Chapter 3

Indexicality, the *De Se*, and Agency

Is there a philosophically significant connection between indexicality\(^1\) and agency? According to orthodoxy the answer is a resounding “yes.” It is more or less common ground in contemporary philosophy that some such significant connection exists. It is taken for granted to such an extent that it by now seems unnecessary to provide arguments for it—all one needs to do is include a footnote referring to Perry 1979 (or Lewis 1979) and that’s supposed to settle it. Consider the following passages (some of which were quoted above):

The present suggestion, then, is that indexical concepts are ineliminable because without them agency would be **impossible**: when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I *eo ipso* imagine myself deprived of the power to act. (McGinn 1983, p. 104, bolded emphasis added)

It is widely agreed that agents need information in an egocentric form: they must think of places as “here” and “there”, times as “now” and “then” if they are to be able to act on what they know (Perry 1979). (Owens 2011, p. 267)

A strong contemporary current runs to the effect that the **ability of an agent to project knowledge of the world into relevant action in the world depends upon the ability to think indexical thoughts.** (Millikan 1990, p. 723, bolded emphasis added)

...practical guidance is, in Perry’s phrase, essentially indexical, in the sense that its function depends not only on which of many propositions it expresses but

\(^1\) Or egocentric thought, or the first person, or the *de se* or (insert your favorite piece of relevant terminology).
Also on how that proposition is determined by the context—specifically, on its being determined in the same way as the reference of indexical expressions such as “I”, “you”, “here”, and “now”. (Velleman forthcoming)

This chapter has two goals: First, to try to make precise exactly what connections are supposed to obtain between indexicality and agency. Second, to evaluate arguments that can be given for those alleged connections. Two points are worth highlighting right away:

a) It is probably uncharitable to read Perry (1979) as even attempting to present arguments for such a connection. It’s not that he failed to establish a philosophically interesting indexicality–agency connection—he didn’t even try.

b) Nonetheless, the assumption that such a connection exists is so entrenched that it is worth trying to articulate the arguments we suspect are tacitly (and sometimes embryonically explicit) appealed to.

The chapter is structured as follows. We first present a natural reading of Perry’s paper “The Essential Indexical” according to which it aims to show nothing more than the opacity of action-explanation contexts. Sections 3.2–3.4 make some preliminary remarks about how to articulate stronger (and more interesting) theses about the connection between indexicality (and the de se) and action. We then discuss four possible motivations for these stronger theses. Sections 3.9 and 3.10 outline our positive account of action rationalization and respond to some objections. Our overall conclusion is pessimistic: it turns out that the role of indexicality in explanation and rationalization of action turns out to be minimal and trivial.

3.1. Perry on the Opacity of Explanation Contexts

Despite its familiarity by now, it will be helpful to have in front of us some crucial passages from Perry to draw on. Perry relies on several examples that are structured more or less the same way. One of them is this:

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other,
seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch...I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn’t believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in my cart. My change in beliefs seems to explain my change in behavior. My aim in this paper is to make a key point about the characterization of this change, and of beliefs in general. (1979, p. 3)

The claim Perry makes about this and the related scenarios is most naturally read as a claim about opacity of explanation contexts. Here are two summarizing passages from Perry where that reading is particularly salient:

At first characterizing the change seems easy. My beliefs changed, didn’t they, in that I came to have a new one, namely, *that I am making a mess*? But things are not so simple.

The reason they are not is the importance of the word “I” in my expression of what I came to believe. *When we replace it with other designations of me, we no longer have an explanation of my behavior and so, it seems, no longer an attribution of the same belief. It seems to be an essential indexical.* But without such a replacement, all we have to identify the belief is the sentence “I am making a mess”. But that sentence by itself doesn’t seem to identify the crucial belief, for if someone else had said it, they would have expressed a different belief, a false one. (p. 3, bolded emphasis added)

One of the key arguments throughout Perry’s article is this (and variations on it):

Suppose I had said, in the manner of de Gaulle, “I came to believe that John Perry is making a mess”: I would no longer have explained why I stopped and looked in my own cart. To explain that I would have to add, “and I believe that I am John Perry”, bringing in the indexical again. After all, suppose I had really given my explanation in the manner of de Gaulle, and said “I came to believe that de Gaulle is making a mess.” That wouldn’t have explained my stopping at all. But it really would have explained it every bit as much as “I came to believe John Perry is making a mess”. For if I added “and I believe that I am de Gaulle” the explanations would be on par. *The only reason “I came to believe John Perry is making a mess” seems to explain*
my action is our natural assumption that I did believe I was John Perry and didn’t believe I was de Gaulle. So replacing the indexical “I” with another term designating the same person really does, as claimed, destroy the explanation. (pp. 4–5, bolded emphasis added)

Here is one way to summarize Perry’s claim:

**Indexical Opacity.** There’s a set of indexicals, $I$-SET, that cannot be substituted *salva veritate* in action-explanation contexts by any other expressions.²

It should be clear that this is an instantiation of the more general thesis:

**Generic Opacity.** Co-referential referring expressions cannot be substituted *salva veritate* in action-explanation contexts.

To see that Generic Opacity is true, note that cases analogous to those appealed to by Perry are easily constructed:

**Superman/Clark Kent.** Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn’t find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn’t believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior.

We take the S/CK case as showing that in general action explanations don’t have their explanatory force preserved by substitution of co-referential singular terms. Seeing Indexical Opacity as an instance of Generic Opacity suggests that there’s nothing deeply central about indexicals here.

² There are many reasons one might believe Indexical Opacity. For example, you might think that the belief *that I am F*, as had by $A$, is a different belief from the belief *that A is F*. If an action explanation then cites the first belief, the substitution of “$A$” for “I” would then change the cited belief, and potentially undermine the explanation.

