Intentionality: A Problem-Solving Approach

by

Alison Ann Springle

B.A., Philosophy & English, Dickinson College, 2011

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Alison Ann Springle

It was defended on

November 18, 2021

and approved by

James Shaw, Associate Professor of Philosophy, University of Pittsburgh Department of
Philosophy

Rick Grush, Professor of Philosophy, University of California San Diego Department of
Philosophy

John McDowell, Distinguished University Professor of Philosophy, University of Pittsburgh
Department of Philosophy

Susanna Schellenberg, Distinguished Professor of Philosophy and Cognitive Science, Rutgers
University Department of Philosophy

Robert Brandom, Distinguished Professor of Philosophy, University of Pittsburgh Department of
Philosophy

Dissertation Directors: James Shaw, Associate Professor of Philosophy, University of Pittsburgh
Department of Philosophy & Rick Grush, Professor of Philosophy, University of California San
Diego Department of Philosophy.

Intentionality: A Problem-Solving Approach

Alison Ann Springle, PhD.

University of Pittsburgh, 2021

Language, maps, pictures, and models all exhibit "intentionality"—they're "about" something. Many philosophers hold that the intentionality of these sorts of representations can only be understood against the background of a more fundamental kind of intentionality, namely, the intentionality of psychological representations such as beliefs, desires, and intentions.

Alas, the intentionality of psychological representations has proven an elusive creature. I diagnose and critique two widespread assumptions as the underlying cause: (1) That all representations have constitutive alethic success conditions (i.e., truth, veridicality, accuracy, or satisfaction conditions), and (2) That the intentionality of action derives from psychological intentionality.

I develop and defend a unified "practical epistemic access" analysis of representation and a "problem-solving" account of intentional action. The problem-solving account analyses the means-end form of intentional action in terms of the notion of an "application-accountable solution to a practical problem." On this view, an intentional action is a process that functions to non-accidentally satisfy a need for flourishing. Needs for flourishing define "practical problems," and intentional actions are the processes that function to solve them. On the practical epistemic access analysis, representations constitutively function to position agents to *appropriately* intentionally respond to facts. So representations are a kind of ability or capacity for intentional actions. I develop a detailed account of fundamental psychological representations in terms of "embodied instructive representations." These are a hitherto overlooked species of representation. They function to provide *direct* practical epistemic access to facts and have constitutive *appropriateness conditions* rather than alethic success conditions. They are in this way distinct from the more familiar "surrogative" species of representation. The latter function to provide *indirect* practical epistemic access and alethic success conditions figure as a constitutive aspect of their appropriateness conditions.

So, against assumption (2), I argue that the intentional directedness of actions cannot be explained as deriving from the intentionality of psychological states. Quite the contrary, psychological intentionality is fundamentally rooted in the intentional directedness of actions. And in light of this, against assumption (1), I argue that fundamental psychological representations have constitutive appropriateness conditions, not alethic success conditions.

Table of Contents

List of Tables

List of Figures

For Bill

The chief defect of all hitherto existing materialism…is that the thing, reality, sensuousness, is conceived only in the form of the *object* or of *contemplation,* but not as *sensuous human activity, practice,* not subjectively.

~Karl Marx, *Theses on Feuerbach*

It was an otherwise unremarkable day in 12th grade. I was sitting in the hallway, between the cafeteria and the auditorium, again skipping class. I was probably running late, and I certainly hadn't done my homework. But I wasn't going to dwell on that. Highschool was ending, and another phase of life was about to begin. What the hell was I going to do? What could I imagine myself doing? And then it hit me: there's such a thing as a professor of philosophy. And at that moment, for the first time in my life, I felt motivated enough to properly set myself a goal: I was going to go to graduate school. I was going to be a philosopher.

But first I needed to clean up my act.

I had been truant so many times, I barely graduated high school. Not surprisingly, my grades were not particularly good. Alright, fine, they were garbage. So I started afresh at Bucks County Community College. At around the same time, I adopted my cat Peanut. And for the first time, I took a yoga class. Thanks to the support I received from my dedicated professors at Bucks (in particular, Andrew Freeman and Maria McGrath,) and the therapeutic effects of yoga and playing with Peanut, I excelled. After two years, I transferred to Dickinson College on a scholarship.

The philosophy department at Dickinson was small but glorious. Thanks in particular to Susan Feldman and Chauncey Maher, I was able to write an honors thesis on Wittgenstein, and after college, I managed to get into Pitt's philosophy PhD program. I credit Susan and Chauncey for teaching me all of my basic skills in analytic philosophy. Susan advised my thesis, and at one point, after reading my early materials, she noted that I had a tendency to start strong and then trail off. She suspected my problem was anxiety, probably exacerbated by caffeine, so gave me a decaf cinnamon spice tea. While my anxiety returned with a vengeance later, for a time, she cured me. And when I have noticed myself getting anxious while writing, I often think of her and grab a cinnamon spice tea. Chauncey spent I don't know how many hours giving me feedback on draft after draft of my writing sample, and my writing completely transformed under his guidance and example.

Between Dickinson and Pitt, I lived with my parents, Paul and Holly Springle, back in Bucks County, and attended colloquia hosted by the philosophy departments in the general area— U Penn, Princeton, Rutgers, and Temple. In the process, I met Susanna Schellenberg and Elizabeth Camp, both of whom have continued to inspire and support me ever since. After attending some talks at U Penn, I was encouraged to audit some classes. So I did. I learned *so much* from Gary

gave me the chance to visit UCSD where I met my co-advisor, Rick Grush. Rick's faith in me—which he made evident by inviting me to work with him on various projects—has been absolutely invaluable. And teaching my wonderful Pitt students was always inspiring and rewarding, and I was lucky to have Japa Pallikkathayil as my teaching adviser.

I'm also extremely grateful to my fellow graduate students, without whom graduate school wouldn't have been nearly as much fun. I value my relationship with all of the graduate students I've had the pleasure of interacting with grad students and visiting lecturers and reserachers in Pitt Philosophy, HPS, the Center for Philosophy of Science, and CMU philosophy. I've especially enjoyed many wonderful conversations with Vivian Feldblyum, Tom Breed, Stephen Mackereth, and Robbie Howton, and I will always fondly remember late nights in the Cathedral with David de Bruijn. I *treasure* having met my dearest friend, Alessandra Buccella, and my incredibly loving and supportive partner, Konstantin Genin. You two are forever family to me. I'm also grateful for my Yoga Hive family, especially Tracy, Jordan, Dezza, and Kimberly, and wonderful friends I made through the yoga hive, in particular, Caroline Record and Kelly Regan. Pittsburgh helped me grow intellectually as well as spiritually, and I am forever grateful to my amazing therapist, Dr. Lisa Maccarrelli, for helping me heal myself. You've all changed my life for the better. I learned so much from all of you and you made Pittsburgh absolutely magical.

