/713752551

Recanati, F. (1986). On defining communicative Intentions. *Mind and Language*, *1* (3), 213-242. 10.1111/j.1468-0017.1986.tb00102.x

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L. & Schmidt, R. C. (2007). Rocking together: Dynamics of unintentional and intentional interpersonal coordination. *Human Movement Science*, *26* (6), 867-891. 10.1016/j.humov.2007.07.002

Rizzolatti, G. & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11* (4), 264-274. 10.1038/nrn2805

Schuch, S. & Tipper, S. P. (2007). On observing another person's actions: Influences of observed inhibition and errors. *Perception & Psychophysics*, *69* (5), 828-837. 10.3758/BF03193782

Sebanz, N., Knoblich, G. & Prinz, W. (2005). How two share a task: Corepresenting stimulus–response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, *31* (6), 1234-1246. 10.1037/0096-1523.31.6.1234

Sebanz, N., Knoblich, G., Prinz, W. & Wascher, E. (2006). Twin Peaks: An ERP study of action planning and control in co-acting individuals. *Journal of Cognitive Neuroscience*, *18* (5), 859-870. 10.1162/jocn.2006.18.5.859

Sperber, D. & Wilson, D. (1986). *Relevance: Communication and cognition.* Oxford, UK: Blackwell.

Tollefsen, D. (2005). Let's pretend: Children and joint action. *Philosophy of the Social Sciences*, *35* (75), 74-97. 10.1177/0048393104271925

Tomasello, M. (2009). *Why we cooperate.* Cambridge, MA: MIT Press.

——— (2011). Human culture in evolutionary perspective. In M. Gelfand (Ed.) *Advances in Culture and Psychology* (pp. 5-51). Oxford, UK: Oxford University Press.

Tomasello, M. & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10* (1), 121-125. 10.1111/j.1467-7687.2007.00573.x

Tsai, C.-C., Kuo, W.-J., Jing, J.-T., Hung, D. L. & Tzeng, O. J.-L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research*, *175* (2), 353-362. 10.1007/s00221-006-0557-9

Velleman, D. (2007). What good is a will? In A. Leist & H. Baumann (Eds.) *Action in context* (pp. 193-215). Berlin, GER: de Gruyter.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1* (6), 209-216. 10.1016/S1364-6613(97)01070-X

Conscious Intentions: Do We Need a Creation Myth?

A Commentary on Elisabeth Pacherie

Andrea R. Dreßing

We experience ourselves as agents, performing goal-directed actions in the world. In her paper about *Conscious Intentions: The social creation myth* Pacherie develops a creation myth about the function of conscious intentions, based on her hierarchical concept of individual motor actions and joint action. In this creation myth, conscious intentions are not understood as internal mental states with a teleofunctional role. Having a conscious intention exerts a specific contribution to motor control and conscious intentions might have a potential causal power in this myth.

In this commentary I want to postulate, that Pacherie's social creation myth is more than a myth but rather the search for an explanation of the function of conscious intentions in the physical world. It tries to explain the feature of the intention *being conscious* that endows it with its particular causal function. Yet — speaking about a causal function — the potential analytical and neuroscientific limitations of a causal function of conscious intentions in the social creation myth have to be analysed with regard to the argument of causal closure and results of experimental approaches to the causal relevance of conscious intentions. I argue that despite these limitations the social creation myth could be an important step on the way of finding an explanation about the function of conscious intentions, if the question about the *function* of conscious intentions is slightly adjusted and is not understood in a strictly causal way.

Keywords
Causal closure | Conscious agents | Conscious intention | Creation myth | Intentional action | Joint action | Mental causation | Neuronal correlates of conscious intention

Commentator

Andrea R. Dreßing
andrea.dressing@uniklinik-freiburg.de
Klinik für Neurologie und Neurophysiologie
Universitätsklinikum Freiburg
Freiburg, Germany

Target Author

Elisabeth Pacherie
elisabeth.pacherie@ens.fr
Ecole Normale Supérieure
Paris, France

1 Introduction

We experience ourselves as agents, performing goal-directed actions in the world. This can be a short-term goal of a motor action like grasping a glass of water, or long-term goal, like the plan to call someone later on. One crucial point in both cases is that we know what we do or want to do. We are aware of our goals before and during acting. This awareness constitutes a conscious intention to act. Even further, we seem to control our actions — at least most of the time — through our intentions. We also have a sense of agency for our actions, which is an immediate feeling of control and authorship (Gallagher 2005). Common sense teaches us that consciousness of our intentions seems to be of unquestionable relevance for our everyday acting.

This experience raises two kinds of questions: *Why* do we experience our intentions as

Table 1: Overview over the different approaches to the explanation of the function of conscious intentions

Anscombe 1963	
Epistemic creation myth	Conscious intentions "provide us with a special kind of self-knowledge" (Pacherie this collection, p. 5)
Bratman 1987	
Pragmatic creation myth	Conscious intentions "[turn us into] temporally extended agents" (Pacherie this collection, p. 3)
Velleman 2007	
Conscious intentions as a spandrel	Conscious intentions are a "by-product of some more general endowments of human nature" (Pacherie this collection, p. 6)
Pacherie this collection	
Social creation myth	conscious intentions "[..are] not just representations of goals but also [...] a specific set of monitoring and control processes, organizing and structuring motor processes that themselves generate movements" (Pacherie this collection, p. 10)

conscious? What is the function of the phenomenal experience of conscious intentions and *how* do intentions exert their role in our acting? These questions address the problem of conscious intentions at two levels. One is about identifying the function and benefits of conscious intentions for our human nature — it is about a myth. The other seems to be above that about understanding, having to do with how the conscious intention exerts its function. It is an attempt to find a scientific, mechanistic explanation about the function of conscious intentions in not only analytical, but also empirical terms (see also Anderson this collection, and Craver this collection).

In her target article, Conscious Intentions: The social creation myth, Elizabeth Pacherie wants to elucidate the function of conscious intentions and reviews teleological approaches on the role of conscious intentions offered by Velleman, as well as his interpretations of Bratman and Anscombe. In addition, she addresses above-mentioned question about the "how" of the causal role of intentions. Based on her hierarchical concept of individual motor actions and scientific data about joint action, Pacherie develops her own approach to the function of conscious intentions. Her idea is supported by the consideration of the potentially striking role of conscious intentions in joint actions (inter-individual actions) regarded as one of the major achievements of the human species. Pacherie's idea is that conscious intentions have the function of controlling motor action and to intra- and inter-individually align our actions with each other.

Answering the initial question of whether we need a creation myth or not, I would like to answer: no, we do not need a myth. We need, as Pacherie tries to give in her target paper, an explanation. What I perhaps like best about the paper is her focus on the role of conscious intentions in action, while the other creation myths described in her paper only consider a more abstract level. We experience the function of conscious intentions strongly and immediately in individual and joint action. Understanding the function of conscious intentions in this context might therefore be one of the most difficult but promising approaches, as it is so essential for human existence. Her social creation myth has the aim to find an explanation of the function and potential causal role of conscious intentions. The importance of this approach, to my mind, is strengthened by Pacherie's attempt to combine empirical data and analytical considerations about motor action and motor control.

In what follows, the teleological and social creation myths are first summarized. Postulating that Pacherie's social creation myth is more

Figure 1: Pacherie's model of intentional action.

than a myth, it should nevertheless fit the current philosophical conceptions and empirical knowledge about the nature of conscious intentions and their causal function. I therefore analyse it according to contemporary approaches in philosophy of mind and I incorporate knowledge of experimental approaches. I argue that according to these approaches, there might arise some difficulties concerning the causal function of conscious intentions in individual and joint action, postulated in Pacherie's social creation myth. Discussing a potential solution, how to understand the "causal" role of conscious intentions in the social creation myth despite those limitations, this commentary could serve as a complementary approach to the social creation myth of Pacherie. I want to argue that a creation myth cannot answer the relevant question, *how* conscious intentions play a role in our acting, without considering the nature of conscious intentions and thereby simultaneously focusing on their causal role.