Our goals, however, are independent of the particular reasons for endorsing Indexical Opacity. We will argue that there is no convincing reason to endorse Indexical Opacity that are not also reasons for endorsing Generic Opacity. So, for example, those who endorse Indexical Opacity because they think $A$’s belief *that I am F* is a different belief from $A$’s belief *that A is F* have given no reasons for thinking that this is not just a special case of the general phenomenon that, when $A$ is $B$, the belief *that A is F* is a different belief from the belief *that B is F*. 
3.2. Causation, Explanation, Motivation, and Rationalization

Most philosophers (certainly the majority of the more than 1,000 who reference Perry’s classic papers on this topic) take Perry’s arguments to have shown something stronger and more interesting than the opacity of explanation contexts (see e.g. the quotes we started this chapter with). It is a challenge to articulate precisely these stronger claims and it will help to start with an overview of different kinds of projects in which indexicality can play a role. Consider François. He ducked under the table. There are multiple questions we can ask about that action:

**Causal.** What is the full causal history of François’s action?

**Explanation.** What is the explanation of François’s performing the action?

**Motivation.** What was François’s motivation for performing the actions (what are the reasons for which he performed the action)?

**Rationalization.** What made the action reasonable/rationalizable/comprehensible for François to perform?
The central question, then, is whether any of these questions require answers that cite “de se” or other indexical mental contents. For ease of exposition we propose to focus on Motivation and Rationalization but we think what we say will apply to Causal and Explanation with only slight terminological changes, if you think they should be the focus.\(^3\)

### 3.3. Impersonal vs. Personal Action

**Explanations/Rationalizations**

We are looking for a way to articulate an insight in Perry’s work that goes beyond the opacity of explanation contexts.\(^4\) One way to look for such insights is to look for ways to show that the impersonal

\(^3\) We think Motivation and Rationalization present the strongest cases for the proponent of Essential Indexicality, so we focus on those. Plausibly the answer to our central question for Causal and Explanation is “no,” and pretty trivially so. For Causal, we can give a full microphysical history of the world, and we’ll get a causal history of the action that doesn’t cite any de se mental states, because it doesn’t cite any mental states at all. For Explanation, we can observe that explanations are extremely context-sensitive, and that there’s typically no explanatory feature that has to be cited in every explanation of a given phenomenon.

There are potential responses to each of these preliminary points. In response to the Causal worries: If we’re type (or maybe even token) identity theorists, then some of the entities cited in the full microphysical causal history will be mental states (although in a different descriptive guise), and hence may be de se mental states. (Although there are worries about the hyperintensionality of “cite” here.) And if we’re not type/token identity theorists, we may worry that the proposed microphysical history does not in fact give the causal history of the action, but only that of its physical correlates. Here’s a plausible principle: a full causal history of an action must, inter alia, imply that the action is an action. A microphysical causal history absent bridging identities looks like it won’t imply that, though.) In response to the Explanation worries: perhaps we can helpfully appeal to the notion of a full or complete explanation. It’s not obvious that there is any such notion that is sufficiently clear, but arguing that there isn’t looks like more than we want to take on here. These are all interesting issues to explore in great detail, but doing so goes beyond the scope of this work. In order to bypass these kinds of discussion (that play no role in, e.g., Perry’s or Lewis’s work on this topic), we focus primarily on Motivation and Rationalization.

\(^4\) As we have already said, we are not claiming these arguments are found in Perry (1979)—they are rather our attempts to articulate what people have read into that paper. We could have just said: “Perry (1979) does not establish anything but opacity of explanation contexts” and so all those who claim he does are wrong and confused, and leave it at that. But we think that would be too simplistic and uncharitable to those who have read more into Perry. So the goal of the remainder of this chapter is to articulate what we think are the underlying assumptions and arguments that make so many philosophers think there’s a significant indexicality-agency connection (some of the arguments we rely on are inspired by
action rationalizations we are about to cite are incomplete in some significant way. Consider a case to get us started. François thinks François is about to be shot. He doesn’t want François to be shot. So he ducks under the table. Here are two models of action rationalization:5

**Personal Action Rationalization 1.**

- **Belief:** François is about to be shot.
- **Belief:** I am François.
- **Belief (Inferred):** I am about to be shot.
- **Desire:** I not be shot.
- **Belief:** If I duck under the table, I will not be shot.
- **Action:** I duck under the table.

**Impersonal Action Rationalization 1.**

- **Belief:** François is about to be shot.
- **Desire:** François not be shot.
- **Belief:** If François ducks under the table, he will not be shot.
- **Action:** François ducks under the table.

For a second illustration of the two models, consider Herman who believes Nora is in danger:

**Personal Action Rationalization 2.**

- **Belief:** Nora is in danger.
- **Desire:** Nora not be hurt.
- **Belief:** If Herman closes the door, Nora will be safe.
- **Belief:** I am Herman.
- **Action:** Herman closes the door.

comments Perry makes in later papers). In other words, if you suspect we are attacking strawmen in what follows, feel free to either ignore the remainder of this chapter (and agree with us that those who think Perry established something of an interesting indexicality–agency connection are all wrong) or come up with better arguments than those we have presented on behalf of our opponents.

5 We are not committed to any particular view of what combination of beliefs, desires, or intentions goes into an adequate action explanation. The points we are making here can be adjusted to whatever your favorite format is for action explanations.
Impersonal Action Rationalization 2.

- **Belief:** Nora is in danger.
- **Desire:** Nora not be hurt.
- **Belief:** If the door is closed, Nora will be safe.
- **Action:** Herman closes the door.

Note that in this case the impersonal action rationalization doesn’t attribute “Herman”-beliefs or “Herman”-desires to Herman. Instead, the rationalization is entirely third-person. We take our target to be the following view:

**Impersonal Incompleteness Claim (IIC).** Impersonal action rationalizations (IAR) are necessarily incomplete because of a missing indexical component.⁶

Our guiding question is thus whether there are any arguments to the effect that the IAR-explanations must be incomplete. We’ll argue that there are none. First, it is not necessary for an indexical element to enter into the rationalization. Second, on our view the agent doesn’t even need to be represented in a non-indexical way in an adequate action rationalization. Nor does any part of the agent’s body need to be represented (indexically or not). We think there could be a god, who can bring about states of the world just by intending them or maybe just by thinking them. The god thinks, “The door is closed,” and straightaway the door is closed. On our view, this god’s actions can be rationalized even if we don’t specify any kind of de se state (or, indeed, even a de re representation of the god, or of his body parts).

3.4. Preliminaries: Some Weaker Theses and Why We Focus on IIC

We’re inclined to think that the presumed indexicality-agency connection is the source of much of the fascination with indexicality.