Part of this dissertation was composed outside of Pittsburgh: with my parents in Bucks County, with my beloved Aunt Ann in Collegeville, PA, and finally in Tübingen, Germany, with my partner and incredible colleagues—Hong Yu Wong and Krisztina Orbàn in particular. Writing a dissertation makes one, or at least it made me, feel extremely vulnerable. Feeling vulnerable can be bad for forming or deepening relationships insofar as it makes one want to hide, but when hiding isn't an option, it can be a marvelous thing indeed. I have become closer with my partner, my

parents, my aunt, and my colleagues through the experience of writing this dissertation. Indeed, my dissertation brought me to my aunt's house in the summer of 2020, and in the time we spent living together we really got to know one another. I'm incredibly grateful for this. There's something strikingly philosophical about getting to know one's own family better. After all, philosophy is in large part about making the familiar unfamiliar so as to gain a new and deeper familiarity with it. I've known my parents, aunt, uncle, and grandparents my whole life, yet I've known so little about them. It has been especially inspiring to learn more about my grandma Betty's passion for genealogical research, and discussing various philosophical topics with my mother, father, aunt, and uncle. It saddens me that my grandpa Dick passed away before I was mature enough to get to really get to know him, as I am certain we'd have had a lot to talk about. He was an engineer/inventor, and he loved explaining things and conceiving of new possibilities, especially if he would bring them to life in the workshop. When he found out I was going to do a PhD in philosophy, he gave me some philosophy books a professor he used to drive around in college had given him, including early editions of books by Whitehead, Santayana, and Dewey. When I was at my aunt's in the summer of 2020, she showed me some notes of his she had found tucked away in one of his books—hardly legible, distinctively philosophical scribbles. I felt perhaps the most connected to him I ever had, and feeling this connection gave me courage. In his *Seventh Letter*, Plato describes the transmission of philosophical knowledge through philosophical exchange. While Plato does this in the context of explaining why he does not put his thoughts down to paper, I have found his description apt to my experience of reading my grandfather's notes, and likewise reading Aristotle's *De Anima*, and other philosophical works: it is as if "suddenly a light, as it were, is kindled in one soul by a flame that leaps to it from another, and thereafter sustains itself."

Before moving to Pittsburgh, I adopted my darling kitty, Arya (a.k.a., Ruchka). Peanut was happy at my parents' house where she could go outside, and I didn't want to move her. So I needed a house, and Arya needed a home. A year later, she was joined by her sister, Pearl (a.k.a., Bill). I cannot imagine my life without my girls. I could not have gotten through graduate school without them. I spent a lot of time alone, in my apartment, reading, writing, or grading. I would have been so lonely if it were not for my girls who were inevitably on, next to, or very near me. Typically, during crunch time—near a deadline or the end of a semester—my diet suffers. Sometimes I took to just eating Brie or Port Salut, and I have fond memories of my girls climbing all over me to share in the cheese. They were a constant source of comic relief and welcome distraction. They kept me grounded and at the same time, there were often my muse. I can't tell you how many drafts of papers and chapters started with examples involving Bill and Ruch (indeed, they show up in examples in this dissertation). When I was feeling bad about work and other things, I'd think about how sweet and well behaved my babies are, how well they have traveled from apartment to apartment and even across the sea with me. I felt proud of them, like I must have done something right to have such wonderful cats.

I was *devastated* when, six months after moving to Germany, Bill suddenly fell ill and passed away. Part of me died with her, and in truth I wasn't sure I'd ever finish this dissertation. Philosophy helped me get through this period—not my work, which I avoided like the plague, but rather the ancient philosophical idea of reincarnation. About four months after Bill's passing, Bennu (a.k.a. Bihui) entered my life. Bennu is her own cat, but somehow remarkably like Bill, including in her interactions with her sister, Arya. And over the subsequent five months, I resuscitated my dissertation—rewriting large portions of it, though the soul of it remained the same. Now I'm proud to say that it's done, though I expect to continue working on and revising

the ideas it develops for many years. I like to think it's got Bill's personality—bold, original, at times a little sassy. Alas, it isn't nearly as cute. In any case, this dissertation is dedicated to you, my sweet, sweet Bill.

1.0 Introduction: A Problem-Solving Approach to Intentionality

1.1 The Origins of Intentionality

This dissertation attempts to provide a broadly naturalistic analysis of intentionality.[1] Importantly, however, the analysis is not reductive in the sense that I do not attempt to reduce intentionality to something that is not intentional. Instead, my analysis constitutes an account of the *origins* of intentionality. My purpose is to illuminate the kinds of intentionality that have been of interest to philosophers-- in particular, the intentionality of mental states, but also of representations more generally, and of actions-- by way of exposing the most fundamental kind of intentionality and their relationship to it. Since the analysis is not reductive, it might not meet everyone's standard for a naturalistic analysis. A further reason it might not do this is that it is committed to irreducible facts about flourishing.[2] Another way to put the point is to say that my account is teleological in a more loaded sense than most teleosemantic accounts-- which are often considered thoroughly naturalistic-- purport to be. By my lights, the goal of an analysis of intentional phenomena like representation and action is not to show how they could emerge out of a world devoid of normative facts, but rather to clearly conceptualize the relevant normative facts and how things like intentional actions and representations are related to them. So my objective is not to satisfy the sorts of devout, austere naturalists who consider it their task to reduce away

---

[1] I will bracket the topic of consciousness in this dissertation. I am of the view that when we explain fundamental intentionality, we explain consciousness. But there isn't time or space to argue for this general claim in this dissertation nor the more specific claim that my account of intentionality successfully accounts for consciousness (or phenomenology).

[2] That said, I suspect a more austere naturalist could rework much of the account without (at least explicitly) committing to irreducible facts about flourishing

normativity, but rather to provide at least the beginnings of what I think a lot of people who consider themselves naturalists are looking for: an account of intentionality that illuminates its nature, how the different ways it manifests are connected, and how it's the kind of thing that can be realized by the sorts of systems described in physical science.