2 Different myths about conscious intentions

According to Bratman's pragmatic teleological creation myth (1987), intentions are future-directed action plans, offering humans the capacity to "become temporally extended agents" (Pacherie this collection, p. 3). By forming an intention, which is inert and stable, we are able to predict the future and our future planning and form the basis for further intentions. Pacherie criticizes the future-directedness of conscious intentions and says that the pragmatic account of Bratman is incomplete, as it leaves non-pragmatic and present-directed intentions out of sight. Answering to the non-pragmatic function of conscious intentions, Anscombe's teleological creation myth is (1963). Anscombe (1963) gives the whole debate about conscious intentions a highly interesting epistemic turn; her idea of conscious intentions is that they "provide us with a special kind of self-knowledge" (Pacherie this collection, p. 5). Her view of conscious intentions is that they provide immediate knowledge of our intentional actions as they provide an immediate non-observational and direct access to the content of our intention. Velleman's myth about the function of conscious intentions is different (2007). He proposes that conscious intentions are a spandrel and do not have a teleological function on their own, they are a mere "accidental by-product" (Pacherie this collection, p. 6) of two features of human nature: curiosity and self-awareness. From these features arises the concept of intentions that allows human individuals to understand their actions in the world. Pacherie argues that Velleman's approach only shifts the problem of the function of conscious intentions to the function of curiosity and self-awareness.

Pacherie's suggestion is an approach based on empirical knowledge and conceptual considerations about motor cognition. The central element is the suggestion that conscious intentions have a function in motor control. She proposes a three-step hierarchical concept of generation and control of motor actions, developed elsewhere (Pacherie 2008). Motor actions are controlled in an inverse and forward model, comparing error signals on different levels with each other. On the highest level I-Intentions are formed, referring to an abstract goal. These I-intentions allow for the selection of a fitting motor programme, the P-intention. Based on the P-intention the action underlies an online motor control via the M-Intention.

Although providing evidence for unconscious motor control on the lowest level, Pacherie argues that a control function of intentions cannot be denied and remains a "central function of intentions" (this collection, p. 9), mainly on the highest level. Unconscious corrections are sufficient for small misalignments on the lowest level, whereas conscious intentions are necessary in the case of large misalignment between the different levels of motor control. Pacherie declares that "[i]n such case[es] of large

misalignment, error signals would propagate upwards, we would become aware of them, and would shift to a conscious, intentional compensation strategy" (this collection, p. 10). Pacherie also offers a new definition of intentions. She thinks

> of monitoring and control processes as intrinsic to intentions, that is, of intentions as encompassing not just representations of goals but also as a specific set of monitoring and control processes, organizing and structuring motor processes that themselves generate movements. (Pacherie this collection, p. 10)

To summarize, one can understand Pacherie's conscious intentions as having a causal function.

One step further Pacherie suggests that conscious intentions have a coordinative and communicative function in joint action on the basis of her idea that they arise through a hierarchical action control mechanism. Joint action between humans needs a common goal, and success of the joint action is based on our capacity to coordinate actions and share goals, and also to correct and control the individual actions according to the co-agent's actions. Shared actions can, in analogy to the hierarchical model of individual motor control, be controlled on a subconscious low-level. In planned action, however, a hierarchical high level of motor control is needed with which agents represent other's actions and control their own actions according to a shared goal. Mechanisms for joint action discussed in recent empirical science focus on a perceptual framework with joint attention and allocentric spatial orientation (Tomasello & Carpenter 2007; Böckler et al. 2011 cited from Pacherie this collection). The question however, is whether this perceptual information is sufficient for successful joint action. Pacherie concludes that *the conscious intention* is necessary to control intra-individual and inter-individual alignment of actions. One major aspect in joint action is communication of joint goals—so the *conscious* intentions help us to communicate our intentions to others and the other way round, to receive information about the intention of others and to represent them. The influence of other's intentions then guides our own intentions and the following actions.

After this overview over the different creation myths, we should think about the concept of a "myth" itself. A myth in general tries to find an explanation for a phenomenon that we cannot entirely understand. There seems to be a missing piece of knowledge, a gap, which is filled with an idea—the myth. Defining characteristics of a myth since ancient philosophy are its narrative or descriptive character, without being completely irrational. A myth in Plato's sense can neither be falsified nor empirically verified (Partenie 2014). So a myth offers a possible explanation about a phenomenon, without making a claim about truth and without offering a potential empirical approach to the content of the myth. A creation myth about specific functions of conscious intentions is developed, as they seem so unquestionable in our everyday life, and nevertheless, we do not understand, why we have them. The myth—however—does not necessarily need to fit the rules of the physical world.

The social creation myth endows conscious intentions with the important function of a structuring and organizing part in motor action. To my mind, Pacherie develops even more than a myth. The above mentioned characteristics of a myth do not fit Pacherie's empirically based approach. She wants to understand and explain the function of conscious intentions, and her myth wants to *prepare* us for such a deeper understanding. That is an important step, yet it brings certain difficulties. An explanation has to fit into the framework of current scientific knowledge. Most creation myths and most explanation myths make some implicit or explicit assumptions about the nature of conscious intentions, so do the above-described myths. To make full use of Pacherie's contribution, we now should begin by *adding* constraints. Pacherie herself knows these limitations and discusses some of them in her recent paper (2014). What I want to add is a step-by-step-comparison of the empirical and analytical, metaphysical constraints and her hierarchical model in the following sections.

3 Conceptual constraints: The problem of mental causation

Folk psychology tells us that our bodily movements, our actions, are guided by our intentions. One prominent conception of this assumption was developed as part of a non-reductive approach taken by Searle. For Searle, an action is "a causal and intentional transaction between mind and the world" (1983, p. 88). Searle distinguishes between two kinds of intentions, a *prior intention* and an *intention-in-action*. This distinction serves to preserve the difference between an intention as a basic idea or plan, preceding an action and an intention while carrying out an action. If a person P has a prior conscious intention for an action A, P has a representation of A without actually doing A. This is—according to Searle—a deliberative state and represents the *action as a whole*. Contrary, the *intention-in-action* occurs simultaneously with the action, representing the actual *conditions of the action*. Conditions can be regarded as certain steps, an action needs to be carried out. P has an intention-in-action-while A. But, the prior intentions are causally responsible for the intention-in-action and the action itself (Searle 1983).

This is, what—to my understanding—Pacherie's social creation myth stresses as well. At the beginning of Pacherie's paper about conscious intentions, a crucial point is made about the causal connection between intentions and actions: "Roughly, the notion on intentions is of a mental state that represents a goal (and a means to that goal) and contributes, through the guidance and control of behaviour, to the realization of what it represents" (this collection, p. 1). Her considerations about intentionality are about practical intentionality, as they concern conscious intentions in action, not only theoretical or cognitive intentions as mental representations. On the level of metaphysics, her statement could be interpreted along the lines of two kinds of property dualism. First it could be interpreted in a functionalist way in which conscious intentions are *abstract* mental properties possessing a causal role for our actions, in which they have a neuronal realisation or implementation in the background (Lycan 1987; Clark & Chalmers 2002). Secondly, it could be interpreted in a way that declares conscious intentions to be non-reducible, non-physical mental properties, to be local instantiations, which are preceding or accompanying our actions. This notion of conscious intentions describes the conscious intention as a supervening or emerging mental property, which has a physical basis but is not identical or causally dependent with it (Davidson 1980; Kim 1998).