---

⁶ Note that in the examples we have focused on an allegedly missing “I,” but the point could generalize: variations will include thoughts or desires involving “here,” “there,” “now,” “we,” etc.
throughout philosophy. It is therefore important to be crystal clear both about what we think the central indexicality–agency connections might be and how the topics of this chapter relate to the remainder of the book.

We will refute IIC. Here’s a reaction we expect to get from many of our readers (and what we got from readers of earlier drafts). You might then be inclined to think something like this: Okay, well, maybe this was all a set-up. Maybe the indexicality–agency connection that matters isn’t IIC. Maybe it’s some weaker thesis. The focus on IIC is just a strawman. First, some remarks about the strawman remark and then some thoughts about some weaker theses that could be advanced. To see that we’re not attacking a strawman, look back at some of the quotes we gave you earlier:

The present suggestion, then, is that indexical concepts are ineliminable because without them agency would be impossible: when I imagine myself divested of indexical thoughts, employing only centreless mental representations, I eo ipso imagine myself deprived of the power to act.7 (McGinn 1983, p. 104, bolded emphasis added)

Kant’s framing his maxim in the first person is no accident. He could not have restated it, for example, as “Immanuel Kant will make lying promises when he is in need.” Such a third-personal thought would not be a maxim of action, since it could not be acted upon by the thinker until he reformulated it reflexively, in the first person. Insofar as the target of universalization is a practical thought, it is essentially first-personal. (Velleman 2006, p. 121, bolded emphasis added)

When I act intentionally on an object, I care neither about the Russellian (Kaplanian) propositions beyond as such nor about their Fregean (Kaplanian) first referents. These are not crucial. What I need is to bring those doxastic referents somehow into my experience through my thinking indexical (that is,

7 Here is the sentence immediately preceding this passage: “As I remarked in chapter 5, indexicals have recently been associated with practical reasoning, the thesis being that indexical thought is a necessary condition of agency” (p. 104). And here’s what McGinn says in ch. 5,“And as a number of writers have observed, dispositions to action induced by indexical thoughts correspond with the linguistic meaning of the associated indexical sentence: you and I are disposed to act in the same way if we both think ‘I am about to be attacked by a bear’ (other things being equal)” (p. 65). Both quotes are immediately followed by a citation of Perry 1979.
experiential) references...the singular proposition that mobilizes my powers of deliberation and action is the indexical truth within my grasp. (Castañeda 1989, p. 126, bolded emphasis added)

It is widely agreed that agents need information in an egocentric form: they must think of places as “here” and “there”, times as “now” and “then” if they are to be able to act on what they know (Perry 1979). (Owens 2011, p. 267, bolded emphasis added)

We take IIC to be a fair summary of these kinds of claims (and we could have added many, many more citations). Now you might think these claims overreach. Maybe they should have been more modest and so charity requires that we look also for some weaker indexicality–agency connections. What might such weaker connections be? Consider first the following claim:

**Impersonal Incompleteness Claim 1 (IIC1).** In some action explanations/rationalizations, indexicals occur ineliminably.

Important point: *We don’t know how to read IIC1 as anything but a trivial corollary of the opacity of action explanations/rationalizations.* In other words, we grant it, but it shows nothing distinctive about indexicals. For any content, C, there are contexts in which substituting C for any other content can result in a change in truth value. There are, for example, action-explanation (rationalization) contexts in which you can’t substitute “Superman” for some other co-referential term. In the same way there are contexts in which substituting an occurrence of “I” for a co-referential term can’t be done salva veritate. None of this should be a surprise and we don’t need to read papers by Perry and Lewis to learn about it. Recall, we assume throughout that (i)–(ii) are in the common ground among participants in this debate:

i) Belief attribution and action explanations exhibit opacity.

ii) Indexicals have non-stable characters.

IIC1, as we understand it, commits one to nothing more than that. Now, you might think that non-substitutivity as it applies to indexicals is distinctive in some way. If you think that, it is very important to keep

---

8 See for example the passages from Perry quoted in Sect. 3.6.
the structure of these first chapters in mind. In Chapter 4 we explore and then refute the thought that the behavior of indexicals in opaque contexts is distinctive in some way. So if that’s your reason for interest in IIC₁, hold off until Chapter 4 (or maybe go directly there). In sum, we conclude that IIC₁ is true but uninteresting;⁹ to see the full argument for this conclusion, though, you’ll have to look at this chapter in conjunction with the next.

If you agree with that response to IIC₁, you might be tempted to look for a different weaker version of IIC—one that isn’t just about opacity but also isn’t as strong as IIC. We can imagine various versions of IIC₂:

**Impersonal Incompleteness Claim 2 (IIC₂).** As a matter of deep physical necessity, indexical beliefs are needed in order for human beings, constituted as they are, to act.

We talk more about this reply at the end of this chapter (see Objection 4), but in brief: none of the arguments (or motivations) we are going to consider have IIC₂ as a potential conclusion. That should not be surprising—none of the arguments have facts about the physical structure of humans as premises. There is no attempt in the arguments to study in detail the underlying physical structure of humans and their ability to act. That would require arguments and evidence of a completely different kind from what we find in the philosophical tradition we engage with in this work.

So our strategy is to engage with IIC. We should note right away that this isn’t an easy task since there are few real arguments out there to engage with. Put simply, the remarks prominent philosophers make about the indexicality–agency connection don’t seem very plausible candidates for establishing IIC. As a result, we structure the remainder of the chapter around what we call “Motivations.” They’re not really arguments, but they are observations that seem to us to play an important role in the widespread endorsement of IIC. There is a danger here that we are very much aware of: Suppose A endorses a thesis T, but doesn’t tell us why. We then try to figure out why she endorses T. If in

⁹ Again, on the assumption, which we are granting arguendo, that there are opaque contexts.
doing so, the candidates we come up with all fail to establish \( T \), there’s the immediate suspicion that we’ve been uncharitable and that either there are some other (and better) reasons for endorsing \( T \) or that \( A \) didn’t endorse \( T \). The second option we just dealt with. With respect to the first option: we would be delighted if someone could respond to this book by constructing better arguments for IIC. That would really move the debate forward.

