Synthese (2016) 193:2129–2158
DOI 10.1007/s11229-015-0835-9

Chimps as secret agents

Caroline T. Arruda[1] · **Daniel J. Povinelli**[2]

Received: 26 September 2014 / Accepted: 18 July 2015 / Published online: 5 August 2015

Abstract We provide an account of chimpanzee-specific agency within the context of philosophy of action. We do so by showing that chimpanzees are capable of what we call *reason-directed* action, even though they may be incapable of more full-blown action, which we call *reason-considered* action. Although chimpanzee agency does not possess all the features of typical adult human agency, chimpanzee agency is evolutionarily responsive to their environment and overlaps considerably with our own. As such, it is an evolved set of capacities for goal-directed behavior, which solves problems that chimpanzees (and humans) naturally encounter. Thus, it ought not be understood as a deficient instance of human agency.

Keywords Animal agency · Chimpanzees · Animal minds · Animal reasoning

1 Introduction

If you were to see a friend reach to take a book off a shelf and begin reading it, you would correctly conclude that she had engaged in an intentional action, and that she possesses the capacities necessary and sufficient for intentional agency. Likewise, if you saw a chimpanzee named Megan pick up a stick and pull a banana toward her, you might want to conclude that she too is engaged in an intentional action and possesses

✉ Caroline T. Arruda
ctarruda@utep.edu

Daniel J. Povinelli
djp3463@louisiana.edu

1 Department of Philosophy, The University of Texas at El Paso, El Paso, TX 79902, USA

2 Department of Biology, University of Louisiana, 104 E University Ave, Lafayette, LA 70504, USA

some form of the capacities—perhaps more minimal—necessary and sufficient for intentional agency. But is the latter inference warranted?

Using existing and fairly uncontroversial resources in philosophy of mind, philosophy of action, and comparative psychology, it may seem like it would be easy to show that Megan is an agent, particularly of the more minimal kind. As it turns out, this is more difficult than it appears.

Let us begin with the difficulties that arise within philosophy of mind. Despite their natural inclinations, philosophers of mind do not speak with a unified voice about this issue. Broadly understood, there are three ways that the philosopher of mind could approach the question of whether Megan, or non-human animals in general, are agents. Some arguments, such as those advanced by Carruthers (1996; but compare with Carruthers 2013), Stich (1979), Davidson (2001b, d), Dreckman (1999), McDowell (2007a, p. 343; 2007b), and, to some degree, Fodor (1975, but compare with 1990, pp. 152–153), suggest that we need language or other unique cognitive abilities (see Clark 2005) to have concepts, concepts to have propositional attitudes, and finally, propositional attitudes to form intentions. Note that some of these philosophers entertain the possibility that animals do indeed have propositional attitudes, regardless of their precise capacity for language (Carruthers 1996; Fodor 1990, pp. 152–153; Dreckman 1999). Nonetheless, if one does indeed need language to form intentions *and* one doubts that chimpanzees possess language in the relevant form, then it will be difficult to show that chimpanzees are agents.[1] Of course, other arguments, such as those advanced by Andrews (2000), Bermúdez (2003, 2006, 2009), Beisecker (1999), Carruthers (2005, pp. 92–94), DeGrazia (2009), Dennett (1987, 1996), Dretske (1998), Dreyfus (2007), Glock (2010), Hurley (2003a), Lurz (2003), Pacherie (2011), Rowlands (2006), and Saidel (2009), deny this claim and its consequent.

A second approach, exemplified by the work of Bermúdez (2003, 2006), Dretske (2006), Hurley (2003a, b, 2006), and Millikan (1995, essays 7–8; 2006), focuses on whether chimpanzees are able to *reason*, rather than focusing specifically on their capacity to form *intentions to act*. Yet as long as we think that reasoning about action (or practical reason) and reasoning about the truth of our beliefs (or theoretical reason) are distinct, it is not clear that establishing chimpanzees' ability to reason will be sufficient to show that they are agents.

A third approach in the philosophy of mind, such as that developed by Bermúdez (2003, 2006) and Hurley (2003a, b, 2006), appeals to the possibility that chimpanzees indeed hold attitudes relevant for forming intentions, but not in the form embodied by standard propositional attitudes such as "I intend that I eat the banana." This approach may be one way of successfully tackling the problem of chimpanzee agency, but it requires that we make significant commitments to the possibility of non-conceptual content—a theoretically burdensome move, which, in this paper, we will seek to avoid.

We are not trying to engage directly with these debates in philosophy of mind. But we do want to determine whether we can establish that chimpanzees are intentional agents without making significant commitments in the debates about animal rational-

[1] There are philosophers who defend the claim that some animals, despite not having language, possess propositional attitudes. See Saidel (2009) for an example.

ity (broadly understood) and in debates about non-conceptual content, in particular. We believe that being able to provide such an account would advance the debates about chimpanzee agency in that it would show that it is possible to develop such an account without settling the aforementioned contentious debates. It would also entail that being an agent (ape or otherwise) may not depend on the ability to engage in higher-order reasoning or to possess non-conceptual content. To do so will, however, require us to fully engage with important accounts of animal rationality, as exemplified by the work of Bermudez and Hurley (see Sect. 4.2).

If current resources in the philosophy of mind prove unhelpful, what might philosophy of action provide? One way to do so is to lower the bar for what counts as genuine intentional action. For example, many argue for a distinction between merely purposive action and what some philosophers have called "full-blooded action" (Butterfill 2001; Butterfill and Apperly 2013; Bratman 2007a, p. 197; Korsgaard 2009, pp. 98–99, 174–175; Velleman 1989, p. 124) in an attempt to show that non-human animals, such as chimpanzees, are agents. Can these resources allow one to argue that while chimpanzees are not agents like adult human beings, they are nonetheless capable of purposive action?

Possibly, but problems immediately emerge when we examine other commitments held by the philosopher of action. She typically holds that the necessary and sufficient conditions for agency are far more robust than simply being able to engage in purposive action (Anscombe 1957; Bratman 2007a, b; Davidson 2001b; Frankfurt 1978, 1998; Korsgaard 2009; Nussbaum 2006; Paul 2009; Setiya 2007; Velleman 1989, p. 125-ff; Wallace 2006a, b). On this view, agency requires that one must be able to engage in action *on the basis of reasons* that *one takes to be one's own* so that one can *deliberate about how to act* (among other capacities). As we will argue in greater detail below, we take this to mean that the distinction between "purposive" and "full-blown" intentional action is sufficiently underspecified as to mask two inconsistent commitments: (1) there are robust necessary and sufficient conditions for agency (specified above), and (2) chimpanzees, while not able to meet these conditions, are nonetheless (and inexplicably) "agents" by virtue of engaging in purposive action.

So why are we bothering to show that chimpanzees are agents when it appears that we already have reason to abandon the project?

The reason is because we take Megan's actions with the stick and banana quite seriously. In particular, we believe there is a way to show that our opening intuition that Megan is an agent is correct if we can successfully modify the conception of what it is for her to be an agent. Two options arise. The first is to lower the bar for agency. The second is to determine if there is a conception of agency specific to chimpanzees. We favor the second option.[2]

In what follows, we offer a justification for chimpanzee-specific agency within the context of philosophy of action. Here one should note that 'specific' does not entail a form of agency that it *exclusive to* chimpanzees. It merely underscores that it is a form of agency that takes into account abilities that chimpanzees have been

[2] As will become apparent, our account clearly extends to other non-human species. However, any attempt to distinguish human-specific agency from non-human animal agency might profit by examining our evolutionary closest living relations, chimpanzees and other great apes.

shown to possess. We do so by showing that Megan is capable of what we will call *reason-directed* action, even though she may be incapable of a more full-blown form of action, which we will call *reason-considered* action. That is, she is able to act for reasons even though she does not take those reason as her own in the more cognitively demanding way in which adult human agents do so. Some may interpret this as a claim that chimpanzees are deficient agents compared to us. This is not our argument. We claim that even though chimpanzee agency does not possess all the features of typical adult human agency, chimpanzee agency is evolutionarily responsive to their environment. As such, it is an evolved set of capacities for goal-directed behavior, which solves problems that chimpanzees naturally encounter. Thus, it ought not be understood as a deficient instance of human agency. Instead, it should be thought of as an evolutionary variant of agency. Human agency could then be envisioned as an elaborated or hypertrophied instance of agency. If we're right, the type of agency exhibited by chimpanzees is ubiquitous in the natural world. Human agency is an evolutionary recent, and highly unusual form of the capacity.

Naturally, this is not the only route that the philosophy of action provides. Some philosophers, such as Hurley (2003a, b), Bermúdez (2003), and Steward (2009), have also tried to make the case for non-human animal agency by reconsidering the capacities that are necessary and sufficient for agency.[3] Since we take a similar approach, albeit with very different conclusions from the aforementioned philosophers, we critically examine both Hurley's and Bermúdez's respective arguments

2 What's so difficult about showing that chimpanzees are agents?

If it were uncontroversial that chimpanzees were capable of second-order thought and possessed propositional attitudes, then the case for chimpanzee agency would already be made. Notably, some philosophers, such as DeGrazia (2009), Jamieson (2009) and Saidel (2009), might claim that there are empirical or conceptual reasons for thinking that chimpanzees are capable of second-order thought and that they possess propositional attitudes. If you think that this is the case, we invite you to stop reading now. You already have the tools necessary to show that Megan is an agent.

But we want to deal with the more difficult, empirically motivated case, wherein we must show that Megan is agent even though she differs qualitatively from humans in her capacity for second-order propositional thought (Bermúdez 2003; Penn et al. 2008; Call and Tomasello 2008).[4]

Broadly speaking, there are three existing approaches to deal with this more difficult case, all of which are ultimately insufficient for showing that chimpanzees are agents: the intentional stance, theories of merely purposive agency in the philosophy of action and, finally, comparative psychology's account of the pursuit of goals. Given that we

[3] As discussed above, other philosophers of action take their view to accommodate non-human animal agency, but they do not provide a specific defense of it. These views, by contrast, specifically address this question.

[4] Nothing in what we are about to argue depends on the assumption that they do not have propositional attitudes; rather, if it turns out that they do, then all the better for the case that chimpanzees are agents. Our account is not inconsistent with this possibility.

have set out to develop an account of chimp agency without commitments about the nature of animal rationality, it is incumbent of us to briefly survey the other available routes to establishing chimpanzee agency, and highlight why they are unable to do so.