These non-reductional understandings are challenged from a variety of directions. Psychophysical correlations can also be conceptually interpreted using metaphysical models like identity theory (Feigl 1967; Place 1960; Smart 1959) or reductive or eliminative materialism (Churchland 1981), leaving no room for any causal function of conscious intentions. So, according to the most popular models developed after World War II, no conscious intention is a distinct mental entity or an ontological substance in a Cartesian sense. Nevertheless, different assumptions about the nature and the causal function of conscious intentions do exist. To present these in a provocative and simplified way, conscious intentions can either be a mental phenomenon in a physical world and have a causal role (compare: functionalism or non-reductive approaches), or they are causally irrelevant, since they are a by-product of our actions, an epiphenomenon, and as such non-existent (compare this to eliminative materialism).

I now want to focus on non-reductive approaches, as they seem to be relevant for the understanding of Pacherie's social creation myth. Non-reductive approaches, which are supported by the common sense of conscious intentions and intentional action, and which all suggest that our conscious intention initializes the following action, however, might lead to a dilemma. As Heil and Mele put it: "We confront a dilemma. Either we concede that 'purposive' reason-giving explanations of behavior have only a pragmatic standing, or we abandon our conception of the physical domain as causally autonomous" (Heil & Mele 1995, v). The intuition that mental states have causal power is opposed by the rule of causal closure of the

physical world. Kim develops one notion of causal closure with the argument of causal exclusion and supervenience in his essay: *Mind in a Physical World* (1998). In a physical world, in which we do not have a complete physical monism but a non-reductive physicalism with supervenience, two premises are true: (1) every mental property M needs a physical basis P, which is sufficient for the existence of M and on which it supervenes and (2) every physical effect has a sufficient physical cause. Suppose M causes another mental property M*. M* has a physically sufficient basis P*. The problem which arises then is that M and P* as a causally sufficient basis are both responsible for the occurrence of M*, so M has to cause the physical basis P* of M* in a way of mental-to-physical causation. This result conflicts with the premise of causal closure of the physical world, according to which every physical event that has a cause has a *physical* cause (P causes P*). Facing now an over-determination of P*, with two different causally sufficient events *competing* for the causation of P* (M and P), and as P is causally sufficient for M, P seems to be causally sufficient for P*, and M does not have any causal effect itself. A mental phenomenon, according to this view, seems to be causally irrelevant. This is a rather short version of the causal closure-argument; the whole discussion about mental causation and causal closure cannot be displayed here (for an overview see e.g., Heil & Mele 1995). But the causal closure argument seems to be a problem for both: non-reductionist and functional approaches.

What consequences have to be drawn from these considerations about the causal function of conscious intention? Asking for the function of conscious intentions, the different creation myths face the problem of causality in a different way. Both the pragmatic (Bratman) and the epistemic (Anscombe) creation myths are set on a rather abstract mental level of description. Now, coming back to the two-level distinction of intentions introduced by Searle, both teleological myths are about prior intentions. One could say that neither teleological myths require any assumptions about causality, as they do not involve a mind-world directed causality, but rather an intra-mental causality. In the pragmatic creation myth, intentions are preceded and followed by other intentions or intentions to act. Intentions are merely theoretical intentions as they only have a representational character. We can think of the pragmatic creation myth without any real action going on, as an abstract framework for an explanation of the existence of conscious intentions. The epistemic creation myth does not affect the debate about mental causation either. As only a correlation between conscious intention and action is necessary for the epistemic creation myth, it only draws conclusions about self-awareness and does not make any claim about a causal relation of this self-awareness and an action. In Vellemann's view, conscious intentions are a spandrel, a by-product. This model does not imply any explicit claim about causality either. One could go even one step further and postulate that these myths only address the structure of phenomenological experience of conscious intentions and not the intention itself.

In Pacherie's social creation myth, one cannot deny a causal role of conscious intentions any more. This is what I outlined above, referring to Pacherie's definition about conscious intentions. Intentions are not "just" a representation of abstract goals, but of ongoing control and they structure motor processes and "themselves generate movements" (Pacherie this collection, p. 10). Even more, if conscious intentions are needed to modify a joint action, the perception of goal-directed movements of others leads to a mental representation of this action and the formation of a conscious intention for another action follows from this. The problem with conscious intentions in Pacherie's social creation myth could arise when we understand —as outlined above—the I-Intention as purely functional or supervening mental properties in a non-reductive metaphysical framework. Regarding Searle's distinction between *prior intentions* and *intentions-in-action*, I assume that in Pacherie's model the I-intentions could be regarded as *prior intentions* and the P- and M-intentions rather correspond to *intentions-in-action.* To be more precise in this comparison we should talk about the experience of an inten-

tion as conscious mental representation (I-intention). As outlined above, one major analytical constraint against this understanding is the argument of causal closure. If the conscious intention (the I-intention) as a mental phenomenon or a mental representation has a causal function in action, we have to accept downward causation to understand this (which would be against the rule of causal closure). If the I-intentions only supervenes or emerges from its neuronal activity, or is identical with it, then the intention as a conscious mental representation is causally irrelevant and not necessary for the function of motor control. My claim here is that Pacherie's social creation myth needs the causal function of conscious intentions as mental representations to work. Yet requires that we accept the idea of mental causation. As long as the social creation myth is only a myth, we can break the rules of causal closure easily and just offer the gist or general structure of a potential explanation about the function of conscious intentions. Yet if the myth is an explanation, it has to fit the rules of causal closure, and we have to reconsider either the myth or our understanding of causal closure. Last, we could try to create a myth fitting our physical knowledge, yet have to deny the causal effect of the conscious intentions in motor control.

4 Empirical constraints: Current neuroscientific knowledge about the status of conscious intentions

The question about the function of conscious intentions cannot be answered by conceptual considerations alone. The status of practical conscious intentions can be analysed in motor action—as it is done by Pacherie as well—but not only on the level of theoretical hierarchical models of motor initiation and control but on a mere neurophysiological level. Let's begin with a classical example—the Libet-experiments (Libet et al. 1983, 1985) and their modified versions by Haggard & Eimer (1999). Libet and his colleagues designed an experiment to investigate the temporal connection between a voluntary motor activity and the conscious decision—the conscious intention—for this action. They instructed their test persons to voluntarily move their hand and to detect the time at which the urge or the conscious intention to move their hand developed. In parallel, muscle activity was detected via electromyography (EMG) and the readiness potential, a neuronal potential at the beginning of a motor action, was recorded using electroencephalography (EEG). Libet and his colleagues found that the readiness potential can be detected in average 350ms earlier than the test persons experienced the *urge to move* and postulated that according to this finding the decision to move cannot be causally responsive for the action due to a time-based difference. One interpretation of the experiments is that neuronal activity (the readiness potential for the motor activity) occurs before the conscious knowledge of the action itself. So, the conscious intention itself cannot be responsible for a volitional motor action as it occurs later than the subconscious neuronal changes. These findings initiated an on-going debate about the connection between motor activity and the being conscious about this activity, with many neuroscientists supporting the initial hypothesis. Haggard & Eimer detected a lateralized readiness potential (1999). Libet's experiment has been replicated in various alternations, supporting the view that conscious intentions follows pre-conscious brain activity fitting to the movement (Trevena & Miller 2002; Siguru et al. 2004; Rigoni et al. 2011). Similar results were shown for the inhibition of an action (Filevich et al. 2013). fMRI studies (Lau et al. 2004; Soon et al. 2008; Haggard 2008) and transcranial magnetic stimulation-studies postulated a neuronal preceding to motor action similar experimental paradigm (for reviews see Haggard 2005; Shields 2014). One recent fMRI-study for example, reported successful prediction of free choices (addition or subtraction) in the study persons due to fMRI data analysis (Soon et al. 2013). Even single-cell recording in humans—as an objective approach to the self-initiated action—detected neuronal recruitment prior to the intention to act (Fried et al. 2011). The conclusion of above-mentioned experiments frequently is, that the conscious intention of a movement is either an illusion or a post-hoc attribution, generated by the movement itself.