According to my "problem-solving account," the most fundamental kind of intentionality-- the kind of intentionality that grounds all others-- is that of a need for flourishing, or what I call a "practical problem." The next kind of intentionality is that of an intentional action, which I'll define as a kind of "solution" to a practical problem. I'll propose to understand representational intentionality as an aspect of the intentionality of action. Specifically, I'll argue that representations constitutively function to provide "practical epistemic access (PEA)" to facts, and that to have PEA to a fact just is to be able to intentionally respond to it.[3] And this, in turn, just is for an intentional action that is itself intentionally directed at the relevant fact to be potentiated. I'll distinguish between "direct" and "indirect" PEA and between two corresponding species of representation, namely, "embodied instructive representations (EIRs)" and "surrogative representations (SRs)."[4] EIRs are constituted by activations of embodied abilities-- "problem-

---

[3] For the purposes of this dissertation I want to remain as neutral as possible on the ontological status of facts. But to give the reader an idea of what I have in mind: In the case of what I'll call "embodied instructive representations (EIRs)," "facts" are paradigmatically particulars, arrangements of particulars, situations, property instances, or objects. But they may include things like kinds (or instances of kinds) and possibilities (or powers of things). In the case of what I'll call "surrogative representations (SRs)," facts may include the sorts of things we might call "abstract objects," e.g. public interpretations of facts (descriptions, propositions), objects that are imaginatively created, and objects that are posited to explain other objects (models, theories).

[4] Although direct-PEA is essentially intended as an account of acquaintance or the "direct reference" of "thought" (scare quotes because thoughts are not fundamentally propositions on my view), the distinction between direct- and indirect-PEA does not map neatly onto the distinction between direct and indirect reference in the philosophy of language. An SR, i.e. the kind of representation that provides indirect-PEA, may provide direct reference-- in the philosophy of language sense-- to an object in that it does not stand-in-for the object by way of a description but rather by way of same causal history plus conventions. There are a number of different ways SRs can stand-in-for objects and thus provide indirect-PEA to them. I'm inclined to think that the fact that linguistic SRs refer (when or to the extent that they do) is ultimately because they provide indirect-PEA to whatever they refer to. But there's still a question of how they do that, and debates about reference in the philosophy of language address this and related questions.

2

solving dispositions" -- for types of intentional actions. They have "*de actu*" contents which constitute "generative" modes of presentation of facts and are successful or correct when "appropriate." SRs, on the other hand, are constituted by vehicles agents access via their EIRs.[5] They constitute often extended abilities for special kinds of actions that solve "proximity problems" and alethic conditions, i.e., truth, veridicality, accuracy, or satisfaction-- are an important element of their success or correctness. Although philosophers tend to talk exclusively about SRs, or at least about representations with properties that are distinctive of them, I'll argue that EIRs are more fundamental.

To bring out what is distinctive about my project, it is helpful to compare it with Tyler Burge's project in his 2010 *Origins of Objectivity*. One obvious difference between these projects is that Burge aims to illuminate the origins of *objectivity* while I aim to illuminate the origins of *intentualit*y. But these projects are closely related,[6] so much so that at various points in his book, Burge characterizes what he is doing as giving an account of the origins of intentionality. According to Burge, it all starts with perceptual representation. Burge's origins story takes place against a proto-intentional backdrop of what he considers pre- or proto-intentional phenomena, specifically, primitive agency and sensation. Genuinely intentional phenomena begin when objectivity becomes the aim of a system capable of primitive action and sensation. When this

---

[5] A single vehicle (or "surface format" or "sign design" to borrow from Millikan (2017) who borrows from Sellars) or vehicle type may constitute multiple SRs: it may have multiple indirect-PEA functions at once. Surrogative content goes with indirect-PEA functions, so a single SR vehicle can have multiple surrogative contents. It is also the case that distinct vehicles or vehicle types can have the same indirect-PEA function. For example, the sentence "snow is white" and the sentence "schnee ist weiss" constitute different vehicles but convey the same SR contents (propositions are SR contents). So the relationship between vehicles (tokens and types) and SRs may be many-to-one and one-to-many.

[6] To some extent I will explore the topics of objectivity and subjectivity in relation to representation in Chapter 2, but neither that relationship, nor the nature or origins of objectivity, will be a significant (explicit) focus of this dissertation.

happens, genuine perception, representation, and intentional action are born. On Burge's view, representations mark the difference between primitive agency and intentional agency.

In contrast, as we just saw, my origin story begins with creatures (life forms) with needs, i.e., practical problems, and with abilities to solve those practical problems. Accordingly, what Burge considers proto-intentional, i.e., primitive agency and sensation, I consider intentional. I agree with Burge that the kind of intentionality philosophers care about-- representational and agential intentionality-- haven't yet entered the scene, but their intentional roots-- solutions to practical problems-- have. My view also agrees with Burge's that when intentional action and representation enter the scene, they do so at the same time. However, on my view, representation is not what explains the transition to intentional actions.

The root of my disagreement with Burge centers on how we understand representation. For Burge, representations have veridicality conditions essentially. Therefore, they only enter the scene when objectivity becomes part of the aim of a system. Because Burge thinks that veridicality conditions are constitutive of representation, and for reasons I'll get into in the chapters to come (in particular, Chapter 5), Burge argues that it is probably not possible to provide a reductive account of representation (308). So, he does not attempt one. Instead, he treats representation as primitive: he defines a representation as something that functions to represent veridically. Burge describes some background conditions for representation primitive agency, and he describes an empirical mark of the most primitive kind of representation, namely, the perceptual constancies. But he does not think that the concept of representation that matters to philosophers can be further reduced. In addition, while representation (perception) and intentional agency co-emerge, representation is what makes the difference; when representations of perceptual targets and goals guide agency, agency is intentional. In contrast, I believe that the concept of representation can be

reduced and rather than reducing intentional action to action plus guidance by representations, I'll

analyze representations as potentiations of intentional actions.

My disagreement with Burge can be further crystalized around the key issue of

representational *telos*. According to Burge,

> There is...a root mismatch between representational error and failure of biological function… Fitness is very clearly a practical value...Explanations that appeal to biological function are explanations of the practical (fitness) value of a trait or system. But accuracy is not in itself a practical value. Explanations that appeal to accuracy and inaccuracy— such as those in perceptual psychology—are not explanations of practical value, or of contributions to some practical end. (301)

> It is part of the nature of perceptions and beliefs that they have the representational function of representing veridically. These states succeed or fail, in one respect, depending on whether their representational contents are veridical. It is part of the nature of intentions, willings, and non propositional representational actional states that they have the representational function to help induce fulfillment of their representational contents...Of course, beliefs, perceptual states, and actional representational states can have other functions besides their representational functions. Some have practical or biological functions. Fulfilling these functions marks other sorts of teleology, distinct from the teleology involved in representational function. (75)

Much of what I do in this dissertation is aimed at showing that Burge is wrong:

representational *telos* or function is fundamentally practical. However, part of this will involve

bringing attention to something I think Burge is absolutely right about, namely, that veridicality

or accuracy conditions (and alethic conditions more generally) are not *just* practical success

conditions. Accordingly, in claiming that representational *telos* is fundamentally practical I'll be

claiming that representational success conditions are not fundamentally (or generally) a matter of

veridicality (or alethic success or correctness). Instead, I'll argue that practical success, and thus

representational success, is fundamentally a matter of appropriateness. This does not mean that

representations *just are* actions, nor that alethic conditions never get a hold on them. Rather, I'll

argue that representations are potentiations of actions, where potentiations do not entail action

performances. I'll also suggest that the Practical Epistemic Access (PEA) analysis of representation makes available a fairly inflationary account of alethic success conditions as an aspect of appropriateness conditions (and values) for surrogative representations (SRs), i.e., indirect-PEA. (See Chapter 3).

