White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/121715/

Version: Accepted Version

Article:
Morgan , Daniel (2018) 'Impersonal Intentions'. The Philosophical Quarterly, 68 (271). pp. 376-384.

https://doi.org/10.1093/pq/pqx053

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Impersonal Intentions

Abstract: Matthew Babb offers a strikingly elegant argument for, and explanation of, the essential indexicality of intentional argument. His two key thoughts are that intentional action always involves intentions, and intentions are essentially indexical. In particular, every intention is indexically *about the agent whose intention it is*, i.e. *de se*. In this paper, I set out two models on which at least some intentions are not *de se* – they are *impersonal* – and I show that these models are compatible with the data Babb points to. I also set out some more data that an account of essential indexical cases ought to be responsive to. Its span suggests that the claim that all intentions are *de se*, even if true, cannot be what explains what is going on in essential indexical cases involving intention.

Key-words: intention, action, essential indexicality, *de se*, rationality.

Mathtew Babb (2016) offers the following strikingly simple argument for, and explanation of, the essential indexicality of intentional action and agency:

1. Intentions are essential to intentional actions and agency.

2. Intentions are essentially indexical

3. Therefore, intentional action and agency are essentially indexical (2016:447).

There are probably readings of 'intentional' on which 1 is false. For example, one might say that an action is intentional if and only if is directed by the whole agent as opposed to some subsystem of the agent. On this reading, spiders are intentional agents despite not having any intentions. But it is also very natural to think that there are other readings on which 1 is true and, indeed, truistic. Babb comments regarding 1: 'That intentional action requires having intentions is as close to a truism as one can get' (2016: 448). I won't query 1.

Babb's reason for accepting 2 is that he thinks that any intention will be *indexically about the agent whose intention it is*. Of course, there are other ways in which an intention might be

indexical – e.g. by being indexically about a time, or a place. But the case Babb makes for 2 does depend on just one specific kind of indexicality. To keep this in view, I'll talk of intentions, as Babb is thinking of them, as being *de se*, rather than just indexical. 2 is supposed to be is true because all intentions are *de se* – *i.e.* indexically *about the agent*.

What considerations does Babb offer for the claim that all intentions are *de se*? One consideration is that *the idea of one agent making another agent's decisions for them is incoherent*. Another consideration – independent, it seems to me – is that *agents are capable of resolving cases of conflicting intention*. I'll provide two models of intentions on which at least some intentions are not *de se* – they are *impersonal*. On the first impersonal model, some intentions are not *about the agent* at all. On the second impersonal model, some intentions are about the agent without being *de se*. It's an open question whether either of these models is correct. But I'll show that neither of them has a problem with either of the two considerations Babb points to. In section I, I set out Babb's point about the impossibility of deciding for someone else, and introduce the two non-*de se* models of intention. In section II, I consider Babb's argument from conflicting intentions. In concluding, I set out some data that I think an account of essential indexical cases ought to be responsive to. The span of the data suggests that the claim that all intentions are *de se*, even if true, cannot be what explains what is going on in essential indexical cases involving intention.

Section I: The impossibility of making another's decisions for them.

Very often we ascribe intentions in a way that makes it clear that they are about the agent who has the intention. E.g. 'I intend that *I* go running', 'I intend *to go* running'. 'You intend that *you* go running'. In all these cases, the embedded clauses have subject terms

(an unvoiced one – PRO – in the case of 'I intend *to go* running') that co-refer with the subjects of the main clause.

We don't always speak that way. We also say things like: 'Jill intends that Jack go running' or 'Jill intends Jack to go running'. But, in every such case, Babb's thought is, we need to supply some unarticulated material. Once this material is supplied it will be clear that the intention ascribed really is about Jill. We need to understand the claim to be: 'Jill intends that she see to it that Jack goes running' or 'Jill intends to see to it that Jack goes running'.