2.1 Dennett's intentional stance

The first approach is Dennett's (1987) intentional stance. Dennett (1987) argues that a creature[5] is an intentional agent just in case the strategy of assuming that said creature is a rational agent who exhibits intentional actions is successful. The strategy is successful—that is, the entity ought be treated as an intentional agent—if one is able to continue to predict reliably the creature's actions after having initially assumed that it is an intentional agent (p. 15). When we adopt the intentional stance, we assume that we can justifiably attribute beliefs, desires and reasons to agents (or would-be agents) as long as we can suppose what a "real" agent in the same context would reasonably hold and do. If our "possible" intentional agent acts in a similar way, we can treat the creature as if the creature were an agent.[6] In this regard, we attribute beliefs, desires and reasons to possible agents not because they are, in fact, agents, but because these attributions function as predictive tools. Our attribution will be right if and only if the possible agent behaves in ways that are *best explained* in light of these attributions.[7]

On first pass, it would seem that Dennett's view is perfectly suited to make the case that chimpanzees are agents. Megan's use of the stick to obtain an appetizing banana is best explained by the intentional rather than the design stance. Unlike an alarm clock or a thermostat, we can reliably attribute relatively rich belief states to chimpanzees like Megan in a variety of circumstances—including cases where they initially use, often unsuccessfully, startlingly inappropriate means to obtain desired objectives (Köhler 1925; Povinelli 2003, 2012).[8] This reliable attribution indicates that the intentional stance is predictively successful such that we can conclude, using Dennett's view, that chimpanzees can be treated as if they were agents.

But there are two problems with his view as it applies to the question of whether chimpanzees are agents.[9] The first problem is that when used in this way, the inten-

[5] Note that Dennett's account allows for the possibility that this interpretative strategy can accommodate non-human systems.

[6] Traditionally, the intentional stance in interpreted as justifying the following judgment: We can treat the creature *as if* the creature were an agent, but we are not warranted in inferring that the creature *is in fact* an agent.

[7] Of course, it is not the case that all entities deserve this initial attribution (Dennett 1987, pp. 16–17). As Dennett notes, an alarm clock that alerts its owner to wake up at six o'clock in the morning is best explained by the design stance (p. 17), whereby we explain the "actions" of the clock by citing its designed internal workings. The design stance is a better (read: "more predictively accurate") stance than the intentional stance in this case.

[8] The issue is not that the design stance cannot accommodate evolutionary explanations that cite evolutionary adaptation as the relevant "design" in question. In fact, as Dennett notes, it is well suited for these types of explanations. The pressure that pushes us toward the intentional stance in cases such as that of the chimpanzee is the issue of the complex changes in behavior that the chimpanzee produces.

[9] We intend these criticisms to apply directly to the case of using the intentional stance to show that chimpanzees are agents. It goes without saying that there is a large body of literature on Dennett's view,

tional stance seems to fall prey to circularity. Dennett's view allows us to side-step metaphysical worries about whether individuals are *in fact* intentional agents who possess intentional states. Instead, we can simply adopt the intentional stance toward those creatures whose action is best explained by attributing these states to them. Recall that we adopt this stance when other explanatory stances, such as mechanistic ones, fail to do a better or comparable job of making sense of the behavior in question.[10] Thus, on Dennett's view we treat those creatures *as if* they were intentional agents; we do not assent to the claim that they are *in fact* intentional agents.[11] What's worse, his criterion of explanatory traction (whether it is simplicity or some other criterion) requires that we must already have an idea of whether the creature in question *might in fact be* an intentional agent and how being an intentional agent would get expressed in action in order to be properly used.[12,13]

Dennett (1987, pp. 34–35) argues that this type of criticism of his view confuses two distinct kinds of empirical questions, one to which the intentional stance is committed and the other to which it need not be. The intentional stance can be used to establish that there is some objectively verifiable pattern in the world that we think counts as the patterns of intentional agents. The existence and nature of the neural mechanisms by which these patterns get produced are independent empirical questions which, Dennett suggests, are not independently established by the intentional stance. What's more, one could, he argues, deny the existence of these mechanisms while accepting the intentional stance.

But even if Dennett is right that the intentional stance commits one only to the existence of *some* objective pattern in the world, this commitment itself is tantamount to the commitment that we attributed to his view above—namely, that it must already have

Footnote 9 continued
and, in this regard, our criticisms should be taken in the limited context of the question under present consideration.

[10] Cases where, say, the intentional stance does no better but no worse than mechanistic explanations present a puzzle for this type of view. Presumably, empirical observation would be the only way to resolve the impasse, further supporting one of our central objections (see below).

[11] Even if one were satisfied with this relatively unstable foundation on which to base one's attribution of agency to chimpanzees, there is a second, more significant problem with the commitments on which the intentional stance depends. Contrary to its stated aims, the intentional stance ends up making commitments about the existence of the states, whether brain states or mental states, that it proposes to attribute to possible agents in the "as if" mode. The reason is clear: the intentional stance relies on the criterion of predictive success (or some other view about how science works). Thus, if you think that the stance has been predictively successful, then you are left thinking that there is some causally efficacious set of states doing the relevant causal work. See Fodor and Lepore (1993, p. 76) for a related criticism. See Dennett (1987, pp. 34–35) for a response to this criticism.

[12] See Fodor and Lepore (1993, p. 76) for a criticism of the claim that an epistemic position—that of the intentional stance (or any stance, for that matter)—can either "make facts" or make them disappear. See also Andrews (2000, p. 19-ff) for a discussion of how this problem applies to using the intentional stance in theory of mind debates about non-human primates.

[13] See Sober (2005, pp. 93–96) for a defense of the claim that Morgan's canon can, for some evolutionarily derived similar traits between humans and non-human animals, license anthropomorphic conclusions as the most parsimonious. Povinelli and colleagues argue that such an approach is unwarranted (Povinelli and Giambrone 1999; Povinelli 2003).

commitments to what counts as actual instances of intentional action.[14] Even under this interpretation of the intentional stance's upshot, it continues to be the case that the intentional stance must make more burdensome commitments regarding the nature of chimpanzees' capacities that enable them to act as agents. If this is correct, then someone who favors the intentional stance must confront the question of whether chimpanzees have the states that can do the sufficiently complex work required to reliably produce behavior of the sort that an agent undertakes—even if we grant that they will never possess the analogous mental states that do the same work in human beings.[15]

The most general statement of our second criticism is that the intentional stance must be committed to the existence of some type of capacities that do the causal work of generating chimpanzees and other intentional agents' actions. If true, the intentional stance cannot—even under the most generous interpretation vis-à-vis the criticism leveled here—be used to deal with the hard problem that we hope to solve in this paper. Namely, it cannot provide an account of how chimpanzees are agents without making substantial commitments either to the possibility of non-conceptual content, or to controversial empirical claims regarding chimpanzees' capacities for second-order propositional thought.

If either of these two criticisms is correct, then Dennett's view (or some view like his) is ill suited for showing that chimpanzees are agents.[16]

2.2 A theory of merely purposive agency

Aside from the intentional stance, are there other available resources to show that chimpanzees like Megan are agents? Some philosophers of action such as Davidson (2001d) and Stoecker (2009), resolutely say 'no.' Many other philosophers of action may say 'yes,' given that they typically distinguish between what we might call "full-blown" agency and merely purposive agency (Alvarez 2009a, footnote 15; Bratman 2007a, p. 197; Butterfill 2001; Davidson 2001d; Glock 2009; Kalis 2011, p. 115; Korsgaard 2009, pp. 98–99, 174–175; Nussbaum 2006, p. 133; Steward 2009; Velleman 2000, pp. 124 and 189; Watson 2004, p. 136).

Full-blown agents have two sets of capacities. First, they have the more basic capacities to intend actions, to pursue goals and to aim at ends. In addition, they have robust capacities for rational deliberation, for second-order consideration of their first-order intentional states and for launching complex series of actions or plans. The capacity for second-order consideration of one's first-order intentional states (including reasons for action) is best understood as *explicit* reasoning about those intentional states or reasons that one understands as one's own. Naturally, 'second-order reasoning' and 'explicit reasoning' are not synonymous. Yet in the case of understanding the capacity

[14] Under this interpretation, the intentional stance would be indistinguishable from standard scientific practice, in which case the former is unnecessary.

[15] For a related discussion of whether chimpanzees need to possess isomorphic neural architecture in order to engage in the same kinds of reasoning as human intentional agents, see Godfrey-Smith's (2003) criticism of Hurley (2003a) and Hurley's response (2003b, pp. 275–276).

[16] We would go farther and claim that Dennett is wrong that the judgment to treat them as if they were intentional agents is exclusively a conceptual matter. We side with Millikan (2000, p. 65) that it is not.

for considering one's own reasons for action and intentional states, any second-order thought must also be explicit. The converse it also true (at least in this case): any explicit consideration of one's reasons for actions or one's intentional states must, by definition, be second-order. Otherwise, it is unclear how would one engage in explicit consideration of any kind.

Less than full-blown agents lack the latter set of capacities, but may possess some version of the former. Alternatively, they may possess a less complex version of the latter set along with a fully developed set of the former capacities. In general, though, these agents are merely capable of purposive agency such that they can aim at ends but they cannot evaluate them, rank them, or reason about them. In our parlance, they cannot "consider" them.

Put more generally, many philosophers of action who think that we have the resources to show that chimpanzees are agents, hold, minimally, two inconsistent views:

1. The hallmark of intentional agency is the capacity to act for reasons.
2. Chimpanzees, among other non-human animals, can be said to engage in intentional action by virtue of the fact that they can engage *only in* purposive action, which does not require the capacity to act for reasons.

These claims are mutually inconsistent. Thus, we must either reject one of the claims or modify one or both such that they are no longer inconsistent. The question is whether to substantially weaken (1), which would require lowering the bar for what counts as an agent, or to retain (1) in its present form and to drop (2). Since we are focusing on views that aim to show or are willing to accommodate non-human animal agency, the latter is not a viable option. Nonetheless, given that we have no reason to doubt that (1) is paradigmatic of agency, rejecting or modifying (1) will not be the route to showing that chimpanzees are agents. Thus, it looks as though we are at an impasse.

Even if one were willing to substantially weaken (1), additional problems would arise with (2). Given traditional ways of understanding intentional action, it will be difficult to form a view about merely purposive action that does not require—in the very least—the capacity to conceptualize a goal as such. Embedded in the idea of purposive action are the building blocks for an account of acting for reasons (Smith 1994). Thus, any worries about whether a chimpanzee like Megan has the capacity to form intentions for reasons that she takes to be her own, might undermine the claim that she is an agent because she can engage in purposive action.

But perhaps there is a way to modify what counts as purposive action and, thereby, to include Megan as engaging in this type of action. Butterfill (2001) argues that actions can be purposive even when they are what he calls "information based but unreflective actions." In this case, agents are not aware of the considerations that inform their actions, and yet the actions are intentional by virtue of being purposive. We might think that many non-human animals, including chimpanzees, would be capable of engaging in this kind of action and thus count as a type of agents.