There's a sense in which my strategy in this dissertation consists in turning teleosemantics accounts of representation "inside-out." This is because I will define representations in terms of functions of potentiating (and thereby enabling) intentional actions (i.e., types of practical epistemic access), and I will employ broadly teleosemantic tools to define intentional action. I will say that an intentional action is a process/mechanism that (constitutively) functions to satisfy a need (solve a practical problem). To say that a process/ mechanism functions to satisfy a need is to say that it exists for this purpose. Clearly, this purpose cannot fundamentally be one that's created by the intentional actions of agents since this purpose is supposed to ground intentional actions. Teleosemantic accounts of representation face the same sort of issue-- they can't explain psychological intentionality in terms of functions that depend on psychological (and hence agential) intentionality. So, they appeal to some notion of natural functions. Some (e.g., Millikan 1984; 1989) argue that natural functions are selected functions-- the sense in which something exists for some purpose is just the sense in which a type of thing did something and its doing that thing is what explains the reason it was reproduced. In other words, only descendants of that thing have the function to do what their ancestor simply did. But there are other ways of explaining what it is for something to exist in order to do something-- other ways to cash-out natural functions. [7] I'm not going to take a position on how to cash-out the source of the relevant functions. For the

---

[7] See e.g., Wright (1973, 1976), Cummins (1975).

time being, at least, I prefer to keep an open mind. I'm even open to the possibility that natural functions might depend on *types* of agents, or at least, species of organisms. What matters, I think, is that natural functions do not depend on the intentions of *particular* agents or particular agent-like entities (like an anthropomorphic God).

1.2 Broader Methodological Orientation

Intentional actions and intentional states are both topics of intense philosophical study. An important reason for this is that both seem to pose the mind-body problem. Specifically, both seem to have a material or physical causal dimension in that they seem to be both susceptible to material causal impacts and capable of producing such impacts. At the same time, neither seems to be exhausted by this dimension: both intentional states and intentional actions have special *intentional* properties. These properties can be recognized in different ways. First, intentional actions are susceptible to a special form of explanation-- teleological explanation-- and intentional states appear to figure in those explanations essentially. Second, both intentional actions and intentional states are "intentionally directed" at something beyond themselves: actions are directed at the satisfaction of ends, and intentional states are typically thought to be directed at the satisfaction of their contents. In connection with this, third, both intentional actions and intentional states have broadly normative properties. Intentional actions, but not other events, can be appropriate or not, or rational or not, with respect to their ends, and intentional states are typically deemed successful (true, veridical, etc.) when their contents are satisfied. As I understand it, intentional states and

intentional actions pose the mind-body problem in that they raise the question of how their material properties (borrowing from Sellars 1962) "hang together" with their intentional properties.

Now, there are different responses to this problem. One response denies that it's really a problem at all. Material properties and intentional properties just do hang together, and there's no positive or constructive philosophical work to be done to show how this is so. The only philosophical work that's to be done is therapeutic-- a matter of pulling the wool off our eyes or finding our way out of the fly jar that gave rise to the appearance of a problem. I'll call this "the therapeutic response."

Another response holds that it's really the intentional properties that are the problem-- since those are the weird ones. The philosophical task is to show that they aren't weird after all by showing that they can be reduced to un-weird material (physical) properties or reduced to useful fictional ways of talking about them. I'll call this "the reductive view"

These views may be thought of as the opposing ends of a spectrum. The reductive view holds that intentional properties are weird and aims to make them not weird by showing that they are actually something else (material) or actually not really anything at all (just ways of talking about real things at a useful level of abstraction). In contrast, the therapeutic view says that the idea that intentional properties are weird is an illusion. They're just different. The idea that we need to make them uniform is confused.

I find myself most sympathetic to a third kind of response that sits somewhere in the middle of this spectrum which I'll call the "rich correspondence view." Like the therapeutic view, it denies the need and possibility of reducing intentional properties to material properties like any others. But like the reductive view, it thinks philosophy has a bigger job to do. This job consists in making sense of how it is that intentional properties and material properties systematically correspond. So,

the rich correspondence view seeks a substantive account of intentional phenomena that illuminates interesting continuities and discontinuities in the different ways such phenomena manifest in nature. It therefore tends to favor a bottom-up approach to theorizing, i.e., theorizing that starts with intentional phenomena as they manifest more generally and fundamentally and builds towards illuminating the distinctive ways they manifest in rational animals as opposed to approaches to theorizing that begin with the distinctive ways intentional phenomena manifest in rational animals.

One motivation for the rich correspondence view is to guide science by developing intentional (in particular, psychological) concepts (i.e. concepts of the intentional/mental) that are rich enough to impose structural constraints on how scientists ought to conceptualize the material properties that realize the intentional properties the intentional concepts describe.[8] This brings out an important contrast between the rich correspondence view and the reductive view: where the reductive view wants to bring intentional properties into the scope of scientific concepts which reference only material (efficient causal) sorts of properties, the rich correspondence view wants to articulate intentional concepts which refer to intentional properties and distill those concepts in sufficient detail to be more usefully superimposed on scientific concepts. What do I mean by "superimposed"?

Here's the basic idea: Physical science gives us a picture of the properties of matter. The sorts of structures that are present in this picture are so general they'd be difficult to discern. You might imagine a picture of variously colored dots. If you relax your gaze, you get the sense that

---

[8] I consider the rich correspondence view a version of the American and Cambridge pragmatist conceptual engineering approach to philosophy.

various patterns are present. But that's about it. Philosophy, in contrast, gives us a picture of forms -- the nature of different kinds of structures and the kinds of laws or principles that govern them where at least some of these structures can be *realized* by material structures. Then sciences like biology and psychology fit these pictures together--- they superimpose or overlay the philosophical pictures of various forms onto the physical scientist's picture of matter. Biologists and psychologists have the job of figuring out how these pictures fit together-- how matter (at different "levels") realizes these forms. The rich correspondence view aims to guide psychological science by producing more detailed pictures of psychological forms.