What if someone resisted that and said that, on the reading they have in mind, Jill's intention really is just about Jack and his running, not about Jill at all? The problem, Babb says, is that such a reading would:

> … imply one person has direct authority over someone else's agency, which is incomprehensible. For example, (4c) ['Jill intends Jack to go running'] implies [if read in this unsupplemented way – *author*] that Jill has direct authority over not just Jack's actions, but over his agency as well. By 'direct' authority, what is meant is not that Jill can cause Jack to do certain things. What is meant is Jill literally makes Jack's decisions. She is not merely making his decisions *for him*; she is making his decisions *full stop*….In order for Jill to make Jack's decisions *full stop* Jill would have to be Jack. That however is beyond the reach of possibility. (2016:451)

I think this point about its being impossible for one agent to make another's decisions is correct, and very insightful. But I also suspect that it is a bit of a red herring in this

context. Notice first that there are plenty of sentences that ascribe intentions that aren't obviously about the agent whose intention it is, but to which no one's else's agency is relevant. Here are four:

'Jill intends that Jack be shot in the back'.

'Jill intends that Ayer's rock cease to exist.'

'Jill intends that the tea be in the cup.'

'Jill intends that someone raise their hand.'

One might, and Babb presumably does, regard these sentences as not making sense without supplementation. E.g. one might think that the first and the second have to be understood as ascribing to Jill, respectively, the intention that *she* see to it that Jack be shot in the back, and the intention that *she* see to it that Ayer's rock cease to exist. But the reason that the four sentences do not make sense without supplementation (if they don't) cannot be that the unsupplemented readings imply that Jill has direct authority over someone else's agency. In the case of the first sentence, there is a second agent on the scene – Jack – but the predicate chosen ('be shot in the back') makes it clear that his agency is not in question. In the other three cases, there is no second agent. Relatedly, someone who insists that these four sentences *do* make sense unsupplemented is in no way committed to the claim that, sometimes, one agent makes another agent's decision.

I'm going to describe two models of how some intentions could fail to be *de se*. On the first, the relevant intentions are not about the agent *at all* – they are *worldly* (so, on this

sense, the unsupplemented readings of sentences like 'Jill intends that Ayer's rock cease to exist.' do make sense). On the second, the relevant intentions are about the agent but not in a *de se* way. I'll be emphasizing how neither model has a problem with the point that nobody can make anybody else's decisions for them. [1]

The worldly intentions model

Here is a self-directed-intention free explanation of Jill's intentionally seeing to it that Ayer's rock ceases to exist. Jill formed the worldly intention that Ayer's rock cease to exist. This did *not* cause her to start deliberating about appropriate means to that end. In particular, it did *not* lead her to form an intention to the effect that *she* press a button that sends a missile flying towards Ayer's rock. But, of course, Jill had to do something with her body to act on Ayer's rock. Some means had to be selected. And the particular means selected was the pressing of the button. This selection of means was achieved subpersonally, rather than via the formation of a self-directed intention. As is very familiar, one sometimes intentionally moves one's arm *by* twitching certain muscles without the twitching of the muscle's being something one intends. Similarly, the thought is, Jill has intentionally seen to it that Ayer's rock ceases to exist *by* pressing a button that sends a missile flying towards Ayer's rock. But she didn't form the intention that *she* press the button, or that *she* send the missile flying, or even that *she* see to it that Ayer's rock ceases to exist. The only intention she formed is the worldly intention that Ayer's rock cease to exist.

[1] I make no claim for originality in relation to these models. Both of them are present in Cappelen and Dever (2013). The first is also present in Magidor (2015). The second is also present in Millikan (1990), (2001).

I doubt that this explanation would be the correct one for most realistic instances of a human agent intentionally seeing to it that Ayer's rock ceases to exist. In imagining a realistic case of a human agent doing that, one naturally imagines the agent as being able to say something authoritative about the means they are employing (e.g. 'I am pressing this button'). But they wouldn't be in a position to say that, if pressing the button had just been selected by their subpersonal systems.

Perhaps this isn't just a point about human agency. Perhaps there are *in principle* limits on what kind of thing can be outsourced to the subpersonal level which means that the idea of an intentional agent (a God, say) simply forming a worldly intention and then executing it without forming any self-directed intention about appropriate means is subtly incoherent. But, to make that credible, one needs to give some account of the *source* of the incoherence. Babb's point about the impossibility of making someone else's decisions for them is compelling, and it seems to be a general point about agency as opposed to something special to human-agents. But it doesn't engage with the model just outlined. Acting on a worldly intention need not involve making anyone else's decisions for them.

The self-directed but not *non-de se* intentions model.