One could interpret Butterfill's view as a promising route to understanding chimpanzee agency that relies on a deflationary view of what agency requires. We argue, however, that there are few good reasons to take this more radical route; rather, we will argue that careful consideration of what types of actions count as acting for reasons

will reveal ready-made conceptual resources for building an account of chimpanzee agency. We intend to show that there are creatures "act for reasons" even though they do not represent those reasons to themselves as such. For the moment, however, what is clear is that *the standard distinction between full-blown agency and merely purposive agency does not allow us to show that chimpanzees are agents*.

2.3 Comparative psychology's account of the pursuit of goals

Comparative psychologists routinely argue that many species are able to pursue the same goals by different means, and different goals by the same means, and over a century of empirical data would seem to support this conception (see overviews by Tomasello and Call 1997; Penn and Povinelli 2007; Povinelli and Penn 2011).

The problem with this approach's ability to show that chimpanzees are agents lies in its use of the term 'goal'. Many comparative psychologists fail to distinguish among different senses of what it means to pursue a goal (for a detailed discussion, see Penn and Povinelli 2013). On the one hand, an organism might pursue a goal that it does not take as its own goal. Satiation of hunger is an external goal insofar as the organism need not say to itself—"I am hungry and I desire to be satiated, so I will eat the banana"—in order to count as genuinely pursuing a goal. On the other, one can pursue a goal that one takes as its own, whereby an organism represents to itself a particular aim as one that it wants itself to pursue. One's goal of exercising regularly might be an example of the latter.

Given this distinction, one might think that only those creatures that are capable of the second sort of goal-pursuit count as genuine agents. If this claim is correct, then the attempt to use the comparative psychologist's data that chimpanzees pursue goals using various means as an indication that they are intentional agents will be unsuccessful. But even if this claim is false, the aim of this paper is to deal with the difficult (and empirically motivated case) wherein chimpanzees do not represent internal goals but are still agents. And as we have just seen, the comparative psychologist's resources are unable to provide the foundation for doing this.

3 Varieties of actions done for reasons

We begin our efforts to provide a positive account of chimpanzee agency by considering what are typically taken to be the hallmark(s) of agency. Our motivation here is to identify the *prima facie* reasons for including more varieties of actions in the set 'actions done for reasons'. Note that this account is not specific to chimpanzees and arguably applies to many different kinds of agents. It will, nonetheless, lay the foundation for showing that chimpanzees are indeed agents by identifying a kind of action done for reasons heretofore ignored by philosophers of action.

There have been numerous attempts to provide accounts of the capacities necessary and sufficient for agency and the conditions for a particular case of action to count as an intentional action. However, accounts that endorse Davidson (2001a, b) causal theory of action, as well as those that reject it, all agree on one aspect of his view: creatures that are agents are (among other things) capable of forming intentions to act on the

basis of reasons (Anscombe 1957; Brand 1984; Bratman 1999, pp. 26–29, 111; Korsgaard 2008, p. 13; Mele 1992; Paul 2009; Setiya 2007, pp. 30–36, but compare with p. 40; Velleman 2000, pp. 197–199).[17] This is also the case for so-called teleological theories of intentional action (Sehon 2005; Wilson 1989), volitional theories (Ginet 1990; Wallace 2006a, b) as well as those that take agents' guidance of (or relationship to) their actions to be essential for intentional action (Frankfurt 1978, 1998; Hornsby 2004; Velleman 2000, p. 199).[18]

Although not all views of intentional agency take the causal theory of action to be correct, they nonetheless take actions done for reasons, whether explanatory or justifying, to be the hallmark of intentional agency. Here these accounts understand 'reasons' to be subjective states rather than, as Dancy (1993; 2002, pp. 85–99, 103–107) and others (Kearns and Star 2009) argue, as facts of the matter about the action itself. We too understand reasons in the former fashion because we aim to understand how an agent, such as Megan, might count as having the *right relationship* to reasons for action without having any explicit, second-order relationship to them. More important, understanding reasons in this way does not represent a competing account of what constitutes *genuinely good, or normative, reasons*.

The fact that, for our purposes in this paper, we understand reasons in this way does not entail any substantive conclusions for the debate about what constitutes *genuinely good, or normative, reasons for action* (Alvarez 2009b; Dancy 1993, 2002; Kearns and Star 2009). This is because our account of reasons is intended to track agents' relationship to considerations in light of which they act, not how considerations justify or require certain courses of action. In this context, reasons must be subjective states given that agents must be capable of taking them as considerations or acting in light of them. Understanding reasons in this way *does not exclude* understanding *genuinely good reasons* as facts of the matter; rather, it simply underscores that for agents to act in light of reasons in the way that is paradigmatic of intentional agency, they must *take the form of subjective states*.

The standard Davidsonian picture of intentional action provides one picture of this hallmark. According to the causal theory, intentional actions exist in *causal* relationships with their antecedent mental events (Davidson 2001a, p. 12).[19] But what are those antecedent mental events? Davidson's (2001a, p. 4) version of the causal theory does not posit intentions as the cause of actions; rather, he contends that having a primary reason for an action is necessary and sufficient for it to be intentional. A "primary reason" comprises a pro attitude toward the action under consideration, a belief about it, or both (p. 4). A desire is a salient, although not an exclusive, example of a pro attitude.

[17] Naturally, these accounts do so on a variety of different grounds and, in some cases, understand what "acting for a reason" constitutes in diverging ways.

[18] Velleman's brand of constitutivism does not take acting for reasons to be the hallmark of agency or intentional action. Nonetheless, we do not think that Velleman would deny that acting for reasons is an important part of what pursuing this aim involves.

[19] Davidson (2001a, p. 12) notes that reasons (*qua* beliefs plus a pro-attitude toward a particular state of affairs), are typically understood as states but do not manifest themselves as states. Their "onslaught," as he puts it, *is* an event.

On this view just described, there is no requirement that agents understand their action as their own. But given worries about the problem of causal deviance, it is reasonable to think that any modified version of the causal theory of action will need to acknowledge that agents must take the reasons that cause their actions to be those that they endorse *as their own* (for views on the problem of causal deviance, cf. Brand 1984, p. 23; Chisholm 1966, pp. 29–30; Frankfurt 1978; Mele and Moser 1994, pp. 41–43; Schlosser 2007, p. 189; Searle 1983, p. 82).[20] In these cases, endorsing reasons as one's own reasons for action is tantamount to taking an explicit, second-order relationship to those reasons *as reasons.*

Given this basic account, we propose to distinguish between two classes of actions—what we will call *reason-considered* and *reason-devoid*. An example of a reason-considered action is as follows: John decides to bake a cake, but, given that he is only familiar with making flan, decides that he needs a cookbook to guide him. He sets out the ingredients, but he realizes that he has not taken the flour down from its storage place on the top shelf of his kitchen cabinet. John climbs onto a step stool to reach the top shelf, but realizes that he needs an implement to push the flour bag from the shelf and into his outstretched hand. He can choose between his broken fly swatter and a long cane with a looped top. He believes (correctly) that the broken fly swatter will not exert the proper force on the flour, so he reaches for the cane and uses its hooked end to drag the bag of flour off the shelf. This is a case of reason-considered action because John has a belief about how to obtain the flour with the cane and a desire to do so, all the while recognizing that he aims to do so because he needs the flour to make the cake.

Reason-devoid cases—here we hesitate to call them actions given that they are not technically actions—are also familiar. When Sally spills the coffee or jumps when she hears a loud noise, her behavior is reason-devoid. This is so because she has no obvious belief-desire pair in light of which she acts, and she lacks the relevant intentional states to undertake the action in question. By 'reason-devoid,' however, we do not mean actions that are "done for no reason whatsoever" but are still intentionally undertaken. These types of action are best exemplified by Quinn's (1993) example of the person with the ken to turn on radios each time she enters a room that contains one. Whether there is such a category of action is a matter of debate (Alvarez 2009a, p. 298; Setiya 2007). It is also worth noting that reason-devoid actions do not include actions that are done for bad, but motivating reasons. Our use of the term 'reason-devoid' is intended to capture those actions that *are not done for reasons* in Davidson's sense of the phrase and, on these grounds, are not intentional actions.[21]

3.1 The puzzle, or a justification for reason-directed actions

On a relatively standard conception of what it means to act in light of a reason, the philosopher of action readily accepts the distinction between reason-considered and

[20] Even Davidson (2001c, pp. 78–80), in his discussion of the infamous climber case, notes that reasons must cause actions in the "right kind of way" in order to count as cases of intentional actions. One might suppose that he had in mind something like the condition that one takes one's reasons as one's own.

[21] We set aside the claim, as defended by Alvarez (2009a), that actions can be intentional and yet not done for reasons.

reason-devoid actions, even if she does not use this terminology to describe it. The problem, however, is that the frequency with which we engage in reason-considered actions is an open, empirical question. We think highly familiar examples show*both* that we do not engage in reason-considered action much of the time *and* also that there is a grey area—namely, instances of actions that are neither reason-considered nor reason-devoid.

Consider a relatively familiar example—that of automatic driving. You get in the car and begin to drive. Five minutes later you find yourself in front of your house or apartment, but you cannot fully recall how you got there.

This does not seem to be a case of a reflex or an accident, and yet it is also unlike the case of full-blown, reason-considered action.[22] If it were merely reflexive or "mindless," we would not be able to explain how automatic drivers end up at home without hitting a pedestrian, running a red light, among other things. In this sense, (at least some) automatic drivers can respond to novel situations. They are relatively flexible, even while being unable to recall the intentional process by which one got oneself home. Note that this lack of recall does not suggest that automatic drivers have no episodic memory of how they arrived at home. What they are unable to do it to reconstruct the *intentional acts* that got them there. Automatic driving is paradigmatic of an action that lies somewhere between reason-considered and reason-devoid actions. How, then, are we to classify these kinds of cases?

If we are right, there is thus a third class of actions that are neither reason-considered nor reason-devoid. But the standard view of intentional action, which we will hereafter call the *highbrow view*, does not account for such familiar (and we believe, ubiquitous) examples.[23] These are what we will call *reason-directed* actions—actions done with the relevant belief/desire pair "in mind" (so to speak) but without the second-order, explicit relationship to those reasons.[24] Here 'relevant' simply means that the belief/desire pair matches up with the intended action (Smith 1994). So my desire for ice cream and my (true) belief that I have some in the freezer are fulfilled by (or serve as a reason in favor of) intending to open the freezer door and remove the ice cream to soften. While some scholars could find our division between reason-directed and reason-considered action artificial, this additional category covers much of human behavior that does not fit in either of the other two categories.[25]

[22] We are grateful to an anonymous referee for pushing us to consider this objection more carefully.