### 3.5. Motivation 1: Judgments about Cases

The cases at the beginning of Perry’s “The Essential Indexical” are intriguing, ingenious, suggestive and have for good reasons captured the imagination of at least two generations of philosophers. We suspect that reflection on these cases motivates many of those endorsing various indexicality–agency connections. But what can we learn from them? One thing to note right away is that these cases are not like what has become known as “the Gettier case” or Burge’s famous “arthritis case” in his paper, “Individualism and the Mental.” Burge and Gettier presented counterexamples to necessity theses (put simplistically, the claims that necessarily all instances of justified true belief amount to knowledge and necessarily all contents are determined by “what’s in the mind of the speaker or thinker”). Burge’s and Gettier’s cases, if they succeed, show that these necessity claims are not true. This is not the dialectical setting of Perry’s cases. There are two opposing views: on one view, a distinctly first-person (or other indexical) attitude is required in an adequate action rationalization. According to the other view, it is not required. Perry’s cases have the form of thought experiments in which it is stipulated that at \( t_1 \) there is no action (and no relevant first-person or other indexical attitude) and then at \( t_2 \) there’s action. By stipulation, the only difference between \( t_1 \) and \( t_2 \) is the addition of a first-person state. So we all say: fine, yes, that could happen—the first-person state could have that effect and so could play an important role in rationalization of that particular action. Note that this is at best super-weak inductive support for a strong thesis such as IIC. Without further argument, this won’t justify a necessity claim like IIC. *That an ignorance of the kind the agent starts out with in Perry’s examples can block*
rationalization of agency, doesn’t show that it has to. Cases can be used as counterexamples to universal claims, but can’t be used to (deductively) establish the latter.\textsuperscript{10}

3.6. Motivation 2: Bodily Movements Require Indexical Thought

Here is a passage from Perry and Israel’s “Fodor and Psychological Explanation” (1993):

…the need for context-sensitive representation never goes away. If he wants to drink the coffee, Jerry still has to pick up the cup in front of him.

Consider, for a different example, Jerry’s making a phone call to Zenon. Here the belief, say, “Zenon’s phone number is 555-5555” is relatively context-insensitive. There is a present-tense marker, but phone numbers are relatively stable properties of persons, so we can ignore that. When Jerry calls Zenon, the perceptions that originally gave rise to the belief may be remote in time, and the effect of the action that fulfills the goal—the ringing of Zenon’s telephone and his answering it—are remote in space. Neither the number nor Zenon needs to be thought of demonstratively or indexically to understand the transaction. But, like all human action—except perhaps pure ratiocination—the crux of the matter involves physical interactions with a physical object. The practical reasoning will terminate with some context-sensitive way of thinking about the telephone. The Fodorian version of half of Kant’s maxim about concepts and intuitions should be “Eternal tokens of mentalese without context-sensitive tokens of mentalese are blind” (p. 311, bolded emphasis added)

In a later paper (1998), Perry says:

Consider a transaction with a fax machine. To press certain buttons on it, I have to move my fingers a certain distance and direction from me. It isn’t enough to know where the buttons were relative to one another, or where the

\textsuperscript{10} Of course, what the cases do show is what Perry says they show: that action-explanation contexts are opaque. This, however, is entirely unsurprising (for the reasons we have already given).
fax machine was in the building or room. I had to know where these things were relative to me (p. 87, bolded emphasis added)\(^{11}\)

Note right away that this isn’t really an argument with premises and a conclusion. It is more of a powerful and persuasive statement of the view we’re opposing. That said, we suspect this kind of rhetoric can be effective. We suspect one might have this reaction: Right, isn’t that just obvious?

Reply to Motivation 2. Here is why you shouldn’t be moved by the kind of rhetoric we find in the Perry quote: In order to act, our bodies have to be brought into engagement with the world around us. That much is trivial. But what is not trivial is that for that to happen, we must represent ourselves in relation to the objects we engage with. Consider what Perry says about the fax machine: why do you need to know or have a belief about where the buttons are relative to you? In general, we don’t need to have knowledge or beliefs of all the facts involved in our action. To move a finger, you have to tense various sequences of muscles. But you don’t need to know or have beliefs about what those sequences of muscles are. Somewhere in the physical architecture, the relation of you to the muscles gets implemented, but it doesn’t need to be a cognitive implementation. There is no reason why it can’t be like that for you and the buttons. On the cognitive-representation level, it’s just all about the buttons and their objective position in space. Then a bunch of neurons fire and our bodies end up doing the right thing. If there is no representation of the fingers (whether conscious or subpersonal), there is a fortiori no crucial indexical representation of the fingers. Everyone agrees that the representational level gives out somewhere. For Motivation 2 to be persuasive, it needs to come with some reason to think that it gives out at exactly the right spot: while the self is still involved in the representation. But the passage from Perry doesn’t even make an effort to develop an argument to the effect that this it is impossible for representation to end elsewhere—no evidence is presented that the action couldn’t take place the way we just sketched.

\(^{11}\) For those inclined to doubt that IIC is endorsed by our opponents (despite the passages we have quoted), note especially the use of “had to,” “have to,” and “the need for context sensitive representation never goes away” in these passages.
A couple of further points to bolster the non-indexical characterization of the action: First, note that Perry’s claim about the need for a representation of the agent is, on a simple reading, obviously false. It is not true that the agent and the buttons have to be brought into engagement. What is true is that the agent’s finger and the buttons need to be brought into engagement. That, however, is just two objects in the world that need to be coordinated. There is no good reason to think that she must know where she is relative to these events. Typically, people are near where their fingers are, but in weird cases, that could change, and it wouldn’t matter to the nature of the action.

Second, note that if she happens to think about the buttons’ location relative to her, no argument has been given that she has to think about herself in a distinctly first-person way (whatever that means). If the agent is Nora and she happens to represent Nora’s relationship to the buttons, she could represent herself in a non-first-person way, e.g. as Nora.

In sum, there are at least two mistakes underlying Motivation 2.

**Over-Representation Fallacy.** Because the body is involved in the movement, the body needs to be mentally represented.

**Relational Fallacy.** If some part of the body has to be represented, it has to be indexically represented.

There simply are no arguments or evidence for either of these claims. Not everything involved in an action needs to be represented by the actor. And even if you maintain that some part of the body has to be represented, it doesn’t follow that it has to be indexically represented. What’s important is just that it gets represented, so that it can be directed or controlled. Indexicality just seems to be irrelevant.