Suppose one sets the worldly intentions model aside and accepts Babb's claim that all intentions are self-directed. So, for example, Jill's intention that Ayer's rock ceases to exist is really the intention that *she* see to it that Ayer's rock ceases to exist. It doesn't immediately follow that all intentions are indexical, or *de se*. Ruth Millikan thinks that each of us has a behaviorally special or 'active' way of thinking of herself, or 'self-name',

that leads directly to the performance of intentional actions.[2] This is, according to Millikan, non-indexical and non-*de se*. Suppose '@RM' is the label for Millikan's own active way of thinking of herself. One way in which '@RM' is behaviorally special will be that it figures in the intentions that are the direct causes of Millikan's actions – e.g. Millikan's intention to go running involves a tokening of the mental sentence '@RM runs' in Millikan's intention-box. Moreover, '@RM' is expressed by Millikan in public language using 'I'.

'In making RM's public self references, in purposefully and competently using "I", I manifest the activity of my active self name, the name of the person I know how to act.' (2001: 732).

One tempting response is to say that an intention that involves '@RM' is trivially *de se* since it is expressed using 'I', and being expressible using 'I' just is being *de se*.

'De se' (and 'indexical') are just words. One can use them any way one likes. But using them in the proposed way threatens to make interesting questions disappear. Millikan will say that just as Millikan has '@RM', Babb has '@Babb'. However, '@RM' and '@Babb' are different. It isn't the case that there's a single way of thinking which when used by Millikan picks her out and when used by Babb picks Babb out, and it isn't true that there's any way of thinking that picks Millikan out *because* she is the person who is deploying it. '@RM' hooks onto Millikan in much the way a proper name does, despite the fact that it is expressed in language using 'I'. Is Millikan right about all of this? Or is there reason to acknowledge indexicality in thought as well as language? These seem like genuine questions. Saying that *all it means* to say that an intention is '*de se*' is that it is

[2] See Millikan (1990) and (2001).

expressed in public language using the word 'I' threatens to obliterate them (or just forces us to find some other way to articulate them). So, we should not use '*de se* self-directed intention' to just mean 'a self-directed intention expressible with 'I' (and Babb nowhere proposes that we do that). If one thinks that intentions are always *de se*, then one has to give a reason to think that Millikan's picture – on which intentions are *never de se* – is wrong. I note that, on the face of it, Millikan's picture does not have any problem with the consideration that making someone else's decision for them is impossible. Jill's '@Jill' way of thinking is precisely not a way of thinking of someone other than Jill (e.g. Jack). It isn't even true that Jill's '@Jill'-intentions are ones that 'for all Jill knows, could be about someone else'. If Jill is asked who one of these intentions is about she will say: 'It's about *me*'.

Section II: The argument from conflicting intentions

Babb has a second (independent, it seems to me) argument for the claim that intentions have to be *de se* that focuses on how conflicts of intention are resolved. Since I think there are different ways in which this argument might be reconstructed, I'm going to quote fairly extensively from the portion of Babb's paper in which it appears:

> For those like Castañeda the process of deciding what to do, or practical reasoning, always takes place within a first-personal, subjective perspective, which is 'personal, ephemeral, confrontational, and executive' (1989: 126). The point Castañeda is pushing is that there is no such thing as deciding what to do, or reasoning, under a third-person, objective perspective, as this would be mere computation or calculation, as a non-agential computer would do. We, however, are not mere computer. Similarly, Burge argues that 'fully understanding the

> concept of reason, and engaging in reasoning in the most reflective and articulated way, require having the *I* concept and being able to apply it for this purpose' (2000: 259). Though Burge's point is weaker than Castañeda's (he is only saying that one cannot *understand the concept* of reason fully unless one can conceptualize oneself in a first-person way), it still serves to mark an important relation between reason and the first-person.
>
> That we are not mere computers can be drawn out by considering cases of conflicting intentions. Suppose that on Monday Jill forms an intention to phone her mother at 4 pm on Thursday. Now it is Wednesday and Jill is trying to schedule a meeting with her boss to discuss a problem at work. The only time her boss is available is 4 pm on Thursday. Not thinking about the intention she formed earlier, Jill agrees to meet then and thereby forms an intention to do so. Soon afterwards, she realizes her mistake. She now has two intentions that cannot both be satisfied. If Jill were a mere computer, this conflict would cause a system crash. But this is not what happens. Rather, Jill (*qua* herself) recognizes the conflict and makes an executive decision to abandon one of the intentions and keep the other (or abandon both). It is here, in deciding between the conflicting intentions, that we see Jill is more than a computer, that decision-making requires an agent herself to be part of the process (2016:453).