[23] On one reading of Davidson's account, it is possible to explain these cases (see Clarke 2010 for an example of such a defense, at least in the case of skilled activity). After all, you had the relevant belief-desire pair—namely, you wanted to get home and you believed that if you drove a certain route, you would arrive there—such that automatic driving does count as an intentional action. But, as is well known, reading Davidson's view in this way allows for the problem of causal deviance and requires as a possible, but problematic, solution that we add the condition that agents relate to their reasons in the right kind of way (Chisholm 1966, pp. 29–30; Frankfurt 1978; Mele and Moser 1994; Schlosser 2007, p. 189; Velleman 2000).

[24] Might the same action could be, in one case, reason-considered and, in another, reason-directed? Yes, in principle, because it is the agent's relationship to the action that matters here. There may, however, be limiting (even non-contentful) conditions on the kinds of actions that can be reason-directed.

[25] For some philosophers of action, 'reason-directed' action may sound close to what is often called 'intention in action,' or the intention that guides the action as we are engaging in it. On our view, reason-directed actions are those actions that bear the right relationship to considerations that function as reasons

To avoid overly metaphorical phrases like "in mind," let us outline what we take to be the necessary and sufficient conditions for an action to be reason-directed rather than reason-considered:

1. The action is not reason-devoid because the action is not best explained by reflexes possessed by the agent, such as a conditioned response or as an accident.
2. The action falls short of being reasonably classified as reason-considered, either because the agent in question lacks the abilities to engage in reason-considered action (as we will argue chimpanzees do) or, if the agent does possess the relevant capacities, she does not relate to her reasons in a second-order, explicit fashion.
3. The action is purposeful or has an end goal.
4. The action is best explained, against available alternatives outlined in (1), in terms of some belief or, for those not capable of propositional thought, belief-like state and some desire, or desire-like state, that the agent in these circumstances is likely to possess.

It may strike the reader that conditions (1), (2) and (4) are epistemic conditions for *knowing* that an action is reason-directed rather than metaphysical conditions for what renders an action *itself* reason-directed. We do not disagree with this characterization. But there is a reason for including epistemic conditions in what is arguably a list concerned with enumerating features that reason-directed actions *themselves* possess. There is a long history in philosophy of action—best represented by Davidson's (2001a) own account but also exemplified by Anscombe (1957), contemporary casual theorists like Mele and Moser (1994), those who defend the teleological view like Sehon (2005), and others—of including epistemic conditions—or the conditions that we need to meet to justifiably classify some action as an intentional action—in an account of what makes an action intentional. And although the fact that there is such a long history is not a reason to include such conditions in our account of reason-directed action, it does underscore what such a reason might be. Namely, to explain the necessary and sufficient conditions for reason-directed action, we must explain what causally efficacious states give rise to the action and those states *may not* be accessible to the agent herself. If so, then the conditions that are accessible to a neutral, third-person observer—what are arguably *epistemic* conditions—will stand in for a straightforward account of the metaphysical conditions. Here the metaphysical conditions that we have in mind are the casually efficacious states that bring about reason-directed actions.

Another worry that may strike the reader is that we use the language of 'beliefs' and 'belief-like' states, and similarly so for desires, in (4). Here one may object that we have simply stipulated the very claim that we are trying to prove—namely, that that there are beings who, regardless of their abilities to engage in propositional thought, engage in reason-directed action and are thus agents. We disagree. Condition (4) specifies that there are some kinds of causally efficacious states that are either beliefs or desires or that *function like* beliefs or desires with regard to the launching of an action. Here 'function like' means that they are inputs that the creature uses in his or

Footnote 25 continued
for the agent. Whether they ought be understood as prospective intentions or intentions-in-action is a further aspect of the capacity for intentional action that is beyond the scope of this paper.

her action-choice. The only thing that condition (4) makes clear, then, is how such states figure in a correct explanation of the action as a reason-directed as opposed to a reason-considered or reason-devoid action.

3.2 Is there really a puzzle?

At this point, we anticipate several possible objections to our argument.

The first, and most natural objection, is whether what we are calling reason-directed actions are simply those actions for which we *could* reflect on our reasons but we simply do not do so in the moment.. But, *ex hypothesi*, humans are not engaging in higher-order relational reasoning in the case of reason-directed actions. That is, we do not take considerations *as our reasons to act* prior to or in the context of acting. By contrast, we can imagine cases where we act in light of some form of higher-order reasoning but are reasons for acting are simply not apparent to us. For example, a cab driver in a busy city takes one cross-town route over another at rush hour. She might reason that she should take one cross-town street over another at rush hour, but when she launches the action of taking such a route her reason—say, that she can get her passenger to her destination more quickly—is largely implicit. Thus, the cases of reason-directed action are definitively different from those cases where the reasons in question are merely implicit.

But this raises a second objection. One might worry that there are very few cases besides automatic driving that count as reason-directed and thus that including an additional category of actions done for reasons on its basis is ontologically and metaphysically unjustified. Contrary to this claim, we contend that much of what we do on a daily basis is best described as reason-directed rather than reason-considered (see Sect. 3.2). Note that this assumption is contrary to the *highbrow view* of human action, and many of the assumptions that guide work in the philosophy of action. Recall that the *highbrow view* does not argue that most of our actions are intentional; rather, it simply claims that, among the behaviors in the set 'intentional actions,' the majority are (to use our terminology) *reason-considered*.

Even if one argues, *pace* our argument, that the majority of human cases of actions done for reasons are reason-considered actions, one must still explain the significant minority of cases, which are neither reason devoid nor reason-considered. This raises the question of how to explain automatic driving and absent-minded actions, the poor execution of plans like putting the milk in the cupboard, among many others (Amaya 2013; Mele and Moser 1994, pp. 41–44). As we have suggested, whether one finds these cases to be common (as we do) or uncommon (as someone else might), they still need to be explained and the only plausible way to do so is to introduce the category of reason-directed action.

But there is still a third pressing objection. Some may think there is no puzzle about these sorts of cases and, if so, that there is no need for introducing the category of reason-directed action. Brownstein (*forthcoming*) provides such an argument. He argues that a skilled, unreflective action— such as automatic driving—is an intentional action because the driver can answer Anscombean "why" questions about the reasons for what she is doing. Bearing in mind Anscombe's (1957) claim that the ability to answer these questions is the hallmark of intentional action, Brownstein's argument

presents a significant challenge to our claim that automatic driving and other similar cases fall short of the requirements for reason-considered (or full-blown) intentional action.

Is Brownstein right? We think his argument has shown only that we ought to be more permissive when using the term "actions done for reasons." Specifically, his view shows that we ought include actions that are done for reasons explicitly (and thus from a second-order perspective) *and* implicitly (but potentially subject to explicit, second-order consideration) While this claim may be important for understanding human action, it does not, we think, provide a foundation for showing that chimpanzees (and other non-linguistic creatures) are agents. Moreover, *human actions,* as we will show below, are not exhausted by simply enumerating those cases for which there are explicit, second-order reasons and those cases for which the reasons are implicit.

3.3 Why we need the category of reason-directed action

With Brownstein's objection addressed, we now turn to enumerating general reasons for thinking that there is a genuine puzzle over the cases discussed above and that they are not fully explained by the currently available categories of action available in action theory. Hence the need for the category of reason-directed action.

First, consider Railton's (2009) discussion of what he calls "fluent agency". Railton argues that genuinely skilled actions, such as playing the piano, typing and, perhaps,[26] automatic driving, are examples of what he calls "fluent agency." But he goes on to make a stronger claim, namely that "*all* action—including in particular paradigmatic premeditated intentional action—has *and must have* unpremeditated action as its source and core. A corollary: Most of the reasons for which we act, and that give us the name of rational beings are not made effective by 'choosing one's reasons"' (p. 102). If Railton is right, then automatic driving is merely less obviously fluent than considering, say, what we want to write in this paper, but both have the same initial set of "unpremeditated" actions at their respective cores. Although Railton's view does not directly help us to make the case that chimpanzees are agents, it does expose why the standard view that there are only reason-considered and reason-devoid actions is flawed because it has no easy way of categorizing these kind of cases.

With this point in mind, let us return to the case of automatic driving, as we described it earlier. Our purpose in using this example was to show that, on a relatively standard conception of what it means to act for a reason, automatic driving does not fit squarely into either the category of reason-considered or in the category of reason-devoid action.

To be fair, there is a competing explanation of these cases, which denies that the cases that we have used *are any different from the standard cases of reason-considered action*. On this view, automatic driving *is* actually a form of reason-considered action and thus there is no real problem with the standard distinction between reason-considered and reason-devoid actions. To give voice to such an objection, let us

[26] Note that Railton (2009, p. 96) takes skilled driving, rather than automatic driving, to be a relevant example. It seems reasonable to think that, barring luck, one must be a skilled driver to engage in automatic driving in the first place.

consider Velleman's (1985, p. 33) description of a roughly analogous case of suddenly finding yourself strolling up 5th Avenue without a clear idea of why you are doing so. In contrast to our account, he takes such cases to illuminate the features of what we are calling reason-considered action and, by extension, to be versions of reason-considered actions themselves. If Velleman is right, then such cases count as intentional actions and the category of reason-directed actions is unnecessary. Velleman's view can explain, first, why these actions are patently intentional and thus do not require an additional category of intentional actions to count as such. If correct, his view might provide an easy route for defending the claim that chimpanzees are agents. However, we believe his view is flawed.

But let us examine Velleman's view in more detail. He argues that hypothetical cases in which one does not know what one is doing—as in familiar cases where one suddenly finds oneself putting milk in the cupboard rather than in the refrigerator—provide one with *the opportunity* to identify the reasons for what one is doing. "[Y]our desire to understand what you're doing restrains you from going on with behavior once you realize that you don't know your motives for it," (Velleman 1989, p. 32). Thus, argues Velleman, we usually undertake actions that we do in fact understand, even if we do not necessarily have reasons in mind when we initially launch the action. If we have undertaken an action, then we have knowingly done so.

In this regard, a necessary condition for a series of bodily movements or the like to count as an intentional action is that they are constituted by a "subagential" aim (Velleman 2000, p. 191). The subagential aim, Velleman argues, is conscious control of the action in question, and one cannot have the aim of conscious control unless one has a second-order aim *to aim at it* (pp. 192–193). Simply put, what makes action genuinely intentional is that I, as an agent, constitutively aim to have control over the action as my own (p. 193). But since the aim is constitutive, I need not have any second-order attitudes toward my action before I am actually engaged in it; rather, on this view, the fact that I aim at controlling my action simply by virtue of launching an action is necessary and sufficient for the action to be an intentional one.