A second motivation for the rich correspondence view is to articulate rich (detailed) intentional concepts so that we humans may better understand ourselves and one another. The basic idea is that, by further developing our intentional concepts, we can better articulate theories of human nature. We may bring new theories to light, and we may supply existing theories with a stronger psychological conceptual backbone than they previously had. This, in turn, may help those theories compete with more dominant theories that depend on certain assumptions about our psychological makeup that themselves have gained a position of social-epistemic dominance. This is because the further articulated intentional concepts might provide the resources to systematically challenge such dominant assumptions.

One might hope, as I do, that because the rich correspondence view is an in-between view, proponents of the more extreme views will welcome it as a kind of synthesis, and everyone will love it. One might hope for this, there's no doubt that I do, but one probably shouldn't bet on it: there's a good chance that committed proponents of the more extreme view ain't gonna like it at all.

On the other hand, I think a good number of philosophical folks are not so much committed

to either sort of view as they are simply caught somewhere in-between them, finding themselves

really quite attracted to certain *parts* of *both* of them, sometimes more of one than the other. Such

folks might feel themselves pressured, perhaps in part due to the influence of their most proximate

philosophical peers, to simply care less about parts of one so that they can finally be admitted into

one of the philosophical clubs. I've often found myself in this position, and I venture to guess I'm

not alone. Perhaps, then, it's realistic to hope that the rich correspondence view might capture

enough of the features of the other two views to provide a safe home for those who, like me, have

often found themselves violently torn between extremes, unable to find a stable refuge in either of

them. I consider such folks (assuming they exist) my philosophical kin, and in this dissertation,

they'll be my primary audience.

1.3 Theses & Antitheses

This dissertation will be framed around rejecting and replacing the following two theses:

The Derivation Thesis: The intentionality of intentional actions derives from the intentionality of intrinsically intentional original representations, i.e., beliefs, desires, and intentions.

The Descriptive Thesis: The essence of representation is "aboutness" understood as a kind of directivity that's exemplified by declarative sentences, pictures, models that stand-in-for their represented and have "alethic"-- i.e. truth, veridicality, satisfaction, or accuracy-- conditions; a representational content -- as distinct from attitude (force)-- always describes, portrays, or otherwise puts forward or presents some fact as such, i.e. as what is the case.

These theses are connected, and every chapter will explore aspects of their connection, more or less explicitly. Each chapter will also criticize both theses, though, with the exception of Chapter 2, the descriptive thesis will be the more explicit target.

I'll ultimately argue that the descriptive thesis gives rise to problems in the context of illuminating original representation, that on inspection it appears not be particularly well-motivated, but most importantly, that it is not the only game in town. I will try to motivate dropping the descriptive thesis in favor of the thesis that the essential feature of representation is its function of providing practical epistemic access (PEA). According to the PEA analysis of representation I'll propose,

> **PEA Def. Representation:** R is a representation of (about, directed at) a fact X iff R constitutes a form of PEA to X, where R constitutes a form of PEA to X iff, in having R, an agent is positioned to produce an intentional action (in a sense that extends to non-human animals) in response to X.

This is an analysis of the *genus* representation. It says that anything that represents provides PEA; providing PEA is necessary and sufficient for representation. According to the PEA analysis, the fundamental norm of representation is appropriateness: representations function to potentiate (or to be abilities for) appropriate intentional responses to facts. I'll suggest that alethic success conditions might be analyzed as an aspect of the kind of appropriateness condition (and value) that goes with one of the two major forms of PEA. So, the idea that representations must have alethic success conditions may result from conflating these forms. The two major species of PEA are Direct-PEA and Indirect-PEA:

> **Direct-PEA:** An agent has direct-PEA to X if she's positioned to produce an intentional action that is a direct response to X, where a response is direct when its immediate target is X, that is, when the action consists in interacting with X itself.

> **Indirect-PEA:** Agent has indirect-PEA to X if an agent is positioned to produce an intentional action that is an indirect response to X, where a response is indirect when its

immediate target is something that takes the place of (stands-in-for) X-- something one acts on directly so as to act on X indirectly.

And corresponding to these different species of PEA are different species of representation:

Embodied Instructive Representations (EIRs) provide *direct*-PEA. EIRs have *de actu* contents (Chapter 1 and 4) which constitute *Generative Modes of Presentation* of one's own actions (Chapter 4) -- the actions one is in a position to produce or in the process of producing-- and the particulars at which those actions are directed. An EIR is a representation that stands-*for* an intentional action and the facts at which the action is directed by virtue of being the potentiation (determination or priming) of that very action. For EIRs to present us with their targets, it is neither necessary nor sufficient that we be presented with EIR vehicles. Rather, EIRs present us with their targets by virtue of being constituted by "vehicles" *in us*--vehicles in virtue of which we are causally positioned to intentionally respond to the relevant targets, i.e., our problem-solving dispositions (Chapters 4 and 5), as these realize embodied instructive concepts applications of which constitute EIRs.

Surrogative Representations (SRs) provide *in*direct-PEA. SRs are the philosophically familiar species of representations. SRs are *surrogative modes of presentation of facts*, which is to say, they are broadly *propositional* or *attributive modes of Presentation*. They stand-*in*-for facts thereby potentiating (or determining) some indirect action involving those facts. For SRs to present us with the targets they-stand-in for, we must have PEA to their vehicles.[9]

Where EIRs stand-for the direct actions of which they are the priming or potentiation, SRs stand-in-for particulars (or facts) which may include but need not be actions, and thereby potentiate

---

[9] I consider it a wide open question how exactly this works for psychological SRs.

indirect actions with respect to those particulars.[10]   A representation that provides indirect-PEA, that is, an SR, potentiates an indirect action that consists in an interaction with a vehicle Y that stands-in-for a target X and that constitutes an agent's being in a position to produce an intentional action that is an indirect response to a situation X; the action is itself directly directed at Y, and the interaction with Y constitutes an indirect interaction with X. Accordingly, there can be EIRs without SRs, but there cannot be SRs without EIRs. SRs piggyback on EIRs because an agent must have EIRs giving her access to the vehicle Y of an SR that stands-in-for X in order for that SR to provide her with indirect PEA to X. An EIRs' success conditions are identical to the conditions under which the action it instructs (of which it's the potentiation or determination) is locally appropriate. [11]

SRs, on the other hand, have truth, veridicality, accuracy, and satisfaction conditions. Because SRs are genuine stand-ins, they can potentiate non-accidentally successful indirect actions with respect to their targets iff they bear certain sorts of structural (though they may be heavily conventional) correspondence relations to the facts for which they stand-in. Consequently, I'll suggest that alethic standards of success such as truth, veridicality, accuracy, and satisfaction may be understood in terms of the properties an SR must have to constitute the potentiation of an appropriate indirect action. (See in particular Chapter 3.)