It is helpful to contrast our points here with one made by Millikan. Millikan—one of our few allies against Essential Indexicality—says:

Now it is trivial that if I am to react in a special and different way to the knowledge that I, RM, am positioned so in the world, a way quite unlike how I would react knowing anyone else was positioned so in the world, then my inner term for RM must bear a very special and unique relation to my dispositions to act. But what does that have to do with indexicality? My inner name “RM” obviously is not like other names in my mental vocabulary. It is a name that hooks up with my knowhows, with my abilities and dispositions to act, in a rather special way. . . . My inner “RM” is indeed special. Let us call
it @"RM", or RM’s “active self name”. It names a person whom I know, under that name, how to manipulate directly; I know how to effect her behavior. But in order to know how to manipulate this person, why would I need to think indexical thoughts? What has knowhow to do with indexicality? (1990, p. 730)

Millikan is right, and importantly so, in pointing out that there’s not even the beginning of an argument for anything properly labeled “essential indexicality” here. If anything is shown, it has to do with something like what Millikan calls an “active self-name” and it need not be indexical. However, we think that even the claim about an essential role for the “active self name” (or “de se concept,” if you prefer that label) is unsupported. Why think that the agent has to represent herself at all in order to push the button? What we get is simply a claim that this is required, no evidence or argument to the effect that the act couldn’t happen without a representation of the agent.

3.7. Motivation 3: Generalizations about Actions

At the end of “The Essential Indexical,” Perry presents a positive view about indexical content. Perry summarizes his view as follows:

I propose we look at things in this way. The shoppers, for example, are all in a certain belief state, a state which, given normal desires and other belief states they can be expected to be in, will lead each of them to examine his cart. But, although they are all in the same belief state (not the same total belief state, of course), they do not all have the same belief (believe the same thing, have the relation of belief to the same object). We use sentences with indexicals or relativized propositions to individuate belief states, for the purposes of classifying believers in ways useful for explanation and prediction. That is, belief states individuated in this way enter into our common sense theory about human behavior and more sophisticated theories emerging from it. (1979, p. 18)

The following argument can be extracted from this passage. Explaining François’s action using “François”-beliefs rather than “I”-beliefs fails to capture important generalizations. We don’t explain why François and Josh both act in similar ways in similar circumstances. When someone
has a belief at $t$ that would be expressed by the believer (were he an English speaker) by an utterance at $t$ of the sentence type “I am about to be shot,” that person ducks under the table. A bit more generally: suppose the Impersonal Action Rationalization model can succeed. Now suppose that there is some rationalization $R$ of François’s action, featuring various “François”-beliefs and -desires, but no indexical beliefs and desires. Here’s a claim: any agent $A$ who had the analogous set of beliefs and desires (that is, beliefs and desires with “François” uniformly replaced with “$A$”) would be rational to act in the same way as François. The Impersonal Action Explanation model doesn’t explain that. The Personal Action Explanation model does explain it: on this model, exactly the same explanation (involving indexical beliefs and desires) applies to both agents.\footnote{Not every story about indexical contents or de se states will yield the result that the Personal Action Explanation captures the generalization by attributing the same mental states to François and to $A$. The crucial question is whether François’s first-person belief that he is $F$ has the same content as $A$’s first-person belief that he is $F$. Accounts vary in their answer—Fregean pictures of first-person thought, for example, tend to attribute different contents to François’s and $A$’s first-person thoughts, while Lewisian accounts attribute the same content.}

Reply to Motivation 3. Note that the argument we just sketched didn’t end with a conclusion. One conclusion one might draw is this: There are some generalizations that impersonal action explanations cannot capture. Note that this is very far from an interesting thesis like IIC. Suppose it is true that there are some generalizations involving indexicals that help us classify “believers in ways useful for explanation and prediction” (in Perry’s phrase). That’s compatible with there being other, non-de se/first-person generalizations that also are useful in that way. There are all kinds of generalizations we might want to capture, some metalinguistic, some not (and this bolsters our claim that generalizations that appeal to allegedly special de se content have no special status). Here are two generalizations that do not involve indexicals and that, at least for some purposes, will help us classify “believers in ways useful for explanation and prediction”:

**Useful (U1).** Any criminal engaged in a criminal act who has a belief she would express by “Superman is right around the corner” will, ceteris paribus, run.
INDEXICALITY, THE DE SE, AND AGENCY

**Useful 2 (U2).** For any person, x, if there’s a button, y, that x wants to push, y is close to x, y is within x’s visual field, and y is easily within x’s reach, then *ceteris paribus*, x will push y.

Given the immense contextual flexibility in what counts as useful for explanation and prediction of action, there are very natural settings in which U₁ and U₂ can serve that purpose at least as well as a de se-involving generalization such as U₃:

**Useful 3 (U3).** Anyone who has a belief she would express using “I am being pursued by a bear” and believes she can avoid being eaten by that bear by climbing up a nearby tree, will, *ceteris paribus*, climb up the nearby tree.

To re-emphasize the main point here: even if we grant that generalizations like U₃ are true and in some significant sense involve indexicality, that’s too weak to support IIC.¹³

### 3.8. Motivation 4: Anti-Generalizations about Actions

The previous argument focused on the need to capture sameness of beliefs and desires among agents who are motivated to act in the same way. Now we focus on the need to capture difference in belief and desire among agents who are motivated to act differently. Consider a case where two agents have all the same objective, or non-indexical, information about the world. Not only can they be motivated differently, but what it is rational for them to do can differ. To see what

---

¹³ This is a point at which several of our readers have found that they are unsatisfied with our focus on IIC. They agree that Motivation 3 fails to establish or support IIC, but find that it is still interesting that there are generalizations about action involving indexicality. We attempt to address this concern in Sect. 3.4. To summarize: if you think that these generalizations are not just a manifestation of Generic Opacity, and the fact that any different expressions will give rise to different generalizations, but are evidence for a form of Essential Indexicality other than IIC, then you need to formulate the appropriate Essential Indexicality thesis. We are unable to see that there is an interesting Essential Indexicality thesis supported by Motivation 3.
the proponent of IIC might have in mind, consider this example from Perry:

I shall use the term “locating beliefs” to refer to one’s beliefs about where one is, when it is, and who one is. Such beliefs seem essentially indexical. Imagine two lost campers who trust the same guidebook but disagree about where they are. If we were to try to characterize the beliefs of these campers without the use of indexicals, it would seem impossible to bring out this disagreement. If, for example, we characterized their beliefs by the set of “eternal sentences,” drawn from the guidebook they would mark “true”, there is no reason to suppose that the sets would differ. They could mark all of the same sentences “true”, and still disagree in their locating beliefs. It seems that there has to be some indexical element in the characterization of their beliefs to bring out this disagreement. (1979, p. 5)

Perry’s focus here is on how to bring out the disagreement between the two agents. We cannot do so without appealing to indexical content, he claims. We want instead to focus on a separate point: what it is rational for the two agents to do will differ. They both want to reach Mount M. For one of them, it will be rational to turn left, for the other to continue straight ahead, even though they have the very same non-indexical beliefs.

Reply to Motivation 4. As it is stated, this argument doesn’t conclude with anything as strong as IIC. Even if it were true that in some case what makes it rational for A to perform an action, but not for B, is that A has an indexical belief that B does not have, nothing general follows about the connection between indexicality and agency. All kinds of difference in beliefs can make a difference to what it is rational to do. The fact that action explanation and rationalization contexts are opaque makes this particularly obvious. Consider the following case:

A and B are being threatened by Lex Luthor. They both see a person in the distance. A believes that the person is Superman; B believes that the person is Clark Kent. Now A has reason to refuse to give in to Lex Luthor’s threats, since he has reason to think Superman will save him. But B has reason to submit to Lex Luthor’s threats, since he has reason to think Clark Kent won’t save him. A and B have all the same objective information and desires here, but have reason to act differently.

Notice that there’s nothing distinctively de se and nothing indexical going on in this case. So the Perry phenomenon can’t in itself be an
argument for the distinctiveness of indexicals. At most it can establish
that there is a way that \( A \) is thinking of \( A \) that isn’t the way that \( B \) is
thinking of \( A \). But that’s no more than a familiar instance of opacity,
nothing distinctive about indexicals.

It might be tempting to try to improve on this argument for IIC in
the following way. In any case where someone, \( A \), is characterized only
by her non-indexical beliefs, we can imagine another agent, \( B \), with the
same non-indexical beliefs, but rationally motivated to pursue a different
course of action because of the presence of an indexical belief not shared
with \( A \). In reply to this attempted improvement, we say: Even if this were
true, it doesn’t show that as a matter of fact indexical beliefs are always a
component of what rationalizes an action. If this isn’t obvious, note that
the following may be true (or in any case, we have as much reason to
believe it as we do the corresponding claim about indexical belief): in
any case where someone, \( A \), is characterized only by her non-Superman
beliefs, we can imagine another agent, \( B \), with the same non-Superman
beliefs, but rationally motivated to pursue a different course of action
because of the presence of a Superman-involving belief not shared with
\( A \). For example, suppose \( A \) and \( B \) are alike with respect to their beliefs
and desires about what to order for dessert: it all points towards apple
crumble as the rational choice. Now imagine that \( B \) has the added belief
that Superman will torture everyone who eats apple crumble for des-
sert. The addition of this belief will make a difference as to what it is
rational for each to do: it is rational for \( B \) to not order the crumble, while
it is rational for \( A \) to order the crumble. This point obviously generalizes,
but that’s not a good reason for thinking that “Superman”-beliefs play dis-
tinctive, essential, or even important roles in an account of rational agency.

3.9. The Selection Problem and a Positive
Proposal: The Action Inventory Model

So far this discussion has pointed out flaws in arguments attempting to estab-
lish IIC. Someone convinced by those criticisms might still wonder why it
is the case that “when you and I both apprehend the thought that I am
about to be attacked by a bear, we behave differently” (Perry 1977, p. 494).
More generally, an agent will have indefinitely many third-person beliefs
and desires about the world that don’t result in or rationalize action by the agent. So we need to explain which ones produce action and which ones don’t; we need a filtering mechanism of some sort. Call this “the Selection Problem.” The de se proposal is that the first-person status is the filtering mechanism; only the first-person beliefs and desires give rise to the right sort of intentions, which then gives rise to action/rationalization of action. What can be said about the Selection Problem if we don’t appeal to de se states? Here’s a toy picture that we offer as a replacement for the de se picture:

**The Action Inventory model.** Every agent has a very wide range of third-person beliefs and desires that give rise to third-person intentions, which in turn rationalize or motivate actions (via their recognition). Not all of these intentions are going to produce action, at least in normal cases (perhaps in a god they would). This is because a given agent has an “action inventory”: a range of actions that he can perform. An agent constantly seeks to match his intentions with his action inventory, and when he finds a match, action occurs. When there’s no match, the intention idles, and doesn’t motivate or rationalize action. So the Selection Problem is solved by appealing to the physical or psychological constraints of the agent: only certain actions result because only certain actions were available in the first place.14

Now, this is going to look like a cheat if we think that the actions have to be specified first-personally—such as “I move my hand thus-and-such” or (without specification of subject, as in Lewis) “kick my left leg.” Then the thought will be that there’s first-person content in the action inventory, and it’s only via the appeal to those contents that we solve the Selection Problem. But there’s no reason we need to think of it that way. The action inventory can include “John throws a baseball through a window,” “That hand lifts a glass,” “The car steers out of the skid,” or even “Kripke releases his unpublished papers” (in the inventory of someone other than Kripke). The thought is that we’re embedded agents who have a range of capacities directly to manipulate things in the environment. Some of those things are ourselves and our parts, but others

14 We hope it is obvious that we are taking the Action Inventory model pretty loosely—we’re not particularly committed to anything like a serious model with a switchboard in the head. We don’t need that this model is really right at all, but just that it’s a possibility.
are frequently and infrequently encountered objects, both proximal and distal. To think otherwise is to pursue the “how to move my body” line of thought, which we have already argued against. According to our alternative picture, the belief-desire-obligation-intention sets produce a bunch of inputs—potential actions that are ready to go. Those inputs (i.e., potential actions) then hit the “action center,” which is a big switchboard with a bunch of available actions. If an input matches an available action on the switchboard, an action results. But no one has to “look” to see if there’s a match, and hence no one has to “think about the available actions” in any way, let alone a first-person way. The actions have already been “thought about” by the time they emerge as potential actions (and thought about just as the third-person action “that p”); all that remains is to see if the actions are among the things that can be done.