In response to Velleman's view and its implicit challenge to our defense of the need for the category of reason-directed action, it is useful to consider two facts. First, it is not clear that, even if Velleman is correct, the type of action that he is describing covers actions performed by creatures other than fully developed, adult human beings or, at least, creatures with the capacity for complex, second-order thought. If this describes adult human intentional action rather than intentional action in general, then his account does not raise a genuine alternative to our view—albeit a slightly modified version of it. Namely, there are some cases of action that are neither reason-considered nor reason-devoid, and they ought still be considered intentional *when certain kinds of creatures other than adult human beings* perform them.

We do not think, however, that this modification is necessary. Velleman's account does not rule out the possibility that there are more and less full-blown instances of intentional action under his description. He acknowledges that cases like finding oneself strolling up 5th Avenue without a clue as to why one originally set out on the walk is not a paradigmatic instance of intentional action. Instead, he uses these examples to show that we aim to know something about our particular aims when we engage in intentional action and that we aim to exert control in so doing. But this does

not show that the 5th Avenue stroll is ontologically equivalent to someone reaching for flour on the top shelf to bake a cake. In fact, his claim is consistent with our claim that these two cases share a family resemblance—they are both, in some sense, done for reasons and that they are thus intentional, albeit in reasonably different ways and to different degrees. So Velleman's account is not a challenge to our view; rather, it is at least consistent with it and, read more strongly, it provides reasons for thinking that we are on the right track.

With this said, the aim of this paper is not to provide a lengthy defense of the concept of reason-directed action.[27] We present it and respond to challenges here because a view like Velleman's would challenge the justification for the introduction of the action category "reason-directed", and, thereby, weaken the foundations for the argument that chimpanzees are agents.

Velleman's view aside, the burden is still on us to show that there are indeed cases of action that are best described as reason-directed. All we shown in responding to his view is that it would be wrong to classify these actions as reason-considered.

3.4 Why most of our actions are reason-directed

The puzzle about reason-directed actions is not unique to the specific cases discussed in the previous section. Even if it were, this would not undermine our argument for a class of actions best described as reason-directed. It would, however, undercut our attempt to use this argument to show that chimpanzees are agents because they engage in a form of agency—exemplified by reason-directed action—shared with full-blown agents. So are there other kinds of cases that would count as reason-directed rather than reason-considered action?

We offer the following, non-exhaustive list to provide a sense of the ubiquity of reason-directed cases in everyday human life:

- Stirring soup, which one knows how to make, while it is cooking on the stove
- Making whipped cream by hand, which requires whipping cream and granulated sugar with a whisk until stiff peaks form
- Various instances of multi-tasking
- Window shopping
- Watering a large flower bed in a garden
- Petting a dog that one knows well and/or with whom one is comfortable
- Most skilled actions, such as piano-playing, horseback riding, ice skating, jumping rope, among others
- Carrying a light package in one's arms or a briefcase in one's hand
- Walking a well-behaved dog or leading a docile horse to pasture
- Going for a stroll through a park
- Reading this sentence

The list illustrates the many actions we perform without considering or evaluating the reasons for doing them. Note that our list covers a wide range of types of cases,

[27] We do so in our (ms) "Two Ways of Acting for Reasons".

from skilled actions to what we might call "mindless" ones such as stirring soup, or shuffling in place while waiting for a play to happen in a soccer match.

One could argue that it is counterfactually true of many, if not all, of the examples in our list that we *could* consider or evaluate the reasons that direct us. Thus, one may claim, it is this fact that makes them intentional actions and not that they are reason-directed.

Earlier we provided a *prima facie* case for answering this objection. But now that we have a list of examples, let us return to this objection.

The examples on our list are those that, in fact, we tend to do without engaging in consideration of our reasons. But the purpose of our list is not to suggest that all reason-directed actions are those toward which we lack directed attention. Many actions can be either reason-considered or reason-directed, but our relationship to them (or lack thereof) is what determines whether a particular instance of an action is one or the other. We suggest that, much of the time, these actions *are in fact* reason-directed, even though they may sometimes be reason-considered.

But our point is not that the same actions will be reason-directed for both the fully developed human being and the chimpanzee. Rather, it is to show that human beings engage in reason-directed action and that, in these cases, they are acting as agents. This sets the stage for our argument in Sect. 4, where we will argue that chimpanzees such as Megan that are able to engage in similar sorts of actions will also count as kinds of agents. This is not, however, an argument by analogy; rather, we will show that the independent case for the category of reason-directed action sets the foundation for the case for chimpanzee agency. For the moment, we are simply building the foundations for the argument that reason-directed action is genuinely action.

3.5 Mixed cases

Thus far, we have maintained an analytical distinction between reason-directed and reason-considered actions. But what we have been calling reason-directed actions might be best understood as a string of intermediate (or more basic) actions that agents undertake in the process of undertaking an action toward which their attention is directed.[28] If this is right, then it would seem that it is best to classify most actions as mixed cases—namely, as some combination of reason-directed and reason-considered actions. What's more, we suggest that many of our everyday actions are nonetheless closer to reason-directed actions than they are to reason-considered actions.[29] As Railton (2009, p. 83) notes, if most of our actions were in the latter group, we would be treading dangerously close to an implausible, quasi-homuncular view of how agents relate to their actions.

To see why this is the case, let us return to the case of John the sometimes-baker taking the flour off the shelf with the cane. It is possible that John is engaging in reason-considered action in this case. But if we are right, it is more likely that he is

[28] This claim is exemplified by the distinction between basic and non-basic (or complex) actions.

[29] We take it that Railton (2009, p. 102) has something like this claim in mind when he argues that all actions have an "unpremeditated core."

not—he is simply scanning the environment for tools available to achieve the end of getting the flour off the shelf. If so, then he would be engaging in reason-directed action rather than reason-considered action. Or, perhaps more accurately, he is engaging in an action that is some combination of reason-directed action (pulling the flour off the shelf with the cane) and reason-considered action (attempting to bake a cake when he has never done so).

Aside from understanding John's action in the right way, the argument in defense of reason-directed action as genuine action sets the stage for the argument that chimpanzees are agents of kind. If reason-directed action were the evolutionarily ideal case of executing goal-directed action, then we wouldn't want to side with the *highbrow view* and categorize it as deficient, or worse yet, as a non-action.

4 Chimps as secret agents

We have now made the general case for the existence of reason-directed action. How does this argument help to make space for chimpanzee agency?

4.1 Goal-directed behavior as reason-directed action among chimpanzees

Many species, including chimpanzees, seem plainly capable of goal-directed behavior (Tolman 1932; Povinelli 2003, 2012; Tomasello and Call 1997). Behavior is goal-directed when the behavior is done in the pursuit of some end, perceived or actual (Tolman 1932). But what makes goal-directed behavior relevant for our purposes? Goal-directed behavior is genuinely action insofar as it is neither accidental nor reflexive and yet does not require full-blown agency for it to be pursued. At the same time, the traditional definition of goal-directed behavior does not, we think, cover the capacities necessary and sufficient for engaging in this kind of action. On the account above, we end up being forced to claim that even protozoa engage in goal-directed behavior. In light of this worry, we suggest that the necessary and sufficient *abilities* that a creature must possess to engage in *goal-directed behavior* are as follows:

1. The ability to form preferences and to engage in some ranking of these preferences.
2. The ability to track and to respond to facts (including merely perceived facts) in the world that are relevant for the pursuit of the goal in question.
3. The ability to pursue the preference by exerting control over one's body and directing one's movements toward the end in question.

Note here that goal-directed behavior is not, itself, reason-directed action, but we will argue that chimpanzees' ability to engage in goal-directed behavior *is a case example* of their ability to engage in reason-directed action. Before considering how chimpanzees' capacity for goal-directed behavior suggests that they are agents, let us consider each of (1)–(3).

It stands to reason that (1)–(3) can, as abilities, be expressed in more and less complex ways and that these expressions are dependent on the complexity of the creature who possesses them. Since our aim is to show that chimpanzees possess these abilities even given their more modest cognitive abilities (in comparison to fully-

developed, adult human beings), we will detail (1)–(3) in their minimal expression. We already possess a straightforward picture of these abilities in their more developed, human forms. But because we lack a clear picture of their more minimal forms, we thus cannot say whether creatures less complex than us possess them.

Possessing (1) represents the ability, minimally, to pick out an end to pursue and to choose to pursue it rather than others that one might otherwise prefer to pursue. The ends in this case are identified by desire-like states. Crucially, being able to rank one's preferences need not mean that one has a second-order relationship with the preferences. Minimally, one can rank one's preferences just in case one opts to pursue one rather than another when they cannot be simultaneously pursued, and one could have pursued another preference on the list in similar circumstances. Furthermore, ranking *need not* require that one has a specific metric in mind that one applies to one's preferences; rather, it requires only that one has a set of preferences, some of which one pursues over others when one has the choice between them. While this is a deflationary view of ranking preferences, it does not exclude the more complex, second-order forms of ranking in which robustly rational agents engage. It merely highlights that the ability to rank preferences need not manifest itself in this form.

Ability (2) above is relatively straightforward. It describes the ability to form some basic representation of how a course of action will satisfy the preference that one is aiming to satisfy. A second aspect of possessing (2) is possessing the sensory and cognitive apparatuses that allow one to identify relevant changes in the environment such that one can change course or adapt one's behavior in light of them at least some of the time (Bermúdez 2003, pp. 56–58). But note that it also makes room for the possibility that a creature engages in goal-directed behavior on the basis of false beliefs or merely perceived facts, as the case of Megan that we discuss below illustrates. Since (2) identifies an *ability that a creature must possess* rather than what must be true of an action for it to count as reason-directed, it would be strange to think that the ability to engage in fact-tracking must also include the ability to be mistaken about such facts. Instead, we typically think that such abilities get masked in such instances, which is consistent with the description that we provide above. Nonetheless, since we want to argue that abilities (1)–(3) *underwrite* chimpanzees' ability to engage in reason-directed action and thus renders them agents, we include the defeasibility condition that the chimpanzee can act in light of incorrect fact-tracking without this mistake counting against their status as agents.

By contrast, (3) above would seem to require that one is able to take a second-order attitude toward one's ends such that one can direct oneself toward realizing them. We contend, however, that a more minimal form of (3) requires only the ability to aim at some end without having any second-order propositional attitudes toward the end in question or toward the act of aiming at it. A creature need merely be able to pick out some end that she would prefer—say, a very ripe banana that can be obtained only by using a tool to pull it towards her—and attempt to obtain it.

Taken together, a creature is capable of goal-directed behavior when the creature minimally possesses desire-like states that the creature ranks. The creature then, by deploying some form of means–ends reasoning, can determine how to satisfy the desire as the preference she aims to pursue over others. Finally, the creature can deploy the

conclusions of the means–ends reasoning in action such that she aims to satisfy her desire.