The descriptive thesis, and the propositional attitude model it inspires, treats all representations as species of SRs-- all representation is a matter of something standing-in-for

---

[10] A single vehicle (e.g., a single diagram or sentence form) will often potentiate a variety of different indirect actions and therefore realize multiple SRs.

[11] That said, EIRs can involve sub-EIRs, and executional success and failure may be a matter of whether the sub-actions composing a token of a type of action (locally appropriate or not) are appropriate. In addition, whether the action an EIR instructs is performed (whether it moves from being potentiated to being actualized) depends on whether it is selected, where selected EIRs are evaluable for both local and global appropriateness.

something. If the PEA analysis is right, then this is a mistake. Only SRs are stand-ins; EIRs are not, and SRs piggyback on EIRs. Therefore, in the light of the PEA analysis, the descriptive thesis conflates the surrogative *species* of representation with the genus and treats properties that are *distinctive* of that species as properties that are constitutive of representations *as such.*[12]

Because the PEA analysis defines representations as potentiations of (or abilities for) intentional actions, it is obviously inconsistent with the derivation thesis. Accordingly, one of the ways in which I will defend this analysis is by arguing against the derivation thesis. In arguing against the derivation thesis, I'll argue in favor of views, such as the problem-solving account of intentional action, that hold that the intentionality of actions is intrinsic-- a matter of their form-- rather than etiologically derived. My account of representation, and in particular embodied instructive representations, will draw heavily on my problem-solving account of action.

---

[12] The derivation thesis also threatens to undermine the PEA thesis by raising the threat of circularity: circularity threatens the PEA analysis if we cannot account for intentional actions without appealing to intentional states. Indeed, since the PEA Thesis proposes to define representation in terms of potentiations of intentional actions, it would be, as already noted, viciously circular if coupled with the Derivation Thesis. What's more, the PEA Thesis entails the denial of the Derivation Thesis in that it holds that representational intentionality is essentially an *aspect* of agential intentionality. Afterall, if representational intentionality is essentially an aspect of agential intentionality, agential intentionality can't very well be derived from representational intentionality. To be so derived, representational intentionality would have to be independent of and prior to agential intentionality. But clearly, if representational intentionality is defined as an aspect of agential intentionality-- if it's defined as the potentiation of an intentional response to a fact-- then representational intentionality is neither independent of nor prior to agential intentionality. They are, rather, two sides of the same coin. The arguments in Chapter 2 will suggest that the PEA analysis is incompatible with the derivation thesis, but it will ultimately suggest that this is a feature rather than a bug. This chapter will argue that theories of intentional action that appeal to intentional states face the opposite incompatibility issue: they can only be coupled with accounts of intentional states that don't reference intentional actions. The PEA analysis will turn out to be in the better position: In chapter 2 I'll argue that one of the problems for causal theories is that there's good reason to think that such accounts won't work. This is just one of the reasons I'll discuss to prefer teleological theories of actions that don't explain them in terms of intentional states. As noted, I'll even develop and defend a version of the teleological theory-- the problem-solving account of intentional action-- as well as substantive conception of psychological representations that coheres with a teleological theory of intentional action in terms of what I'll call "intentional action acorns."

1.4 Clarifications

My dissertation will employ a lot of technical jargon. I apologize for this. To help the reader, I have created a glossary which may be found at the end of the dissertation. In addition to this jargon, this dissertation will employ some common terminology. This terminology has the potential to confuse since the same words often mean important different things in the mouths of different philosophers. So, I hope the following clarifications will prevent confusion.

Representations

Some philosophers use the term "representation" fairly loosely. They are happy to ascribe representations to all sorts of animals, their subsystems, machines, and maybe more. I'm happy to allow that a subject's visual system, a computer, a thermostat, perhaps all varieties of animals and maybe even fungi, etc. "represent" in some derivative, deflationary sense. However, in this dissertation I'm exclusively interested in the core (non-derivative), inflationary or "robust" sense of representation. Like Burge (2010), I consider robust representations the kinds of representations whose distinctively representational properties are ineliminable to their explanatory role, whether in science or everyday life. While I don't agree with Burge's assessment of what the distinctively representational properties are, I do agree that they belong uniquely to the sorts of representations that are only properly ascribed to whole individuals (what Burge calls the "individual-level" (e.g., 376) where such representations are essentially connected with subjects' perspectives (they're "constitutively perspectival" (e.g., 391)) and agency (326).

I will assume a *broadly* Fregean view of representation in that I take it that a representation is constituted by applications of something concept- or sense-like (I'll clarify how I understand concepts below). That is, for reasons I'll discuss (to some extent) in Chapter 1, I take it that all

representation is a kind of representation *as*-- all representation involves "modes of presentation." I emphasize that my view of representation is "broadly" Fregean because I'll argue that there are two overarching species of modes of presentation-- one surrogative, the other generative. And I will focus on EIRs/ *de actu* contents, i.e., applications of "embodied instructive concepts," which constitute generative modes of presentation. For reasons I'll explain in Chapter 4, although in some ways similar to singular or *de re* mode of presentation (or demonstrative contents), EIR/*de actu* contents (or their constitutive embodied instructive concepts) are subjective in ways that preclude them from counting as any kind of Fregean senses. Surrogative modes of presentation may, at least in some cases (where they constitute inter-subjective reasons-- see Chapter 3), be understood as being or as being at least in the vicinity of Fregean senses proper. But I have not yet developed enough of an account of surrogative representations to argue one way or the other.

It should be noted that my account of embodied instructive representations and their constitutive embodied instructive concepts will incorporate aspects of "enactivist" accounts of sense-making.[13] Enactivism is typically thought to be at odds with representationalism. Since it won't be discussed in the dissertation, it's worth briefly noting that a motivation for developing the PEA analysis of representation which recognizes embodied instructive representations (EIRs) as a species is that it arguably provides an ecumenical account of representation. This is because the PEA analysis and in particular the account of EIRs shows that one can be a representationalist while at the same time acknowledging and incorporating enactivist insights.