One way to put our basic point is to say that the Selection Problem is solved by identifying a category of actionable contents. Most likely there isn’t much systematic and general to say about that category. What constitutes actionable content for an agent will depend on the kinds of sensory inputs the agent has—for example, whether he moves around the world only with smell and taste or using sonar and sounds, eyes and ears and fingers, etc. It will depend on the agent’s physical setup: whether he has fingers, claws, a trunk, wings, feet, fins, etc. It will depend on whether the agent’s skills include steering cars out of skids or balancing plates on poles, and whether the agent’s casual causal reach includes having employees publish Kripke’s manuscripts. It is unlikely that there is much to say in general about all content that combines with the propositional attitudes, intentions, and other cognitive states, perceptual inputs, etc. of such varied agents to trigger action. Despite this variability, two negative points are worth emphasizing:

1) It is not necessary that the agent believes or knows that they are within actionable reach. There has to be some part of the agent’s action-generating architecture that recognizes the contents as representing potential actions that are within actionable reach, but that doesn’t have to rise to the level of a belief or be represented as being within actionable reach.

2) Sometimes the actionable contents are represented as being within actionable reach, the agent believes that they are, but in
those cases they are representations of the relationship between the agent and the actionable content. The agent doesn’t have to think of him- or herself, the actionable content or the relation between him- or herself and the content in any special way. They are simply thoughts of the form, “agent \( a \) is actionably related to object \( o \).”

### 3.10. Clarifications, Objections, and Replies

**Objection 1.** *What you say here is incompatible with basic elements of folk-psychological explanations of actions.*\(^{15}\) According to your view, two people can have the same beliefs and desires (the belief that François is about to be shot, the desire that François not be shot, etc.), but act and be rationally motivated to act in different ways. This is an unacceptable consequence.

**Reply.** This objection mischaracterizes our view (which is easy to do since it requires thinking in a somewhat unusual way). On our view, the actions will differ, but not what we are rationally motivated to do. Someone other than François, e.g., Dilip, can be rationally motivated to perform the action that François climb into the tree. Dilip just can’t perform the action. François can perform it, so he does. François is, like Dilip, rationally motivated to perform the action that Dilip call for help. But François can’t perform that act. So François climbs the tree and Dilip calls for help. Now, if you think that folk psychology requires that when two agents have the same belief-desire sets they perform the same action, then we are denying that. But surely that’s not a view worth preserving.

A more plausible principle is that *ceteris paribus*, they act the same. That principle we do endorse. It is because of the difference in their available actions that François and Dilip don’t act the same. Everyone should agree that you can’t explain what someone actually does just by talking about their psychological states. They can have all the psychological states you want, but if they don’t have hands and legs, for example, there are lots of actions they aren’t going to perform. Presumably everyone wants *these*

\(^{15}\) Dilip Ninan has forcefully pressed this concern on us.
sorts of cases to be swept into the *ceteris paribus* clauses. We want a bit more than that. There’s a legitimate debate to be had about how broad the *ceteris paribus* clauses should be in the ideal theory of action, but we see no reason to think that folk psychology as such speaks on where to draw the boundaries.\(^{16,17}\)

**Objection 2.** What about present tense? Maybe agency doesn’t require first-person beliefs, but surely “now”-thoughts are required. Otherwise, it would be a mystery how action could take place. Consider an agent about to turn on the light by flipping a switch. She wants the light to be on, she knows that flipping the switch will turn on the light, she intends to flip the switch, and believes that moving her hand and then finger in a certain way can make it happen. But, the objection goes, these are not tenseless intentions, beliefs and desires. She wants the switch flipped soon (close to now), which is why she now starts moving her hand, and now moves her finger. If she didn’t think, Now the finger is within the reach of the switch, why would she move it then, and not at another time?

\(^{16}\) A comparison with Fregeanism vs. direct reference theory might help here. Fregeanism claims to be better off than direct reference theory because it can explain (e.g.) inferences in terms of contents. The difference in sense between “Superman” and “Clark Kent” explains why inferences from “Superman is \(F\)” and “Clark Kent is \(G\)” to “Superman is \(F\) and \(G\)” aren’t made. Direct reference can’t give such explanations. This looks like a win for Fregeanism, but then we note that Fregeanism can’t give us error-proof explanations: minimally, there’s a “rational agents” proviso; furthermore, there’s just no plausible picture of senses in which there’s no possibility of misidentification of sameness of sense. So both direct reference theorists and Fregeans will have to say at some point that there are just brute facts about whether we “hook things up” mentally. So if you’re looking for error-proof explanations, you’re not going to get it, and if that was your reason for wanting Fregeanism, you should give up on it. Similarly here: no one’s going to get infallible explanations of action, in which having the right mental states just guarantees that everything works out. We’re all going to have explanations which sometimes result in error. See Fodor (1995) for a defense of this line of thought.

\(^{17}\) If you’re a functionalist about mental content, you might think that you just can’t have the sort of story we have, since on that story mental contents are (roughly) just individuated by their potential to produce action. So if François and Dilip act differently, that already establishes that they believe differently. We will return to a detailed discussion of functionalism in Ch. 7, but here we just note that the functionalist point can’t be made quite so easily. Any plausible functionalist story will also need to include various *ceteris paribus* clauses, or its picture of content will bear no recognizable relation to our pre-theoretic mental state ascriptions. So what’s needed here is again an argument that the degree of error-tolerance in the functionalist story needs to be exactly what our opponents say it is, rather than what we suggest it could be.
Reply. Recall the Action Inventory model: our belief-desire-obligation-intention sets produce potential actions that are ready to go; they encounter a set of available actions. If an input’s content matches the content of an available action, an action results. Since no one has to “look” to see if there’s a match, no one has to “think about the available actions” in any way, let alone a de nunc way. Pre-theoretically, this is unsurprising: the actions we care about are complex, consisting of an indefinite number of sub-actions; moving one’s hand, stretching one’s finger, touching a switch, pushing down on one’s finger, etc. To think that each of these will have to be represented not just as an initiated sequence, but that each moment of transition into a new stage has to be represented by a belief-desire-obligation-intention set involving “now” is to attribute unnecessary and baroque representational powers to us.