Let's consider each of (1)–(3) with regard to the question of whether chimpanzees are capable of goal-directed behavior, where this capacity *is a case example* of their status as agents. Naturally, other abilities that chimpanzees have may also allow them to engage in reason-directed action. Given that the question of the abilities that they have is empirically open, we assert that the *best exemplar of their abilities as agents is their ability to engage in goal-directed action*.[30]

We take it as axiomatic that chimpanzees possess the sensory, perceptual and cognitive apparatuses to pursue goals (see Penn and Povinelli 2013). We suggest that such goal-directed behavior is reason-directed action in chimpanzees. Although it is possible that chimpanzees have beliefs and desires, recall that we wish to deal with the difficult case in which they do not. (1) and (2) above show us how they have less complex forms of these attitudes such that we can plausibly say that they could "possess" reasons.

Consider the following illustration of what we take to be a chimpanzee engaging goal-directed, and thus in reason-directed, action (for details, see Povinelli 2012, Chap. 3). Megan has been trained to lift a small, somewhat heavy box (box weight = 7.0 kg) and place it on a short platform to her left. After she does so, she receives a slice of banana. She is given four trials of this type. In between the setup of each trial, Megan waits outside where she cannot see the situation. After the fourth trial, we switch to a *visually identical but much lighter box* (box weight = 1.5 kg). Again, we administer four trials of this type.

In Fig. 1 we present the data concerning the maximum height Megan lifts the box on each trial. Notice that on the first trial when the weight is switched, Megan makes the same "error" you or I would make: she lifts the box much higher than she intends (Povinelli 2012). We know this because of the height of her subsequent lifts—they gradually decline as she adjusts to the new weight.

Let us examine the critical first trial after the switch in greater detail (Fig. 2a, b). In Fig. 2a, Megan is reaching and almost grasping the box. To us, this indicates that she is aiming at lifting it. In Fig. 2b, c, Megan begins lifting the box, but applying too much force and her arm and the box swing upwards. This "error" is the result of an experimental manipulation in the information provided to Megan In Fig. 2d, she places the box correctly (having recovered for the overshoot), indicating that she has returned to the pursuit of her original goal.

In the above illustration, it is clear that Megan is engaging in a form of reason-directed action, although she gets the execution of the action wrong insofar as she overshoots when she lifts the box. But her error in fact-tracking need not indicate that she is failing to engage in, first, goal-directed action and, second, reason-directed action. She aims to lift the box, she directs her body toward doing so, and she is surprised at how high she lifts the box given what she apparently took to be an appropriate

[30] Saidel (2009, pp. 35–36) also argues that goal-directed behavior is, for some non-human animals, an indication of their capacity to engage in genuine actions. However, Saidel suggests that this status provides evidence for concluding that non-human animals have genuine beliefs and desires, which is exactly the controversial content we seek to avoid.

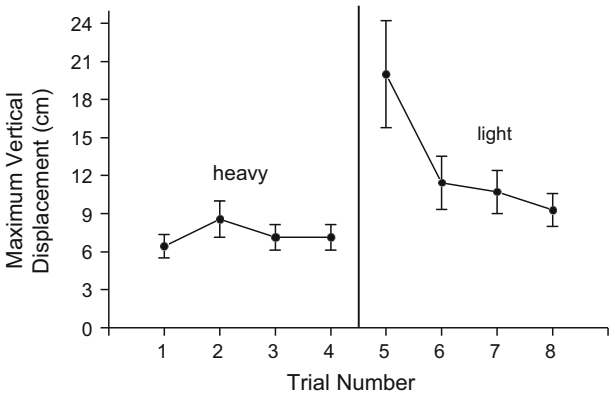

Fig. 1 Data from a group of seven chimpanzees who lifted a small *heavy* box for four trials and then lifted an unexpectedly *light* box for four additional trials. Note how on their initial trials with the *light* box they lift the box higher than they intend (compare trials 5–6 to trials 7–8). (Data from Povinelli 2012)

exertion of force to lift the box. In fact, human subjects make the same performance error (Povinelli 2012), underscoring what we have shown above regarding the complexity of human agency and the need for what we have called *mixed cases*.

The manifest evidence that Megan did not expect this outcome, indicates why we should understand this case of goal-directed behavior as an instance of reason-directed action. First, it indicates that she took her plan for action to have gone differently than what she was aiming to do—namely, to lift the box and for the way that she lifted it to occur in just the same way as it had in the past. Second, it suggests that she did in fact have an aim—that of lifting the box and setting it down.

It would be difficult to say that Megan had this aim and yet lacked any reasons for doing what she did. The trouble, as discussed above, is to show how she is capable of acting for reasons without the relevant capacities for second-order thought or propositional thought more generally. But if our account is correct, her actions were directed by reasons, but she lacks the relevant capacities for second-order thought, or propositional thought more generally.

4.2 What's so difficult about establishing chimpanzees' agency?

Others have attempted to show that creatures that lack the capacity for propositional thought can be understood as agents or as capable of intentional action. These other approaches reflect an interest in developing an account of non-human animal rationality (Bermúdez 2003; Carruthers 2005; Dreyfus 2007; Glock 2009; Hurley 2003a, b; Jamieson 2009; Lurz 2003; Saidel 2009), out of which will hopefully fall a theory of non-human animal agency. However, we believe our approach has a number of advantages over this approach. Considering these advantages in direct comparison to some of these other approaches will help to underscore the difficulty in showing that chimpanzees are agents without appealing to capacities chimpanzees are unlikely to possess.

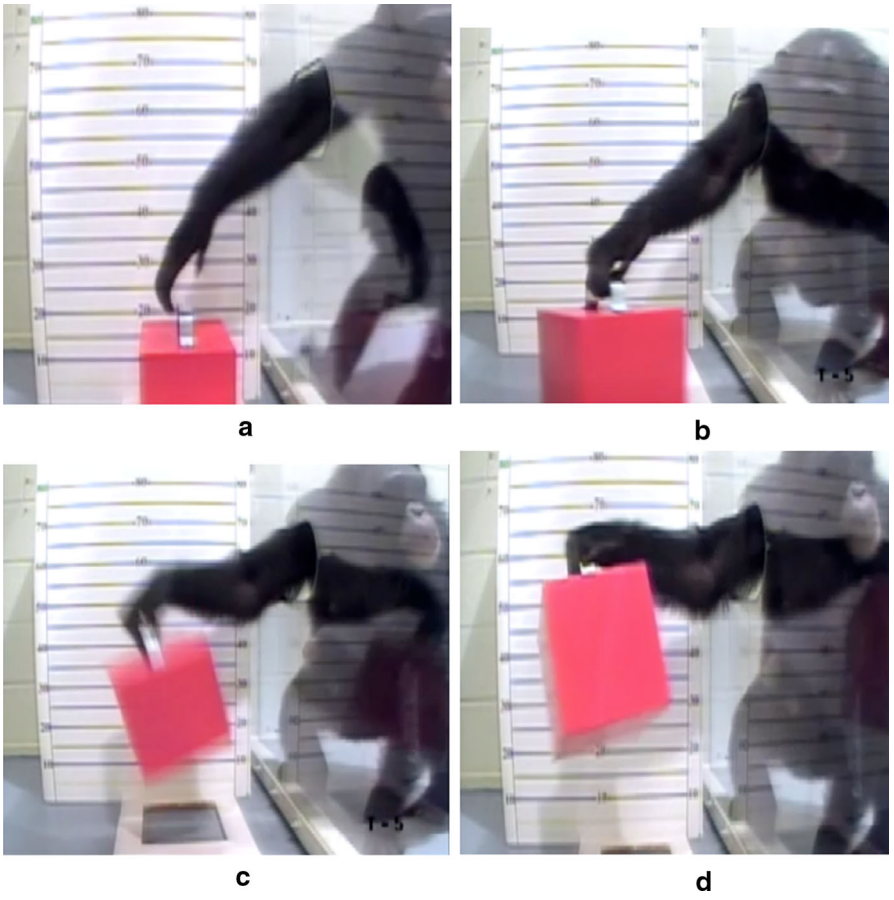

Fig. 2 A chimpanzee (Megan) intentionally reaches and lifts a box that is much lighter than she expects based on previous lifts. Because the weight of the box has been manipulated, she lifts it much higher than she intends (see data in Fig. 2)

Bermúdez (2003, 2006, 2009) provides an account non-linguistic creatures' ability to engage in intentional action on the basis of a theory about their ability to reason. He argues that if we are able to provide a coherent psychological explanation of a creature's behaviors, then we are warranted in judging that they are agents of a sort (Bermúdez 2003, pp. 128–129). On his view, providing a psychological explanation of a non-linguistic creature's behaviors is to provide an explanation of their behavior in terms of their non-linguistic rational capacities. These capacities, Bermúdez argues, provide "representations of contingencies between actions and their outcomes" such that creatures can put together their beliefs about their environments and their desires in the right kind of way and thereby engage in rational action (p. 129). But, adds Bermúdez, it would be overly demanding to require that the beliefs or desires be explicit to the animal in question—they are often "immediately perceptually manifest" (p. 129).

The possibility of providing such an explanation, Bermúdez argues, is determined by the insufficiency of environmental explanations: "[a]n action requires psychological explanation just if its occurrence could not have been predicted solely from knowledge of the environmental parameters and sensory input," (p. 130). These are cases where one cannot cite some law-like relationship among environmental factors, the creature's sensory input and its behavioral output (pp. 130–131). If one must cite some kind of explicit or implicit instrumental reasoning to link these two together, Bermúdez argues, then it is the case that we can attribute the capacity for rationality (of varying degrees of complexity) to the creature and thus the ability to put beliefs and desires together in order to act upon them. As he notes, however, the type of rationality available to non-linguistic creatures is not expressed in the standard practical syllogism that we associate with practical reasoning nor is it best described in terms of decision theoretic utility calculations (pp. 132, 135). Instead, the probabilities of the various outcomes of different action-choices are immediately perceptually manifest such that non-human animals make rational decisions without engaging in traditional forms of reasoning (p. 135).

While Bermúdez may be right that the grounds for attributing the capacity for reasoning to creatures may depend on the negative criterion above (particularly those that are non-linguistic), there are two reasonable questions to ask of his view. First, is his account of non-linguistic practical rationality sufficient for explaining how non-linguistic creatures act as intentional agents? It would seem that the answer to this question is 'yes,' given that the hallmark of intentional action is, as we have granted, acting for reasons. If he is right that the relevant belief-like and desire-like states are immediately perceptually manifest such that they constitute reasons, then it may be the case that we can count his view as establishing that non-linguistic, non-human animals *are* agents. But, on a narrower interpretation of his argument, Bermúdez has merely provided a defeasible evidentiary basis for producing agential explanations of non-human animals' behaviors. He has not, however, shown that we are justified in concluding that those animals are actually agents. To do so, he would need to show (1) how the capacities for reasoning that he discusses are capacities that the creatures actually possess and (2) why they are sufficient for agency.