On a conceptual level, there exist other models about conscious motor control besides Pacherie's hierarchical model. An important idea is the idea of intentional binding (Haggard et al. 2002), where an intentional action is causally linked with a certain sensory outcome. In this case, the action and its subsequent effect are perceived as being closer together in time, this generates the phenomenology of causing and independently originating the action, without an actual causal function of the conscious intention. Another current neurobiological theory of motor control is often referred to as comparator model (Frith et al. 2000). Every action consists of two kinds of representations: inverse models that specify motor commands according to sensory perception and forward models that represent the predicted sensory consequences of the movement. When a comparator signals that the sensory consequences of the movement match those predicted by the forward model, we experience this action as consciously intended. Here again, the conscious intention is not causally responsible for the action.

Transferred to the terminology of intentions, this interpretation could mean that a *prior intention* (or I-intention) cannot be causally responsible for an *intention-in-action* (lower level intention) as the neuronal activation pattern for the prior intention was earlier detected than the intention was reported as conscious. What would be the conclusion regarding the social creation myth? As a conscious intention itself—according to the above mentioned interpretation—is not regarded to be causally responsible for the initiation of a motor activity (only the subconscious neuronal activity is responsible) the conscious mental representation of a motor activity in individual or joint action is not causally involved in the processes of motor control. The function of conscious intentions in the social creation myth either stays a myth, as it contradicts the empirical findings, or the myth fits the nature of conscious intentions and we have to reconsider the interpretation of the experiments.

To support the later alternative, one recent study using transcranial magnetic stimulation, a method which allows generating movements by transcranial stimulation of the neurons of the motor cortex, postulated that motor activity is initiated by conscious intentions. A transcranial stimulus was set in the right motor cortex and introduced a tiny muscle twitch, only recordable by EMG. When test persons intended to move their left hand prior to the transcranial stimulus, the transcranial-induced involuntary movement induced a stronger visible motor response. The authors postulated that the conscious intention prepares volitional motor actions by increasing the excitability of the cells in the motor cortex that can produce the movement intended (Zschorlich & Köhling 2013).

There are further some major limitations to the studies, e.g., the subjectivity of the report of the urge to move, and the highly artificial/constructed experimental situation in which the intentional action is carried out. One common objection against an interpretation of the data in the way of Libet and colleagues is that conscious intentions (e.g., the *prior intentions*) are not comparable to the urge to move in an experimental setting but rather are comparable to the decision to participate in the whole experiment. The urge to move would rather be an *intention-in-action* and by this not comparable to a conscious deliberation about an action. Following from the data, a conscious intention is unnecessary or irrelevant (as it occurs "too late") in conscious motor initiation and control could be a too far-reaching conclusion.

5 The problem of causality and the search for a new myth

The aims of the commentary were first to understand, why Pacherie's social creation myth is more than a myth. Second, I elucidated whether it could, in principle, lay the foundations for an explanation based on and in line with philosophical and experimental ideas about mental causation. This discussion was based on the more general question: do conscious intentions have a causal function in the world? To my mind this question cannot yet be answered conclusively, at least according to our current

knowledge. Postulating a lack of causal function of conscious intentions, as based on analytical considerations and empirical data, might be the only possible solution of the problem. The argument from causal closure postulates that a conscious intention as a mental phenomenon is causally irrelevant, because it is not needed to explain a following physical phenomenon. The experimental data might suggest that an intention becomes conscious only after the neuronal activity is detected. Yet, there still is the strong experience of a causal function for our behaviour.

Now, I want to summarize the problems for the social creation myth, based on the above mentioned discussion and I want to consider possible ways to keep and develop the social creation myth as a potential explanation about the function of conscious intentions. The general question about the function of experienced conscious intentions, as Pacherie puts it, is the question about "the normative sense, in which having these functions confers benefits on intention-forming creatures that explains why these creatures have this capacity" (this collection, p. 1). This general question is one of the interpretation and explanation of human nature and not a question about causality. The creation myths of Bratman and Anscombe mainly address the question of why we experience our intentions as conscious and goal directed. The question about real-world, physical causality seems unessential for a pragmatic or epistemic benefit for our being and self-awareness, because the pragmatic or epistemic benefit of conscious intentions arises from the experience of a conscious intention and not from its causal effect. The intentions remain theoretical intentions or mental representations and no downward causality is needed. This does not mean that they cannot have a specific and more complex function, but a strong claim about a localized control-function in motor action is simply not possible. In addition, the epistemic and the pragmatic creation myth as well as conscious intentions considered as a spandrel remain "narrative" accounts and even if they would break the causal closure of the physical world, this would not matter in the context of a myth.

Pacherie's social creation myth first seems to be of a similar kind, explaining human nature and human interaction on the basis of mutual representation of others' actions and formation of joint actions, which do not necessarily have to be causal for joint action, but only for communication intentions and our understanding joint action. The social creation myth is based on the conceptual, hierarchical model of motor initiation and control. It explains conscious intentions not only in a teleological way, but in an analytical way. It is about practical intentionality. Yet, this confronts it with neuroscientific findings and philosophical considerations about causality:

- Conscious intentions in Pacherie's social creation myth exert an organizing and structuring function in the motor process and therefore might have a causal function.
- According to standard metaphysical models for psychophysical relations, the conscious intentions in the myth could be interpreted as a non-reducible mental phenomenon. But if this is the right interpretation, we are confronted with the argument of causal closure and they are either causally irrelevant or we have to deny causal closure of the world.
- According to neuroscientific data, we only know little about the nature of conscious intentions, yet nevertheless we have a strong general trend underlying empirical research, a trend that increasingly supports the assumption of a generation of the wanting or the urge to move from neuronal activation, simultaneously or after, but not prior to the movement.

What does this mean for the social creation myth? Regarding the outlined considerations about causality, the problem of the social creation myth about the function of conscious intentions can be solved in different ways. Either we could regard it as a myth in line with the teleofunctional creation myths, only trying to answer the "why"-question about conscious intentions and leaving questions about causality aside. This could sidestep the problem of causality in an easy yet unsatisfy-

ing way. But if we stick to a myth without acknowledging the physical rules of the world we live in, then we will never achieve more detailed knowledge about the nature and the function of conscious intentions. There will be no epistemic progress after the formulation of the myth itself.

Or we try to preserve Pacherie's approach and keep searching for an explanation about the function of conscious intentions. Yet, if conscious intentions have a structuring and organizing function in individual and joint motor action but—according to the common interpretation of above mentioned empirical data—cannot have distinct causal function, how else can the function be described?

One possible solution is, that we might have to overcome the problem of causality in another way. Most interpretations of neuroscientific experiments and the analytical argumentation of causal closure are based on a temporal, linear one-way causality in the way that A causes B because A precedes B. Additionally, one single intention is typically regarded as the cause of the action in a quasi-linear model. My claim is that the common interpretation that a conscious intention—qua being conscious—can only be causally relevant if the conscious intention precedes the motor action, has to be revised. A first motivation for this claim is the fact that there are multiple theoretical and practical limitations regarding the experiments themselves (e.g., Mele 2011; Radder & Meynen 2012; Pacherie 2014).