Objection 3: “Consider this dialogue:

A: Why did Francois duck?
B: Because he didn’t want Francois to be shot, and he thought that if Francois ducked under the table, Francois would not be shot.
A: Wait—did Francois realize that he himself was Francois?
B: I’m not sure. And it doesn’t matter anyway: I already explained to you why Francois ducked.
A: No, you didn’t. For suppose Francois had believed that he was not Francois, but that the man standing beside him was Francois. Then surely, given the beliefs and desires you said Francois had, he would have done something like shout “Duck!” at the man beside him rather than duck himself. What Francois would do given the beliefs and desires you said he had depends on who he thinks Francois is, and how he believes that he is related to the person that he thinks is Francois.

On your view, it seems like B’s second response is adequate, and that there is something wrong with A’s last remark. But what A says seems entirely reasonable. What do they think is wrong with A’s remark?”

18 This objection is from an anonymous reader for OUP.
Reply. Maybe the easiest way to see why this response is misguided is to recall that we started out with two Impersonal Action Explanations. One involved François and his “François”-beliefs and -desires. The second concerned Herman and his “Nora”-beliefs and “Nora”-desires. To see where this objection goes wrong, focus on the second case. In that case, Herman wasn’t motivated to act for reasons having to do with Herman and beliefs about Herman didn’t figure into the explanation or rationalization. It was all about Nora. In such cases, the agent’s possible identify confusions are irrelevant.

Objection 4: Suppose I concede that action could proceed non-indexically. I am really doing a form of high-level empirical work on human action. My conclusion is that it’s a contingent (but maybe physically deep) feature of humans that we act only indexically, because we’re embodied and have to move our bodies. I am doing a form of psychology or cognitive science about very general truths about humans (which is why I don’t have to do actual empirical work). In short, your refutation of the necessity claim leaves open that there’s a deep and significant contingent connection between indexicality and rational agency.

Reply. One way to see why this reply is hopeless is to focus on the parenthetical remarks. There’s no such thing as “doing psychology about very general truths about human beings” without empirical backing. The kind of non-empirical approach found in the Perry–Lewis tradition we are engaging with cannot establish contingent truths about how humans act, how they represent the world around them when they act, how they move their bodies, or how, as a matter of fact, they rationalize their actions, no matter how general. Whether indexicality (or “the de se”) is involved is a very, very detailed question about the implementation of complicated mechanisms in the human head. Armchair reflections about us moving our fingers won’t get us such conclusions; nor will philosophical reflections about generalizations or opacity. In saying this we are not taking a stand on how action mechanisms are in fact implemented in human heads. It would be absurd for us to think that such a complicated empirical issue can be settled that way. We’re just saying that our opponents don’t know how
the implementation goes, and that there are feasible non-indexical implementations.\textsuperscript{19}

\textbf{Objection 5.} This is all very unfair to Perry (1979). He didn’t try to argue for IIC. Perry’s goal was to refute what he called “the doctrine of propositions” (p. 6). Even if your arguments against IIC succeed, it doesn’t touch Perry—he still succeeds in his stated aim.

\textbf{Reply.} We should remind the reader of comments made earlier. It is widely assumed that Perry didn’t just present arguments in favor of the opacity of explanation contexts or against one interpretation of Frege’s view of propositions. The philosophical tradition has taken Perry’s papers (and his 1979 paper in particular) as establishing that some important connection exists between indexicality and agency. This chapter has attempted, on behalf of that tradition, to spell out what that connection might be and how it might be argued for. We agree that our pessimistic conclusions don’t show that Perry failed in his stated goals—maybe the charitable reading of Perry is that he had no interest whatsoever in establishing an interesting indexicality–agency connection. If so, the tradition has been mind-blowingly careless in their Perry exegesis.\textsuperscript{20}

\textbf{Objection 6:} You shouldn’t be focused on IIC—it’s a red herring.

Several readers of earlier drafts ended up agreeing with us that IIC is false and that no good arguments or motivations can be found for it. Some reacted by asking: \textit{Why focus just on IIC?} We end this chapter by reminding readers with that reaction to look back to Section 3.4 where we justify our focus on IIC.

\textsuperscript{19} This is true whether or not you are a functionalist about content. In both cases, stories about contingent features of the theory of content are beholden to a wide range of empirical facts that are never examined in the arguments.

\textsuperscript{20} It is of course not entirely unnatural to read Perry as trying to establish some indexicality–agency connection. All of Perry’s examples concern the transition from not acting to acting, and there is a great deal of discussion of action explanation in the paper, which is not just about action–explanation terminology or contexts.
3.11. Brief Summary of This Chapter

There are six central take-home messages from this chapter:

1) The thought that first-person (or indexical) thought has a philosophically significant connection to agency is, without any doubt, a powerful one. It is endorsed by a very wide range of excellent philosophers working across all fields of philosophy. It is interesting and surprising, therefore, that it is difficult to find any precise articulation of that connection and even harder to find arguments for it.

2) On at least one natural reading of the arguments in Perry 1979, they establish nothing more than that indexicals can’t be substituted *salva veritate* in explanation contexts. The claim that explanation contexts are opaque, however, is not surprising. Those drawn to the thought that agency and the first person are significantly connected want more than that.

3) What they say they want is some version of IIC (and it is hard to see any weaker thesis that would be of interest).

4) What arguments can be given for IIC? Here we are left with a bit of a challenge since there are hardly any genuine arguments to be found in written literature (though, in conversations and Q&A’s, we have encountered quite a few). In light of the lack of arguments, we engaged in a bit of philosophical reconstruction: we tried, on behalf of our opponents, to articulate what we suspect are the (often tacit) motivations for endorsements of IIC.

5) We show that none of those motivations come close to establishing IIC.

6) Finally, we sketched an account of action explanation/motivation that did not give any kind of privileged role to indexicality or first-person thought.