I suspect that invoking the idea of a 'mere computer' is not helpful in this context. If a *mere* computer is a computer that is not an intentional agent then, of course, Jill is not a mere computer. If a *mere* computer is an intentional agent whose actions aren't always explained by *de se* intentions, then it isn't obvious that Jill is not a *mere* computer. It's hard to see what intermediate understanding could help move the dialectic forward.

It's clear that Castañeda and Burge, who Babb approvingly cites, think that there is some interesting connection between agency and the *de se*. But Babb also offers an argument, which turns on the phenomenon of resolving conflicting intentions, and it is this that I am most interested in. The question I want to press is whether someone who accepts either of the models of non-*de se* intention I have outlined couldn't give the very same answer to the question of how Jill resolves conflicts in her intentions that Babb recommends. What stops them from saying that Jill simply recognizes the conflict and as a result abandons one or both of the two intentions?

There is no difficulty with the idea that this makes Jill herself part of the decision-making process. Everyone ought to, and can, accept that Jill is part of that process.

The non-*de se* version of the answer won't make any play with the thought that Jill '(*qua* herself)' recognizes the conflict of intention. That is, it won't if the job '(*qua* herself)' does is to invoke the *de se* in some way. But, the question is: why is the answer any the worse for that?

The phrase '(*qua* herself)' might not be intended to invoke the *de se* (note that it doesn't in fact appear within the scope of an attitude verb, which is what makes it a bit hard to interpret). It might just be a reminder that Jill herself, the agent, is involved in the process. But, again, everyone can accept that.

It does seem true and important that conflicts among different mental states (e.g. intentions, but also beliefs) *of a single subject* make for irrationality in a way that conflicts among different mental states of *different subjects* do not. Moreover, one might think that

rational subjects are responsible for monitoring their own mental states and removing conflicts. One might further think that for the monitoring to be effective the *de se* must somehow be involved ('I intend p, and I also intend not p. Something has to give…'). Might this line of thought provide a route towards Babb's conclusion that intentions, and therefore intentional actions, always involve the *de se*?[3]

It seems not to. First, if the thing that makes for irrationality is conflict among the mental states held by a single subject the monitoring of those mental states by the subject seems like a means to an end – that of removing the conflict. And it isn't obvious why the conflict could only be removed via that means. On the face of it, this is something that could be taken care of subpersonally, without the intervention of a reviewing subject. Suppose that is wrong though, and that the intervention of the reviewing subject is essential. It still isn't obvious why the intervention – the belief about the existence of a conflict the subject forms – needs to involve the *de se*. Suppose Jill thinks something that she would express would express using the sentence 'That first intention *of mine* conflicts with that second intention *of mine*', and this causes one of her two intentions to be abandoned. Someone who likes the *non-de se* but self-directed model of intentions sketched earlier can deny that that recognition involves the *de se*. The thing that Jill realizes, and uses the first-person possessive to express, is non-*de se*: 'That first intention of @Jill's conflicts with that second intention of @Jill's'. Finally, even if the intervention does involve the *de se* it doesn't follow that the intentions intervened on need to be *de se*. The belief 'I believe that climate change is real, and I believe that climate change is not real' is *de se*, even though the content of neither first-order belief about climate change is *de se*. Similarly, 'I intend this thing, and I intend that conflicting thing' might be a *de se* intervention. But it doesn't follow that intentions the belief is about need to be *de se*.

[3] Thanks to an anonymous reviewer for the suggestion that this line of thought be considered.

I've argued that the claim that avoiding conflicts among one's mental states requires the *de se* is unsupported. But, even if it were true, appealing to it wouldn't implement Babb's strategy of arguing that intentional action requires the *de se* because all intentions are *de se*. I conclude by giving a reason to be skeptical about Babb's intention-centric explanatory strategy of essential indexicality that abstracts away from particular attempts to implement it.

Conclusion

It isn't obvious that intentions *aren't* always *de se*. There are no clear counterexamples to the claim that they are. For any putatively worldly intention (e.g. the intention that Ayer's rock cease to exist), it will be possible to find a self-directed intention that that intention *might* be identified with (e.g. the intention *that I see to it* that Ayer's rock cease to exist). This self-directed intention will be expressible with 'I'. And there *might* well turn to be reasons to think that, pace Millikan, if a mental state is expressible using 'I', then it must itself be *de se*.