But even if the common conceptual interpretation was right, there might be a further terminological problem. In the whole debate about conscious intentions in the social creation myth, we seem to assume that there must be a certain effect of the *being conscious* of the intention. *Because* an intention is conscious, it has an effect to align and control motor action. If it was not conscious, it would not have this effect. To overcome these problems in the debate of the function of conscious intentions, I suggest that a different concept of causation should be considered. This alternative refers to a parallel generation of a conscious intention and movement planning. As it is a parallel process and we might be confronted with two aspects of one and the same process, the conscious intention neither precedes nor follows the action generation, but occurs simultaneously and both are influenced reciprocally (Desmurget 2013). Even further steps may have to be taken. It has been postulated that we cannot trace back the motor action onto one I-Intention in a linear model or to one single place of neuronal activity in the brain. We rather face a semi-hierarchical, parallel and dynamic network from which the motor action arises, without single, identifiable conscious intentions in a direct line of causality but rather fluctuacting activity (Schurger 2014). This would mean that various intentions exist and each of them can influence, control and generate motor action on a neuronal level in parallel, these intentions are among others generated through the observation and interaction with others. Multiple goal representations might form a context for each other. On a conceptual level there would be different I-intentions and different motor programmes going on at the same time. But let us assume that only some of these I-intentions are conscious. Being conscious, for Pacherie, is a necessary condition to exert a motor function and to align actions with others; being conscious is necessary for the causal role in her creation myth. Maybe the function of being conscious could exert a certain weight to an I-intention, not in the way of a linear causality but in a way of dynamic modelling a given social context.

This could save the social creation myth and sheds new light on the interpretation of neuroscientific findings. Whether or not this move answers the question about the function of conscious intentions remains open. The aim should be to further integrate the analytical definitions of mental phenomena and mental causation into neuroscientific research about conscious intentions and try to find a working definition and a concept of what a conscious intention is like. The focus should be on the function of practical conscious intentions and analyse their causal role and function for the hu-

man nature on a neuronal level. Maybe future attempts to arrive at a satisfactory explanation should try to address the causal power of a conscious intention *while* being conscious and not *because of* being conscious.

6 Conclusion

So, do we need a creation myth after all? One thing is certain: conscious intentions unquestionably exist in our experience. We have at least the phenomenal experience of a conscious intention in our acting. As conscious intentions seem so relevant for our human nature we do need a myth about them. But we need even more. Pacherie's social creation myth to my mind is more than a myth; it is one approach, which combines empirical knowledge with a myth about the function and its history. I have only analysed the question of causality from an empirical and metaphysical point of view and its relevance for the social creation myth. In conclusion, we might have to satisfy some further analytical and empirical constraints. Yet, just denying any function of the experience of conscious intentions due to some experimental data or analytical considerations seems premature. A possible solution could be the reconsideration of the concept of causality, to find an explanation of the function of conscious intentions in individual and joint action. Maybe the creation myth and the experimental approach have to be adjusted and be brought together in concept and content, in order to understand the deeper function of conscious intentions. The search for a creation myth should start with creation facts. These facts should help us to elucidate why and how intentions are conscious or at least achieve their phenomenal character, to define the neural correlates or neural correlation in terms of self-organizing, dynamic networks underlying conscious intentions and the causal function in human action, without the limitations of temporal or linear causality and in a more realistic framework of intentional action.

References

Anderson, M. L. (2015). Beyond componential constitution in the brain: Starburst amacrine cells and enabling constraints. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Anscombe, G. E. M. (1963). *Intention.* Oxford, UK: Blackwell.

Bratman, M. (1987). *Intention, plan, and practical reason.* Cambridge, MA: Harvard University Press.

Böckler, A., Knoblich, G. & Sebanz, N. (2011). Giving a helping hand: Effects of joint attention on mental rotation of body parts. *Experimental Brain Research*, *211* (3-4), 531-545. 10.1007/s00221-011-2625-z

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosphy*, *78* (2), 67-90.

Clark, A. & Chalmers, D. (2002). The extended mind. In D. Chalmers (Ed.) *Philosophy of mind* (pp. 643-651). New York, NY: Oxford University Press.

Craver, C. F. (2015). Levels. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Davidson, D. (1980). Mental events. In L. Foster & J. W. Swanson (Eds.) *Experience and theory* (pp. 79-101). Amherst, MA: University of Massachusetts Press.

Desmurget, M. (2013). Searching for the neural correlates of conscious intention. *Journal of Cognitive Neuroscience*, *25* (6), 830-833. 10.1162/jocn_a_00368

Feigl, H. (1967). *The "mental" and the "physical": The essay and a postscript.* Minneapolis, MN: University of Minnesota Press.

Filevich, E., Kühn, S. & Haggard, P. (2013). There is no free won't: Antecedent brain activity predicts decisions to inhibit. *PLoS One*, *8* (2), e53053. 10.1371/journal.pone.0053053

Fried, I., Mukamel, R. & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, *69*, 548-562. 10.1016/j.neuron.2010.11.045

Frith, C. D., Blakemore, S. J. & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B*, *355* (1404), 1771-1788. 10.1098/rstb.2000.0734

Gallagher, S. (2005). *How the body shapes the mind.* Oxford, UK: Oxford University Press/Clarendon Press.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, *9* (6), 290-295. 10.1016/j.tics.2005.04.012

——— (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, *9*, 934-946. 10.1038/nrn2497

Haggard, P. & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126* (1), 128-133.

Haggard, P., Clark, S. & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*, 382-385. 10.1038/nn827

Heil, J. & Mele, A. (1995). *Mental causation.* Oxford, UK: Oxford University Press.

Kim, J. (1998). *Mind in a physical world. An essay on the mind-body problem and mental causation.* Cambridge, MA: MIT Press.

Lau, H. C., Haggard, P. & Passingham, R. E. (2004). Attention to intention. *Science*, *303* (5661), 1208-1210.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529-566. 10.1007/s00221-011-2625-z

Libet, B., Gleason, C. A., Wright, E. W. & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *The unconscious initiation of a freely voluntary act. Brain*, *106*, 623-642.

Lycan, W. (1987). *Consciousness.* Cambridge, MA: MIT Press.

Mele, A. (2011). Libet on free will: Readiness potentials, decisions, and awareness. In W. Sinnot-Armstrong & L. Nadel (Eds.) *Conscious will and responsibility* (pp. 23-33). Oxford, UK: Oxford University Press.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107* (1), 179-217. 10.1016/j.cognition.2007.09.003

——— (2014). Can conscious agency be saved? *Topoi*, *33* (1), 33-45. 10.1007/s11245-013-9187-6

——— (2015). Conscious intentions: The social creation myth. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Partenie, C. (2014). Plato's myths. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/archives/sum2014/entries/plato-myths.

Place, U. T. (1960). Materialism as a scientific hypothesis. *Philosophical Review*, *69* (1), 101-104.

Radder, H. & Meynen, G. (2012). Does the brain "initiate" freely willed processes? A philosophy of science critique of Libet-type experiments and their interpretation. *Theory & Psychology*, *23* (1), 1-19. 10.1177/0959354312460926

Rigoni, D., Kühne, S., Sartori, G. & Brass, M. (2011). Inducing disbelief in free will alters brain correlates of preconscious motor preparation: The brain minds whether we believe in free will or not. *Psychological Science*, *22* (5), 613-618. 10.1177/0956797611405680

Schurger, A. (2014). Intentions and voluntary actions: Reframing the problem. *Cognitive Neuroscience*, *5* (3-4), 213-214. 10.1080/17588928.2014.950214

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind.* Cambridge, UK: Cambridge University Press.

Shields, G. R. (2014). Neuroscience and conscious causation: Has neuroscience shown that we cannot control our own actions? *Review of Philosophy and Psychology*, *5* (4), 565-582. 10.1007/s13164-014-0200-9

Siguru, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A. & Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience*, *7* (1), 80-4. 10.1038/nn1160

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, *68*, 141-156.

Soon, C. S., Brass, M., Heinze, H.-J. & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543-545. 10.1038/nn.2112

Soon, C. S., Hanxi He, A., Bode, S. & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of the USA*, *110* (15), 6217-6222. 10.1073/pnas.1212218110

Tomasello, M. & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, *10* (1), 121-125. 10.1111/j.1467-7687.2007.00573.x

Trevena, M. & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Conscious and Cognition*, *11* (2), 162-90. 10.1006/ccog.2002.0567

Velleman, D. (2007). What good is a will? In A. Leist & H. Baumann (Eds.) *Action in context* (pp. 193-215). Berlin, GER: de Gruyter.