Bermúdez makes two moves that we would like to avoid. First, his view is based on significant commitments to the possibility of non-conceptual content. We would like to be able to show that they are agents without saddling ourselves with significant commitments about the possibility of non-conceptual reasoning. Our account has the advantage of being able to explain action that is informed by reasons without requiring that the creature in question reasons about the action in question. Second, he uses a view of reasoning to justify *explaining* non-human animals' behavior in agential terms. We would like to be able to show that chimpanzees *are* agents, and we believe that our view more directly does so.

A related problem with Bermúdez's view—at least with regard to the question of whether chimpanzees are agents—is that it uses the need for explanatory completeness to determine the agential status of the creature whose behavior is being explained. Given worries discussed earlier about Dennett's use of this criterion, one might think that it is an unsteady foundation on which to base an argument that shows that chimpanzees are agents. But even if these worries leave our reader unfazed, one might think

that the answer from explanatory completeness does not fully answer the question that we have set out to answer in this paper. Namely, do chimpanzees actually possess the capacities to act as intentional agents and, if so, are they more or less complex agents?

Our account directly addresses this question. First, we provide an independent argument as to why there is such a thing as reason-directed behavior in creatures that we know to be agents—namely, fully developed, adult human beings. From there, we argue that chimpanzees possess the capacities to engage in reason-directed behavior as evidenced by their capacity for goal-directed behavior. While Bermúdez's account may explain how chimpanzees engage in instrumental reasoning in these cases, it cannot, alone, provide the foundation for showing that they are intentional agents.

Our approach also stands in contrast to arguments offered by Hurley (2003a). Hurley argues that non-human animals are capable of having what she calls "non-conceptual reasons" for action, where the reasons in question are attributed to the creature on the basis of its conceptual abilities rather than on the basis of its capacity for possessing second-order attitudes (pp. 232–233).[31] Hurley takes these conceptual abilities, following Tomasello and Call (1997), to be those that track and utilize the flexible relationship between ends and means (p. 237).[32] Non-human animals, including non-human primates, possess the ability to pick out different means to the same end and different ends that might be pursued by the same means. But Hurley suggests that this ability is necessary, but not sufficient, for attributing reasons to non-linguistic creatures such as chimpanzees. The conceptual abilities that are additionally necessary to count as an intentional agents are those that allow creatures to "decompose, transfer and recombine" the relevant elements from one means–ends context to another (p. 239). Only then, Hurley argues, can we attribute reasons to non-linguistic intentional creatures (p. 239).

Although it disavows the criterion of the ability to hold second-order attitudes toward one's reasons, Hurley's view seems to sneak it back in by focusing on capacities that do the same kind of work. The capacities that we discussed in the previous section are focused on the ability to aim towards a goal, which is, on our view, sufficient for attributing to the chimpanzee a reason to act. But this attribution does not require that the chimpanzee has a relationship to that reason *as a reason* for her action to be intentional or for her to be an agent. By contrast, Hurley's view seems to require that this be the case, given her claim that acting for "non-conceptual reasons" is to act for considerations that one can move from one context to the next.

Oddly, Hurley's view might force us to conclude that chimpanzees are frequently not acting as agents at all. Consider Megan's abilities to use sticks of to obtain food across various experimental contexts. Povinelli (2012) shows that Megan can learn (and become proficient at) at using a simple hooked stick to snag a looped platform containing a banana. Nonetheless she does not extrapolate to other, seemingly comparable situations, in which the hook must articulate around a simple post to pull a platform. Nor does she generalize to a stick that has one deformed end and one functional end; instead, she continued to use the deformed end even after it was apparent

[31] See Sterelny (2003, p. 259) for a challenge to Hurley's so-called "interpretativist" strategy.

[32] Compare with Bermúdez (2003, esp. Sect. 3.5).

that it could not exert the necessary force on the banana in order to drag it toward her. Even more striking, consider how expert tool-using chimpanzees like Megan, who have spent a decade learning many separate problems of this kind, now treat the seemingly familiar problem of using a hook stick to retrieve an out-of-reach banana. Here they are given two hook sticks. One of the sticks has a normal, rigid hooked end. The hooked end of the other stick is spring-loaded (it uselessly deforms against another object). Initially they are presented with both sticks laid out in the correct orientation with two bananas nearby. Here, the chimpanzees learn to abandon to quickly abandon the spring-loaded hook after they see it deform against the banana. But when the situation is just slight altered, and they are required to choose one of the tools and carry it over to the banana retrieval context, they are perfectly content to select the incorrect tool—even when the spring-loaded end deforms in their hands as they are selecting it (see Povinelli and Frey, in review). These and related data suggest that although chimpanzees and other animals can exhibit some limited perceptual generalization across similar kinds of contexts, they cannot identify the reasons under which they act nor can they transfer those specific reasons as such into other similar contexts (see Penn et al. 2008). Crucially, if Hurley's view were correct, we would be forced to conclude that chimpanzees, in the cases just detailed, are not acting as agents because they are not able to transfer reasons from one stick-using context to another.[33]

But we find such a conclusion inconsistent with two facts about these cases. First, the cases where chimpanzees *are* successful seem to indicate that they are successfully pursuing the goal of obtaining the banana and thus acting in a way that is directed by reasons even though they may not be able to fully transfer those reasons to other contexts. Second, the cases where the chimpanzees *fail* to use the sticks appropriately are, nonetheless, instances of the chimpanzees trying to pursue the goal of obtaining the banana. They simply lack some of the cognitive capacities to do so effectively. This latter fact, however, does not license the inference that they are not agents, nor does it license the inference that they are not, in our phrasing, engaging in reason-directed action. By contrast, our view accommodates both of these facts.

5 Conclusion

Let us end by considering the commitments that our view, if correct, would force us to accept, both for a view of human agency and chimpanzee agency.

In the case of reason-directed action, our view commits us to the view that both humans and chimpanzees share a complex cognitive system that grounds "belief-like" and "desire-like" states whereas only humans have a cognitive system that grounds beliefs and desires. If what we have argued above is correct, human beings engage in reason-directed actions much more frequently than we would typically think (for related discussions, see Povinelli 2012). It follows further that humans frequently use the language of beliefs and desires to describe what, with respect to many of our actions, might be more accurately described as "belief-like" and "desire-like."

[33] There is even the empirically-motivated possibility that they cannot transfer reasons in this manner in any contexts (see Penn et al. 2008).

This raises the question—one beyond the scope of this paper—of *how* the human system for second-order propositional attitudes resculpted our own evolutionarily primitive, cognitive engines, which currently support the actions of less complex intentional agents such as chimpanzees.[34] We nonetheless wish to highlight two closely related answers to this question. On the one hand, the lower-order systems in both humans and chimpanzees might look identical except for the fact that the second-order human intentional system can translate those lower-order states into beliefs and desires where required. On the other hand, the human capacity for beliefs and desires may have been grafted into the lower order system such that it is difficult to cleave one off from the other.[35]

Aside from questions about how to explain the conclusions of our argument in evolutionary terms, our view entails two conceptual commitments about the nature of agency in general and human and chimpanzee agency in particular. First, there are instances of intentional agency—such at that exemplified by chimpanzees that do not require the presence of propositional attitudes. Second, humans engage in reason-directed action much more frequently than they engage in reason-considered action. Worse yet for the *highbrow view*, these are the evolutionarily prototypical cases. This entails that most human actions are neither deformed instances of intentional action nor idealized ruminations leading to reason-considered action. The calculus of the *highbrow view* does not operate over the bulk of human action.

Acknowledgments The writing of this paper was funded in part by a Grant from the University of Texas at El Paso's University Research Institute. We thank two anonymous referees for this journal. Caroline Arruda also thanks Gregory Nirshberg, who served as her research assistant while she prepared to write this paper.

References

Alvarez, M. (2009a). Acting intentionally and acting for a reason. *Inquiry*, *52*(3), 293–305.
Alvarez, M. (2009b). How many kinds of reasons? *Philosophical Explorations*, *12*(2), 181–193.
Amaya, S. (2013). Slips. *Noûs*, *47*(3), 559–576.
Andrews, K. (2000). Our understanding of other minds: Theory of mind and the intentional stance. *Journal of Consciousness Study*, *7*(7), 12–24.
Anscombe, G. E. M. (1957). *Intention*. Ithaca, NY: Cornell University Press.
Arruda, C., & Povinelli, D. J. (ms). Two ways of acting for reasons.
Beisecker, D. (1999). The importance of being erroneous: Prospects for animal intentionality. *Philosophical Topics*, *27*(1), 281–308.
Bermúdez, J. L. (2003). *Thinking without words*. Oxford, UK: Oxford University Press.
Bermúdez, J. L. (2006). Animal reasoning and proto-logic. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 127–138). Oxford, UK: Oxford University Press.
Bermúdez, J. L. (2009). Mindreading in the animal kingdom. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 145–164). Cambridge: Cambridge University Press.
Brand, M. (1984). *Intending and acting: Toward a naturalized action theory*. Cambridge, MA: MIT Press.

[34] For an extended discussion of the relationship between uniquely human and more ancestral forms of cognition, see Penn et al. 2008. Hurley (2003b) also discusses this issue in response to Godfrey-Smith's (2003) claim that she tries to derive an account of the shared architecture of human and non-human animal agents from an account of their roughly shared folk psychology.

[35] Steward (2009, p. 224) defends this claim, although on the grounds that what we are calling the lower-order system is a form of purposive agency.