Similarly, in rejecting the descriptive thesis and arguing that fundamental representations are EIRs, my account of representation might be attractive to some relationalist theories of

---

[13] In particular, Hurley (2002) & Thompson (2010).

perception. Relationalists object to the view that perception constitutively involves representations with alethic success conditions (what Travis (2013) calls "allorepresentations"). On the view I propose, fundamental representations such as perceptual representations are EIRs, not SRs, and EIRs do not have constitutive alethic success conditions (though alethic conditions can be associated with them--see Chapters 4 & 5). EIRs provide an account of the fundamental intentional relation subjects bear to objects and shows that this relation can be direct and at the same time representational (see Chapters 1 & 3). So, it's possible that an EIR account of perceptual representation will provide a middle ground both for representationalism and relationalists about perception and for enactivists and representationalists about cognition more generally. [14]

Concepts

Embodied instructive representations (EIRs) are constituted by applications of what I call "embodied instructive concepts." Since EIRs are the most fundamental kind of representation, this means that, on my view, the most fundamental kind of representation involves concepts. However, there are different views of what concepts are, and on some of them, embodied instructive representations would not count as concepts. Indeed, embodied instructive representations and their *de actu* contents have similarities to various theories of nonconceptual content. Theories of nonconceptual content are motivated by the idea that some psychological states can represent the world even though the subjects of those psychological states do not possess the concepts required to specify their content. As Peacocke (1992) notes,

> The most fundamental reason [for endorsing non-conceptual content] lies in the need to describe correctly the overlap between human perception and that of some of the nonlinguistic animals. While being reluctant to attribute concepts to the lower animals,

---

[14] I argue as much in my forthcoming paper, "Seeing What To Do: Embodied Instructive Representations in Vision" for the *Synthese Library Series*.

many of us would also want to insist that the property of (say) representing a flat brown surface as being at a certain distance from one can be common to the perceptions of humans and of lower animals. (614)

What Peacocke considers the most fundamental reason for endorsing non-conceptual content is my reason for introducing embodied instructive representations/ *de actu* content. I take it that only rational self-conscious creatures engage in the broadly theoretical practice of specifying the contents of their psychological states, and that when they do so they necessarily employ surrogative representations. In developing an account of embodied instructive (*de actu*) contents, I will specify such contents by means of surrogative representations and associated concepts. But part of the point of introducing such contents is to distinguish them and the embodied instructive concepts they involve from surrogative concepts we use in talking about them. In other words, non-human animals do not have the concept of an embodied instructive concept nor the concept of *de actu* content. But they have embodied instructive concepts and their representations of the world are embodied instructive representations which have *de actu* contents.

My preference, then, is to distinguish between different kinds of concepts-- embodied instructive on the one hand and surrogative on the other-- and, corresponding to them, different kinds of contents -- *de actu*/ generative contents on the one hand and surrogative contents on the other-- rather than to introduce a notion of "non-conceptual content." But I don't wish to argue over how to use the word "concept." So, if one prefers, one might think of embodied instructive representations as having non-conceptual contents (one may think of *de actu* concepts as nonconceptual contents). Either way, it's worth noting that embodied instructive representational (*de actu*) contents bear some similarities to accounts of the nonconceptual content of perception (even though *de actu* contents are not merely intended to account for perceptual content).

For instance, like EIRs, on Evans's (1982) account of nonconceptual perceptual content, it essentially involves information links and behavioral dispositions. As Cussins (1990 p. 7) notes, Evans saw that the cognitive significance of singular contents could be captured, and could only be captured, by the theorist's referring to abilities to grasp or otherwise to locate and track particulars through space and time and to be selectively sensitive in judgment and action and memory to changes in a particular's features. For Evans, these abilities are not available to the subject as the content's referent, but they are available to the subject as the subject's experience-based knowledge of how to act on the object and respond to it. While I agree with Evans that such abilities are not fundamentally available to subjects as referents of true/false thoughts, on my view, exercises of abilities (intentional actions) constitute our fundamental mode of reference -- *de actu* contents are about and, in that sense, refer to actions that are themselves directed at particulars, and in this way, they are about (refer to) particulars. They constitute a generative rather than surrogative mode of aboutness, directedness, or reference to particulars (see Chapter 4). For Evans, the realm of reference is that with respect to which the correctness, which he understood as truth values, of the content is determined. For me, the realm of reference is that with respect to which the correctness, which I understand as appropriateness values, of the content is determined. Hence, Evans does not include abilities in the realm of reference whereas I do.

EIR/ *de actu* contents also bear some similarities to Peacocke's (1992) account of perceptual content as "scenario content"-- content which consists of a scenario which is a spatial type, i.e. a representation of the layout of the world built up over time (which may include temporal information), which corresponds to "the *way* of locating surfaces, features, and the rest in relation to a labeled origin and family of axes... in the real world fixed in accordance with the labeling in the scenario" (614). However, scenario contents are more descriptive and less action-oriented than

de actu/ embodied instructive representational contents. But like scenario contents, I take it that *de actu* contents are the most fundamental kinds of representational content in that the "representational properties of all other sorts in various ways presupposes the existence" of this type of content (61), for on my view, all varieties of surrogative content piggy-back on *de actu*/embodied instructive representational content.

EIR *de actu* contents are probably most similar to Adrian Cussins's (1990) ability-based account of the nonconceptual personal-level content of experience in terms of what he called "cognitive trails." Like Evans, Cussins holds that the cognitive significance of content consists in the "experiential accessibility" of abilities "to the subject in experience-based knowing-how." (7) Cussins calls this "the realm of embodiment." And like Evans, Cussins holds that the realm of embodiment, "unlike the realm of reference, is not that which determines the correctness or incorrectness of a representational content." (7) This is because the truth of the matter about what particular a content relates one to (refers to) is not a matter of the abilities and sensory capacities employed in grasping it. Indeed, most of the facts about the particulars one grasps are independent of one's grasping them. While this is true, as I see it, truth is a matter not just of what the facts are but of how one surrogatively represents them; truth is a matter of how a kind of content *fits* its object. On my view, the fundamental embodied form of cognitive significance is not fundamentally a form of cognitive significance that fits its object (the particular to which it refers) to the extent that it is "true" of it. Instead, the fundamentally embodied form of cognitive significance fits its object insofar as the abilities employed in accessing it are appropriate to it. Whether the abilities employed-- the potentiated intentional actions-- are appropriate to the particulars at which they are directed depends, of course, on the nature of those particulars (the facts). In this way, the appropriateness value of an EIR's *de actu* content reflects the facts-- what

is true in a situation. But *de actu* contents do not present the facts in a way that's evaluable in terms of truth (or other sorts of alethic success conditions). Instead, they present the facts in a way that's evaluable in terms of appropriateness.[15]

Finally, since embodied instructive concepts are realized by abilities for types of intentional actions (solutions to practical problems), and since surrogative representations and concepts extend these abilities, I endorse an ability-based theory of concepts. The idea that at least some concepts might be embodied (Lakoff and Johnson 1980, 1999) or constituted by abilities (e.g., Dummett 1993, Bennett & Hacker 2008, Kenny 2010) is not new.[16] However, I will not attempt to locate this theory in relation to other such theories. I believe my account of the problem-solving dispositions that realize embodied instructive concepts, as well as my account of how EIRs can involve SRs (Chapters 4 & 5), will go at least some way towards distinguishing my view including its ability to address some common objections to ability-based theories of concepts, e.g. the extent to which they're informative about cognitive process, and the extent to which they're compatible with common ways of understanding the generativity and systematicity of thought (Fodor 1968, Chomsky 1980). But I will not try to show that this is the case here.