Zschorlich, V. R. & Köhling, R. (2013). How thoughts give rise to action: Conscious motor intention increases the excitability of target-specific motor circuits. *PLoS One*, *8* (12), e83845. 10.1371/journal.pone.0083845

The Causal Role(s) of Intentions

A Reply to Andrea R. Dreßing

Elisabeth Pacherie

In her commentary (Dreßing this collection) on my target article (Pacherie this collection), Dreßing suggests that the story I offer is not just a creation myth but also an attempt to give an explanation of the function of conscious intentions in the physical world and as such answerable to both metaphysical and empirical constraints. Here, I try to clarify which of my claims should be understood as simply speculations about the origins of our capacity of intentions and which I take to be empirical claims. In response to the metaphysical and empirical challenge Dreßing raises, I argue that Dretske's distinction between structuring and triggering causes may help us see how explanations in terms of physical properties and explanations in terms of mental properties may not compete but rather complement each other. I argue that this distinction may also help us assuage certain worries raised by neuroscientific findings.

Keywords
Causal exclusion | Conscious agents | Conscious intention | Creation myth | Intentional action | Intentions | Joint action | Mental causation | Neuronal correlates of intentions | Structuring causes | Triggering causes

Author

Elisabeth Pacherie
elisabeth.pacherie@ens.fr
Ecole Normale Supérieure
Paris, France

Commentator

Andrea R. Dreßing
andrea.dressing@uniklinik-freiburg.de
Klinik für Neurologie und Neurophysiologie
Universitätsklinikum Freiburg
Freiburg, Germany

1 Introduction

In her commentary, Andrea Dreßing (this collection) suggests that I might have been too timid in calling the story I tell about the social function of intentions in my target article (Pacherie this collection) a creation myth. She encourages me take a bolder stance, claiming that the story I offer is not just a myth but also an attempt to give an explanation of the function of conscious intentions in the physical world. Indeed, part of my story is intended as more than a myth and so my first task here will be to clarify where I draw the line between empirical claims and myths.

Dreßing also points out that an explanation, as opposed to a mere myth, has to fit into the framework of current scientific knowledge and is therefore subject to both metaphysical constraints and empirical constraints. I concur. In what follows, I will argue, however, that my general predicament with regard to conceptual or metaphysical constraints is not so different from the predicament of the other myth-tellers I

discuss in my article, as Dreßing suggests. Nor indeed is it direr than the predicament all philosophers of mind working within a naturalistic framework face. Finally, certain empirical findings have been interpreted as showing that conscious intentions play no role in action initiation. I also try to address this challenge.

2 Myths vs. empirical claims

In my target article, I use the phrase "creation myth" first as a dramatization device. Typically, we do not feel the urge to formulate myths about things we deem insignificant. Talking of a social creation myth was thus a way of emphasizing the importance of the social function of intentions, a function largely neglected in traditional accounts of intentions. Second, I also wanted, following Velleman (2007), to convey a note of caution. A myth, as Dreßing points out, can neither be falsified nor empirically verified. It offers a possible explanation about a phenomenon, without making a claim about truth. But I perhaps wasn't clear enough what I was trying to be cautious about and where I drew the line between empirical claims and ultimately unverifiable explanations. So let me now draw this line more firmly.

To do this, let me distinguish three different questions about intentions and examine how they may relate. The three questions are: *what* roles or functions (in a non-teleological sense) do intentions play in human agency? *How* can intentions play these roles? *Why* do we have intentions in the first place? In my view the *what*- and *how*-questions are both empirical questions for which mythical answers won't do. The *why*-question, as I understand it, is a question about the origins of capacity for intention. How come we have such a capacity? Why was it established?

The focus of the account I proposed, as well as the focus of the alternative accounts by Bratman (1987), Anscombe (1963), and Velleman (2007) with which I contrast it in my article, is on the *what*- and *why*-questions. However, I offered my story as a creation myth only to the extent that it was meant to address the *why*-question. As answers offered to the *what*-question, my claims were meant as empirical claims. I take it that the claims made by Bratman, Anscombe, and Velleman about the epistemic and pragmatic functions of intentions, when understood as answers to the *what*-question, should also be interpreted as empirical claims.

Now, how do the *what*- and the *why*-questions relate? One way to relate them is by assuming that intentions do not just have a function or functions in a value-neutral sense—things that they do—but a teleofunction in the evolutionary sense, that is, something that they do that confers some benefit or advantage on creatures with a capacity for intentions, and in this sense explains why these creatures have this capacity.

Velleman cautions us against this teleofunctional move. First, as his discussion of Bratman's and Anscombe's accounts makes clear, the *what*-question about intentions can be given complementary answers in terms of both pragmatic and epistemic roles, leaving us with several possible teleological stories. Second, Velleman also warns us against assuming direct links between answers to the *what*-question and answers to the *why*-question. The spandrel story he tells is meant to suggest that a capacity for intentions may only be a by-product of other capacities and thus that our capacity for intentions could be nothing more than an (admittedly very fortunate) accident. Finally, in calling his own story a creation myth as well, Velleman is also pointing out that our speculations about the origins of intentions are most likely beyond falsification or empirical verification.

Similarly, in offering my social function story as an answer to the *why*-question, I was not making a claim to truth. Rather, I was trying to broaden the terms of the debate to also include consideration of the social dimension of intentions. If we are considering what possible teleofunction intentions could have, then we should pay more attention to the benefits we derive from being able to act jointly in a flexible manner. If we are tempted by a story that views a capacity for intention as simply a by-product of more general capacities, then, among these more general capacities, we should pay

serious heed to our capacity for sociality and cooperativeness.

Turning now to the relations between the *what*-question and the *how*-question, I take it that the empirical standing of an answer to the *what*-question ultimately depends on whether this answer can be backed up by a convincing answer to the corresponding *how*-question. The validity of any empirical claim about the causal roles of intentions in human agency will remain in doubt unless one can see how it is at all possible for intentions to play these roles (Dreßing's metaphysical constraints), and it will also remain in doubt if appears to be in contradiction with well-established empirical facts (Dreßing's metaphysical constraints).

Since my claims about the functions of intentions qua answers to the *what*- rather than the *why*-question are intended as empirical claims, they are not insulated from these metaphysical and empirical worries. Let me address them in turn.

3 Metaphysical worries

Dreßing points out that my claim that intentions have a causal role to play in the online control of action confronts me with the problem of mental causation. She also suggests that this problem is more pressing for me than it is for the accounts of the functions of intentions proposed by Bratman and Anscombe. While I agree that the problem of mental causation is an issue for me, I disagree with her assessment that it isn't as serious a worry for these accounts.

First, let me clarify that when I talk about conscious intentions and their causal role, I am concerned with what Ned Block (1995) calls access consciousness rather than with phenomenal consciousness. In other words, my claims are about intentions qua conscious states exploiting and conveying information globally available in the cognitive system for the purposes of reasoning, speech, and high-level action control. My account thus faces the "easy" problems of consciousness rather than the "hard" problem (Chalmers 1995). I share Chalmers sanguine assessment about phenomena pertaining to access consciousness:

> There is no real issue about whether these phenomena can be explained scientifically. All of them are straightforwardly vulnerable to explanation in terms of computational or neural mechanisms. (1995, p. 201)

This is not to say, however, that in confining oneself to phenomena of access consciousness one can eschew all metaphysical conundrums. In particular, as pointed out by Dreßing, the mere fact that Cartesian dualism has fallen out of favour and that the vast majority of philosophers and cognitive scientists are nowadays willing to embrace some form of materialist monism doesn't insure the dissolution of philosophical worries about mental causation. The version of the problem of mental causation that non-reductive physicalists, whatever their exact persuasion, are confronted with is the Causal Exclusion Problem: how could mental properties play a causal role given that they appear to be screened off by their physical realizers?