Bratman, M. (1999). *Faces of intention: Selected essays on intention and agency*. New York: Cambridge University Press.
Bratman, M. (2007a). Planning agency, autonomous agency. In M. E. Bratman *Structures of agency: Essays* (pp. 197–232). Oxford, UK: Oxford University Press.
Bratman, M. E. (2007b). *Structures of agency: Essays*. Oxford, UK: Oxford University Press.
Brownstein, M. (Forthcoming). Rationalizing flow: Agency in skilled unreflective action. *Philosophical Studies*. doi:10.1007/s11098-013-0143-5.
Butterfill, S. A. (2001). Two kinds of purposive action. *European Journal of Philosophy*, *9*(2), 141–165.
Butterfill, S. A., & Apperly, I. (2013). How to construct a minimal theory of mind. *Mind and Language*, *28*(5), 606–637.
Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*, 187–192.
Carruthers, P. (1996). *Language, thought, and consciousness*. Cambridge, UK: Cambridge University Press.
Carruthers, P. (2005). Why the question of animal consciousness might not matter very much. *Philosophical Psychology*, *18*(1), 83–102.
Carruthers, P. (2013). Animal minds are real, (distinctively) human minds are not. *American Philosophical Quarterly*, *50*(3), 233–248.
Chisholm, R. (1966). Freedom and action. In K. Lehrer (Ed.), *Freedom and determinism*. New York: Random House.
Clark, A. (2005). Word, niche and super-niche: How language makes minds matter more. *Theoria*, *54*, 255–268.
Clarke, R. (2010). Opposing powers. *Philosophical Studies, 149*, 153–160.
Dancy, J. (1993). *Moral reasons*. Oxford, UK: Blackwell.
Dancy, J. (2002). *Practical reality*. Oxford, UK: Oxford University Press.
Davidson, D. (2001a). Actions, reasons, and causes. In D. Davidson *Essays on actions and events* (pp. 3–19). Oxford, UK: Clarendon. [Reprint of Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, *60*, 685–700].
Davidson, D. (2001b). Agency. In D. Davidson *Essays on actions and events* (pp. 43–62). Oxford, UK: Clarendon. [Reprint of Davidson, D. (1971). Agency. In R. Binkley, R. Bronaugh & A. Marras (Eds.), *Agent, action and reason*. Totonto: University of Toronto Press].
Davidson, D. (2001c). Freedom to act. In D. Davidson *Essays on actions and events* (pp. 63–82). Oxford, UK: Clarendon. [Reprint of Davidson, D. (1973). Freedom to act. In T. Honderich (Ed.), *Essays on freedom of action*. London: Routledge & Kean Paul].
Davidson, D. (2001d). Rational animals. In D. Davidson *Subjective, intersubjective, objective* (pp. 95–106). Oxford, UK: Clarendon. [Reprint of Davidson, D. (1982). Rational animals. *Dialectica*, *36*(4), 317–327].
DeGrazia, D. (2009). Self-awareness in animals. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 201–217). Cambridge, UK: Cambridge University Press.
Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
Dennett, D. (1996). *Kinds of minds*. New York: Basic Books.
Dreckman, F. (1999). Animal beliefs and their contents. *Erkenntnis*, *51*, 93–111.
Dretske, F. (1998). Minds, machines, and money: What really explains behavior. In J. Bransen & S. E. Cuypers (Eds.), *Human action, deliberation and causation* (pp. 157–174). Dordrecht: Kluwer.
Dretske, F. (2006). Minimal rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 107–116). Oxford, UK: Oxford University Press.
Dreyfus, H. L. (2007). The return of the myth of the mental. *Inquiry*, *50*(4), 352–365.
Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
Fodor, J. A. (1990). Making mind matter more. In J. A. Fodor *A theory of content and other essays*. Cambridge, MA: MIT Press. (reprint).
Fodor, J. A., & Lepore, E. (1993). Is intentional ascription intrinsically normative? In B. Dahlbom (Ed.), *Dennett and his critics: Demystifying mind* (pp. 70–82). Malden, MA: Blackwell.
Frankfurt, H. (1978). The problem of action. *American Philosophical Quarterly*, *15*(2), 157–162.
Frankfurt, H. G. (1998). Freedom of the will and the concept of the person. In H. G. Frankfurt *The importance of what we care about* (pp. 11–25). Cambridge, UK: Cambridge University Press. [Reprint of Frankfurt, H. G. (1971). Freedom of the will and the concept of the person. *The Journal of Philosophy*, *68*(1), 5–20].
Ginet, C. (1990). *On action*. New York: Cambridge University Press.

Glock, H.-J. (2009). Can animals act for reasons? *Inquiry*, *52*(3), 232–254.
Glock, H.-J. (2010). Can animals judge? *Dialectica*, *64*(1), 11–33.
Godfrey-Smith, P. (2003). Folk psychology under stress: Comments on Susan Hurley's 'Animal action in the space of reasons'. *Mind & Language*, *18*(3), 266–272.
Hornsby, J. (2004). Agency and actions. In J. Hyman & H. Steward (Eds.), *Agency and actions* (pp. 1–24). Cambridge: Cambridge University Press.
Hurley, S. (2003a). Animal action in the space of reasons. *Mind & Language*, *18*(3), 231–256.
Hurley, S. (2003b). Making sense of animals: Interpretation vs. architecture. *Mind & Language*, *18*(3), 272–280.
Hurley, S. (2006). Making sense of animals. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 139–172). Oxford, UK: Oxford University Press.
Jamieson, D. (2009). What do animals think? In R. Lurz (Ed.), *The philosophy of animal minds* (pp. 15–34). Cambridge, UK: Cambridge University Press.
Kalis, A. (2011). *Failures of agency*. Lexington, MA: Lexington Books.
Kearns, S., & Star, D. (2009). Reasons as evidence. *Oxford studies in metaethics* (pp. 215–242). Oxford, UK: Oxford University Press.
Köhler, W. (1925). *The mentality of apes*. London: Kegan. (transl. from the 2nd German edition by Ella Winter).
Korsgaard, C. (2008). *The constitution of agency: Essays on practical reason and moral psychology*. Oxford, UK: Oxford University Press.
Korsgaard, C. (2009). *Self-constitution: Agency, identity, and integrity*. Oxford: Oxford University Press.
Lurz, R. W. (2003). Neither hot nor cold: An alternative account of consciousness. *Psyche*, *9*(1), 1–18.
McDowell, J. (2007a). What myth? *Inquiry*, *50*(4), 338–351.
McDowell, J. (2007b). Response to Dreyfus. *Inquiry*, *50*(4), 366–370.
Mele, A. (1992). *The springs of action: Understanding intentional behavior*. Oxford, UK: Oxford University Press.
Mele, A., & Moser, P. (1994). Intentional action. *Noûs*, *28*(1), 39–68.
Millikan, R. G. (1995). What is behavior? A philosophical essay on ethology and individualism in psychology, part 1 and the green grass grows all around: a philosophical essay on ethology and individualism in psychology, part 2. In R. G. Millikan *White queen psychology and other essays for alice* (pp. 135–170). Cambridge, MA: MIT Press.
Millikan, R. G. (2000). Reading mother nature's mind. In D. Ross, A. Brook, & D. Thompson (Eds.), *Dennett's philosophy: A comprehensive assessment* (pp. 55–76). Cambridge, MA: MIT Press.
Millikan, Ruth Garrett. (2006). Styles of rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 117–126). Oxford, UK: Oxford University Press.
Nussbaum, M. (2006). *Frontiers of justice: Disability, nationality, species membership*. Cambridge, MA: Harvard University Press.
Pacherie, E. (2011). Nonconceptual representations for action and the limits of intentional control. *Social Psychology*, *42*(1), 67–73.
Paul, S. K. (2009). How we know what we're doing. *Philosophers' Imprint*, *9*(11), 1–24.
Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31*, 109–130.
Penn, D., & Povinelli, D. J. (2007). Causal cognition in human and nonhuman animals: A comparative, critical review. *Annual Review of Psychology*, *58*, 97–118.
Penn, D. C., & Povinelli, D. J. (2013). The comparative delusion: Beyond behavioristic and mentalistic explanations for nonhuman social cognition. In H. S. Terrace & J. Metcalfe (Eds.), *Agency and joint attention* (pp. 62–81). Oxford, UK: Oxford University Press.
Povinelli, D. J., & Frey, S. H. (in review). Constraints on the representation of the functional properties of objects in expert tool-using chimpanzees.
Povinelli, D. J. (2003). *Folk physics for apes*. Oxford, UK: Oxford University Press.
Povinelli, D. J. (2012). *World without weight: Perspectives on an alien mind*. Oxford, UK: Oxford University Press.
Povinelli, D. J., & Giambrone, S. (1999). Inferring other minds: Failure of the argument by analogy. *Philosophical Topics*, *27*, 167–201.
Povinelli, D. J., & Penn, D. C. (2011). Through a floppy tool darkly: Toward a conceptual overthrow of animal alchemy. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 69–88). Oxford, UK: Oxford University Press.

Quinn, W. (1993). *Morality and action*. Cambridge, UK: Cambridge University Press.
Railton, P. (2009). Practical competence and fluent agency. In D. Sobel & S. Wall (Eds.), *Reasons for action* (pp. 81–115). Cambridge: Cambridge University Press.
Rowlands, M. (2006). The normativity of action. *Philosophical Psychology*, *19*(3), 401–416.
Saidel, E. (2009). Attributing mental representations to animals. In R. W. Lurz (Ed.), *The philosophy of animal minds* (pp. 35–51). Cambridge, UK: Cambridge University Press.
Schlosser, M. E. (2007). Basic deviance reconsidered. *Analysis*, *67*(3), 186–194.
Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. New York: Cambridge University Press.
Sehon, S. (2005). *Teleological realism: Mind, agency, and explanation*. Cambridge, MA: MIT Press.
Setiya, K. (2007). *Reasons without rationalism*. Princeton, NJ: Princeton University Press.
Smith, M. (1994). *The moral problem*. Malden: Blackwell Publishing.
Sober, E. (2005). Comparative psychology meets evolutionary biology: Morgan's canon and cladistic parsimony. In L. Daston & G. Mitman (Eds.), *Thinking with animals* (pp. 85–99). New York: Columbia University Press.
Sterelny, K. (2003). Charting control-space: Comments on Susan Hurley's 'Animal action in the space of reasons'. *Mind & Language*, *18*(3), 257–265.
Steward, H. (2009). Animal agency. *Inquiry*, *52*(3), 217–231.
Stich, S. (1979). Do animals have beliefs? *Australasian Journal of Philosophy*, *57*(1), 15–28.
Stoecker, R. (2009). Why animals can't act. *Inquiry*, *52*(3), 255–271.
Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.
Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford, UK: Oxford University Press.
Velleman, D. (1985). Practical reflection. *The Philosophical Review*, *94*(1), 33–61.
Velleman, D. (1989). *Practical reflection*. Princeton, NJ: Princeton University Press.
Velleman, D. (2000). The possibility of practical reason. In D. Velleman *The possibility of practical reason* (pp. 170–199). Oxford, UK: Oxford University Press.
Wallace, R. J. (2006a). Normativity and the will. In R. J. Wallace *Normativity and the will* (pp. 71–81). New York: Oxford University Press.
Wallace, R. J. (2006b). Normativity, commitment, and instrumental reason. In R. J. Wallace *Normativity and the will* (pp. 82–120). New York: Oxford University Press.
Watson, G. (2004). The work of the will. In G. Watson *Agency and answerability: Selected essays*. Oxford, UK: Oxford University Press.
Wilson, G. (1989). *The intentionality of human action*. Stanford, CA: Stanford University Press.