Suppose then, for the sake of argument, that intentions are always *de se* – intentions, as a mental state type, are *de se.* I would still want to claim that this leaves the most interesting questions about essential indexical cases, including essential indexical cases that involve intention, open. In clarifying this suggestion, I set out a span of four different essential indexical cases:

> *Case 0:* I intend that the Messy Shopper stop making a mess. I don't know that *I* am the Messy Shopper. I keep on making a mess.

Case 1: I intend to leave for my meeting at 3pm. I don't know that it is 3pm *now*. I stay in my office.

Case 2: I am in pain. I introspect. I don't know that I am the Messy Shopper. I am asked to assign truth-values to two items on a questionnaire – 'I am in pain' and 'The Messy Shopper is in pain'. I say 'yes' to the first, and 'don't know' to the second.

Case 3. I am an unsophisticated and impulsive agent. I never form decisions about how to act and then store them up for future control of action. Accordingly, the most compelling reason for positing intentions does not apply to me. The conative aspect of my psychology is adequately dealt with by desire. I desire 'S eat' but not 'I eat'. I do not eat, even though I am S.

In each case, the distinction between an indexical and a non-indexical ways of thinking of something seems crucial. In *Case 0, Case 1* and *Case 3* having an indexical way of thinking seems essential if one is to intentionally do something. In *Case 2*, having an indexical way of thinking seems essential to articulate what it is one is a position to know. It is very natural to think that the explanation of the four cases will have something in common. Relatedly, I would only have confidence in an explanation of what is going on in *Case 0* if it looked as though the explanation, or some analogue of it, could extend to the other three cases.[4] The point that intentions – as a mental state type – are always *de se* is not

[4] Not everyone will be persuaded by this argument from theoretical unity. I note that, even bracketing it, it isn't clear that the claim that all intentions are *de se* helps at all with *Case 0*. If *all* intentions are *de se*, then even my (pre-light bulb moment) intention that the Messy Shopper stop making a mess is *de se*. Such an intention won't lead directly to action, whereas my (post-light bulb moment) intention that *I* stop making a mess will lead directly to action. Intuitively, our chief interest in *Case 0* is to know why, post-light bulb moment, I am in a mental state that can lead *directly* to action, whereas before I was not. If *both* intentions

going to be the basis of an explanation that extends in this way. *Case 1* involves intention and the *de nunc*. But it isn't true that intentions are always *de nunc* – I can intend at 4pm to do something at 5pm. *Case 2* involves *de se belief*. But it isn't true that belief is *always* about the believer – e.g. my belief that 2+2=4 is not about me. *Case 3* involves *de se desire*. But, it isn't true that desire is *always* about the desirer. Accordingly, I suspect that the claim that all intentions are *de se*, even if true, must give way to a deeper explanation of what is doing on in *Case 0,* one which captures why it is that it belongs in same theoretical natural kind as *Cases 1-3*.[5]

University of Barcelona, Spain

Bibliography:

Babb, M (2016) 'The Essential Indexicality of Intentional Action' *The Philosophical Quarterly*. Vol 66 no 264 439-457.

Cappelen, H and Dever, J (2013) *The Inessential Indexical: on the Philosophical Insignificance of Perspective and the First Person.* Oxford: Oxford University Press.

Lewis, D (1979) 'Attitudes *De Dicto* and *De Se'*, in *The Philosophical Review* 88: 513-543.

Magidor, O (2015) 'The Myth of the De Se', *Philosophical Perspectives* 29 (1): 249-283.

Millikan, R (1990) 'The Myth of the Essential Indexical', in *Noûs* 24 (5): 723-734.

Millikan, R (2001) 'The Myth of Mental Indexicals', revised and expanded version of "The Myth of the Essential Indexical," Nous 24.5 (1990) pp. 723-734.

---

are *de se*, then *de se* ness on its own doesn't explain this contrast. One needs a *de se* intention *of the right sort*. And then the question is: why is the post-light bulb moment intention the right sort of *de se* intention?

[5] For discussion and comments, thanks to Michele Palmira, Léa Salje and an anonymous reviewer for this journal.