Dreßing argues that this problem is more pressing for my view than for the pragmatic (Bratman) or the epistemic (Anscombe) creation myths, the reason being that these latter two teleological myths are about prior intentions and that neither "require any assumption about causality, as they do not involve a mind-world directed causality, but rather an intra-mental mental causality" (Dreßing this collection, p. 6). Dreßing also claims that Velleman's view does not imply any explicit claim about causality either, since on this view intentions are a spandrel or a by-product.

I disagree with this assessment for three reasons. First, as I explained in section 2, while the speculative character of these stories qua answers to the *why*-question may justify labelling them as creation myths, the stories also offer answers to the *what*-question. In that regard their claims about the epistemic or pragmatic roles of intentions should be taken as empirical claims. Thus, even if we go along with Velleman's claim that a capacity for intentions

is a spandrel and that the epistemic and pragmatic functions of intentions are not teleofunctions, they are nevertheless functions in the ordinary functionalist sense and we still need an explanation of how intentions can play these epistemic and pragmatic roles.

Second, the Causal Exclusion Problem is a problem for anyone espousing a non-reductive form of materialist monism,[1] whether their primary concern is with intra-mental causation or with mind–world causation. Suppose that a state S has the mental property M (e.g., the property of being an intention to go to London on Monday) and a physical basis P, suppose that S' has the mental property M' (e.g., the property of being an intention to buy a train ticket to London) and a physical basis P', and suppose that S'' has the mental property M'' (e.g., the property of being a being a belief that one will go to London on Monday) and a physical basis P''. On a Bratmanian pragmatic account of intentions, I would want to be able to say that my intention to go to London on Monday causes, via further means-end reasoning, my intention to buy a ticket to London. But how can the mental property M of S play a causal role in bringing about a state S' with mental property M', given that they appear to be screened off by the physical properties P and P'? Similarly, with regard to the epistemic function of intentions, how could I say that my intention to go to London on Monday causes my belief that I will go to London on Monday, given that mental properties M and M'' appear to be screened off by the physical properties P and P''?

Third, while it is true that on Bratman's account future-directed intentions may only cause behaviour through the mediation of present-directed intentions, still Bratman insists that the whole point of having a capacity for intentions is to produce behaviour that contributes in the long run to our securing greater desire-satisfaction. Similarly, on a reliabilist reading of Anscombe's epistemic claim that intentions embody knowledge of our actions, they do so because intentions reliably cause what they represent. As Velleman puts it, "[u]nless an intention with the content 'I'm going to move my toe' reliably causes my toe to move, it won't amount to practical knowledge" (Velleman 2007, p. 201). Thus, Bratman, Anscombe and Velleman cannot be exonerated from the task of explaining how mental states can cause behaviour.

With respect to the problem of mental causation, we are all in the same boat. The metaphysical standing of my account is no less or more precarious than the standing of these other accounts. Are we then all metaphysically doomed? Readers should not hold their breath; I have no new, unassailable solution to the problem of mental causation to offer. Yet, it would certainly be premature to claim that the problem of mental causation is insoluble. Many lines of response have been proposed and are currently being explored (for a review, see Robb & Heil 2014). I cannot discuss all these accounts here. Let me just say that the approach I find most congenial stems from Fred Dretske's work (Dretske 1988, 2004) on psychological explanations of behaviour. Dretske distinguishes between triggering causes and structuring causes, where a triggering cause is an event that initiates or triggers a causal chain of events, and a structuring cause the cause of the process or setup that makes a given triggering cause produce the effect it does. To take an example from Dretske (2004), moving a computer mouse is the triggering cause of cursor movement, but hardware and programming are the structuring causes of cursor movement. Dretske's central claim is that mental states and events are best analysed as structuring rather than triggering causes of behaviour. On this view there is no competition between physical and psychological or mental explanations, since they have different explananda. While the triggering physical properties explain bodily motion, i.e., explain why bodily motions occur at a certain point in time, the structuring mental properties explain

1 While this issue is not at the heart of Bratman's preoccupations, I think we can safely assume that he would want his account of intentions to be compatible with physicalism. I won't dwell here on Anscombe's metaphysical view, except to say that she was no materialist herself but was also highly suspicious of Cartesian dualism (Anscombe 2008). Suffice it to say that many of the philosophers who nowadays embrace the view that intentions have an epistemic function, would want this claim to be compatible with a physicalist stance.

behavior, i.e., they explain why in circumstances of a certain sort, bodily motions of this kind rather than that kind are produced.

Much work remains to be done in order for us to understand more precisely how structural causes operate and in particular how they can do so in the dynamic way needed to account for the plasticity and flexibility of human behaviour. In this respect, Dretske's account remains largely under-developed (for recent work on this issue, see e.g., Slors 2015; Wu 2011). Dretske's approach in terms of structuring causes has the great merit, however, of offering a potential solution to the Causal Exclusion Problem and to let us see how explanations in terms of physical properties and explanations in terms of mental properties may not compete but rather complement each other. As we will now see, thinking of intentions as structural causes of action rather than triggering causes can also help us assuage certain empirical worries.

4 Empirical worries

The claim that conscious intentions play a causal role in action production should be compatible with our best empirical knowledge on how action is produced. The main empirical worries this claim confronts come from neuroscientific findings that have been interpreted as showing that the time of onset of conscious intentions is not compatible with their being the initiators of actions.

The most famous of these experiments are Libet's studies on "readiness potential" (Libet et al. 1983; Libet 1985). In these studies, subjects were asked to flex their wrist at will and to note when they felt the urge to move by observing the position of a dot on a special clock. While subjects were both acting and monitoring their urges (intentions, decisions) to act, Libet used an EEG to record the activity of prefrontal motor areas. On average, participants reported the conscious intention to act, which Libet called the W-judgement, about 200ms before the onset of muscle activity. By contrast, the EEG revealed that preparatory brain activity, termed by Libet type II readiness potential (RP), preceded action onset by about 550ms. In other words, their brains started preparing the action at least 350ms before the participants became aware of their intention to act. This led Libet to the conclusion that the wrist-flexing actions in his experiments were not initiated by conscious intentions but were initiated instead by the (unconscious) RPs.

These experiments and Libet's interpretation of his findings have been widely discussed (see e.g., Banks & Pockett 2007; Bayne & Pacherie 2014; Mele 2009; Nahmias 2002; Pacherie 2014; Roskies 2011) and commentators have pointed out a number of methodological problems with Libet's paradigm as well as conceptual problems with his interpretation of his results. Let me focus first on one methodological problem and one attempt to address it. I will then consider one conceptual problem

Libet argues that it is the RP rather than the conscious intention that initiate the agent's action. If RPs are the initiators of the action, there should be a robust correlation between them and the actions they cause: we should expect RP events to be "immediately" followed by the appropriate action, or, to put it the other way round, we should expect that when there is no movement, there is also no RP event. As several commentators have observed (e.g., Mele 2009; Roskies 2011), the back-averaging techniques used in the experiment do not allow us to ascertain whether this is indeed the case. Because the RP on any one trial is obscured by neural noise, what is presented as "the RP data" is determined by averaging the data collected on a large number of trials. In order to compute this average, the EEG recordings on different trials need to be aligned, and this requires some fixed point that can be identified across trials. Since in Libet's experiments action onset serves as the needed fixed point for the alignment of EEG recordings, any RPs that are not followed by an action simply won't be measured, and so we don't know how robust the correlation between the RP and Libet-actions is.

In a recent experiment, Schurger and colleagues (Schurger et al. 2012) used a modified Libet task to circumvent the limitations of back-averaging techniques. Their aim was to

test the proposal that RPs correlate with pre-decision activity rather than, as Libet proposed, with activity that coincides with, or is subsequent to, the agent's decision. Schurger and colleagues proceeded on the assumption that the decisions of the participants in Libet's experiment can be modelled—as neural decision tasks typically are—in terms of an accumulator-plus-threshold mechanism: decisions are made when relevant evidence accumulated over time reaches a certain threshold. Given that in Libet's task subjects are explicitly instructed not to base their decision on any specific evidence, Schurger and colleagues proposed in this instance that the decision process amounts to simply shifting premotor activation closer to the threshold for initiation of the movement and waiting for a random threshold-crossing fluctuation in RP. Thus, Schurger and colleagues predicted the same premotor activation build-up as Libet when a movement is produced. However, whereas on Libet's post-decision interpretation of this build-up there should be no premotor activity (and hence no RPs) when no movement is produced, on Schurger and colleagues' stochastic decision model there should be continuous random fluctuations in RPs even when no movement is produced. Schurger and colleagues reasoned that it should be possible to capture these fluctuations by interrupting subjects in a Libet task with a compulsory response cue and sorting trials by their reaction times. On the assumption that the interrupted responses arise from the same decision accumulator as the self-initiated ones, and on the assumption that close-to-threshold activity reflects spontaneous fluctuations of RPs rather than mounting preparation to move building over the course of the entire trial, slow and fast reaction times should be distributed equally within trials. In their *Libetus Interruptus* task, they found, as they had predicted, that slow and fast responses to interruptions were distributed equally throughout the time span of the trial.

These results cast serious doubt on Libet's claim that the neural decision to move coincides with the onset of the RP, since spontaneous fluctuations of RPs happen all the time. Therefore, they also cast doubt on his further claim that since RP onset precedes the urge to move by 350ms or more, conscious intentions can play no role in the initiation of the movement. If instead the neural decision to move coincides with a much later threshold-crossing event, it remains at least an open possibility that this event coincides with and constitutes the neural basis of a conscious urge to move. Schurger and colleagues take no stand on the exact relation between the conscious urge to move and their threshold-crossing event. They insist, however, that this threshold-crossing event should not be interpreted as *the* cause of the movement but rather as just one of the many factors involved in the causation of self-initiated movements. This leads me to my final point.

One conceptual problem with Libet's interpretation of his findings and also, as Dreßing points out, with most interpretations of neuroscientific experiments and a large part of the philosophical debates on mental causation and causal exclusion lies in the conception of causality that is assumed, "namely a temporal, linear, one-way causality" (Dreßing this collection, p. 10). I agree with Dreßing's suggestion that a different concept of causation should be considered, one that allows for multiple causal processes to operate in parallel and to exert influence on one another. This is indeed the spirit of the dynamical model of intentions I have proposed elsewhere (Pacherie 2008). In particular, I insisted that a distal intention does not cease to exist and play a role once a corresponding proximal intention has been formed (and the same goes for proximal and motor intentions). What I suggested is that all three levels of intentions operate simultaneously, each exerting its own form of control, as well as operating together with unconscious processes. Following Dretske's lead, we can think of intentions as structuring rather than as triggering causes of action. On the dynamic hierarchical model of intentions I have proposed, we can further think of the structures set up by intentions as nested. This means that we don't need intentions to initiate actions for them to play a causal role in the production of action. This also means that the intentional online control that I argued was an

important pragmatic function of intention may be best conceived as a form of re-structuring, necessary only when the initial structuring is inadequate.

5 Conclusion

In her commentary, Dreßing suggested that the story I told about intentions should be viewed not just as a creation myth but as an attempt to give an explanation of the function of conscious intentions in the physical world. I tried to clarify exactly what I offered as merely a creation myth, namely the story given in answer to the question "*Why* do we have intentions in the first place?" and what I offered as empirical claims, namely my story as an answer to the question "*What* roles do intentions play in human agency?"

Dreßing also stresses that as an account of the roles intentions play in agency, my story has to meet both metaphysical and empirical constraints. In particular, she suggests that my claims about the role of intentions in action control makes the Causal Exclusion Problem more pressing for me than for other myth-tellers. I argued that the problem is actually equally pressing for all of us who want their views to be compatible with physicalism. I suggested that Dretske's distinction between structuring and triggering causes and his view that mental properties should be understood as structuring causes may offer a solution to this metaphysical problem. Finally, Dreßing remarks that my claims concerning the role of conscious intentions appear to clash with certain findings from neuroscientific experiments. In response, I briefly discussed the most famous of these experiments, Libet's RP experiments, and pointed out some of their limitations. I also questioned, together with Dreßing, the conception of causation with which these debates tend to operate, and suggested that Dretske's distinction between structuring and triggering causes may also help to reconcile neuroscientific findings and claims about the causal roles of intentions.

References

Anscombe, G. E. M. (1963). *Intention (second edition).* Oxford, UK: Blackwell.

——— (2008). *Faith in a hard ground: Essays on religion, philosophy and ethics (Vol. 11).* Exeter, UK: Imprint Academic.

Banks, B. & Pockett, S. (2007). Benjamin Libet's work on the neuroscience of free will. In M. Velmans & S. Schneider (Eds.) *The Blackwell companion to consciousness.* London, UK: Blackwell.

Bayne, T. & Pacherie, E. (2014). Consciousness and agency. In J. Clausen & N. Levy (Eds.) *Springer Handbook of Neuroethics* (pp. 211-230). Dordrecht, NL: Springer.

Block, N. (1995). On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, *18* (2), 227-247. 10.1017/S0140525X00038188

Bratman, M. (1987). *Intention, plans, and practical reason.* Cambridge, MA: Harvard University Press.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2* (3), 200-219.

Dretske, F. I. (1988). *Explaining behavior: Reasons in a world of causes.* Cambridge, MA: MIT Press.

——— (2004). Psychological vs. biological explanations of behavior. *Behavior and Philosophy*, *32* (1), 167-177.

Dreßing, A. R. (2015). Conscious intentions: Do we need a creation myth? In T. Metzinger & J. M. Windt (Eds.) *Open MIND.* Frankfurt a. M., GER: MIND Group.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8* (4), 529-566. 10.1017/S0140525X00044903

Libet, B., Gleason, C. A., Wright, E. W. & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, *106* (Pt 3), 623-642.

Mele, A. (2009). *Effective intentions: The power of conscious will.* New York, NY: Oxford University Press.

Nahmias, E. (2002). When consciousness matters: A critical review of Daniel Wegner "the illusion of conscious will". *Philosophical Psychology*, *15* (4), 527-541. 10.1080/0951508021000042049

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107* (1), 179-217. 10.1016/j.cognition.2007.09.003

——— (2014). Can conscious agency be saved? *Topoi*, *33* (1), 33-45. 10.1007/s11245-013-9187-6

——— (2015). Conscious intentions. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a. M., GER: MIND Group.

Robb, D. & Heil, J. (2014). Mental causation. In E. N. Zalta (Ed.) *The Stanford encyclopedia of philosophy (spring 2014 edition)*. http://plato.stanford.edu/archives/spr2014/entries/mental-causation

Roskies, A. (2011). Why Libet's studies don't pose a threat to free will. In W. Sinnott-Armstrong & L. Nadel (Eds.) *Conscious Will and Responsibility* (pp. 11-22). New York, NY: Oxford University Press.

Schurger, A., Sitt, J. D. & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences*, *109* (42), E2904-E2913. 10.1073/pnas.1210467109

Slors, M. (2015). Conscious intending as self-programming. *Philosophical Psychology*, *28* (1), 94-113. 10.1080/09515089.2013.803922

Velleman, D. (2007). What good is a will? In A. Leist & H. Baumann (Eds.) *Action in Context* (pp. 193-215). Berlin, GER: de Gruyter.

Wu, W. (2011). Confronting many-many problems: Attention and agentive control. *Noûs*, *45* (1), 50-76. 10.1111/j.1468-0068.2010.00804.x