Propositional Attitudes

As the dissertation develops, I will introduce an alternative model of psychological representation in stages, and at times under different names (e.g., practical or *de actu*

---

[15] My account of EIR *de actu* contents also takes inspiration from Susanna Schellenberg's (e.g., 2018, 2019) capacities-based theory of perceptual content.

[16] The view that concepts are abilities is opposed to the view that concepts are mental representations in the sense of being bits of an inner language, symbols, inner pictures, etc. I am arguing that mental representations themselves should be thought of as abilities. The objection that abilities are ill fit to play the explanatory role of concepts such as the productivity of thought and that they can say little about mental processes does not apply to my view. I will not directly argue for this claim, but the account of EIRs and embodied instructive concepts developed in this chapter and the next will contain much material out of which a direct argument could be constructed.

representationalism (Chapter 1), the intentional action acorn model (Chapter 2), the EIR model (Chapters 3, 4 &5), the instructive attitude model (Chapter 5), and the "intentions-first" model (Chapter 5)). In every case, the contrasting model is the propositional model. But what do I mean by "the propositional attitude model"?

First, when I say "propositions" I am not talking about ways of modelling propositions. I'm talking about a kind of representational content -- the kind of thing that sentences purportedly express, and the kinds of things that subjects purportedly have attitudes towards.

Second, when I say that psychological representations are not well-modeled as propositional attitudes, I don't mean that propositional attitude talk is never useful. I think it's obviously pretty useful for, e.g., modelling language. And different ways of modelling propositions will plausibly be more or less useful for modelling different bits of language. But I am not talking about modelling language. Rather, I'm talking about modelling psychological representations. And I'm not talking about modelling attributions of psychological representations-- that's more language. My objective is modelling psychological states themselves-- the things those ascriptions (however we want to model them) refer.

Third, I'm not saying that propositional attitude models of psychological states are never useful. Plenty of useful models are "fictional" or "abstract" or "ideal." I only object to propositional models insofar as they purport to be descriptively accurate of psychological representations. I think philosophers are often unclear about how they are understanding propositional models and go back and forth between treating them as fictional/ideal on some occasions and as descriptively accurate (realist) models on others. My target throughout the dissertation is the propositional attitudes model *qua* realist model of psychological states themselves. Finally, the propositional attitude models that are my target are not necessarily strictly

speaking propositional. They include any realist model of *fundamental* (original) psychological representation according to which psychological representations stand-in-for what they represent and hence cohere with the Descriptive Thesis.

Reasons (for Action)

This dissertation will discuss different kinds of reasons for action. In Chapter 2 I will talk about the relationship between intentional actions and the kinds of reasons that figure in action explanations. In Chapter 3 I will distinguish between "objective" and "subjective" reasons and between two species of the latter, namely "intra-subjective reasons" and "inter-subjective reasons."

The literature on reasons for action is dense and complicated. Philosophers commonly distinguish between "normative"/justifying and "motivating"/explaining reasons (Raz 1975, Smith 1994, Parfit 1997, Dancy 1995, 2000).[17] However, some philosophers also distinguish between normative, motivating, and explaining reasons (Baier 1958, Alvarez 2007, 2009a, 2010, Hieronymi 2011, Smith 1994, Darwall 2003, Mantel 2014). Roughly, a normative reason is a fact that *favors* an action; a motivating reason is a psychological state or fact that *motivates* an agent; and an explanatory reason is a fact about an agent that explains an agent's action (e.g., the fact that an agent was jealous). Philosophers (e.g., Schroeder 2008, Williams 1979) also sometimes distinguish between "objective" and "subjective" normative reasons, where whether an agent has an (objective) normative reason to act depends solely on the facts and not on the agent's psychological states, i.e., not on how agents take, understand, or appreciate those facts. There are also "objectivist" or "externalist" and "psychologistic" views of motivating and explanatory

---

[17] Dancy (2000, 20ff & Appendix) discusses the history of this distinction.

reasons. On psychologistic views, motivating or explanatory reasons are psychological entities such as beliefs and desires, or the fact of agents having certain beliefs or desires.

The main distinction for my purpose is between normative reasons, which refer to actual facts in a situation that count in favor of an agent acting a certain way (Raz 1975, Scanlon 1998), and motivating reasons, which refer to how an agent takes things to be with respect to her acting, that is, what an agent considers as counting in favor of acting a certain way (Audi 2001, Mele 2003).[18] Normative reasons are generally thought to establish a relation of "being a reason for" between a fact, an agent, and an action kind (see Raz 1975, 1998, Dancy 2004, Cuneo 2007), or between a fact, an agent, an action kind, and a time, circumstances, etc. (Skorupski 2010, Scanlon 2014). In this sense, normative reasons are "agent-relative" (Nagel 1970, 1986).

I will use "objective reasons" to refer to normative reasons (or objective normative reasons). I will use "subjective reasons" to refer to motivating (or explanatory) reasons. However, in distinguishing between intra-subjective and inter-subjective reasons, I'll be blurring the lines a bit. Intra-subjective reasons are the way subjects first-personally take things to be such that they are disposed to act in certain ways; they're EIRs. Inter-subjective reasons are objectifications of ways subjects take things to be-- they are public representations that stand-in-for facts in a way that relates them to EIRs and hence intra-subjective reasons. Inter-subjective reasons are like objective reasons in that they're in the world rather than in subjects' heads and they stand-in-for objective reasons. But where objective reasons are "factive" in the sense that they just are what they are (they are not representations of things being some way), inter-subjective reasons can

---

[18] There is debate about what the relevant "facts" are, i.e. propositions or their truth-makers. I will assume that the relevant facts are truth-makers. So I'll be following Raz (1975). For Raz, the facts that are reasons designate "that in virtue of which true or justified statements are true or justified. By "fact" is meant simply that which can be designated by the use of the operator "the fact that …" (17–18).

misrepresent. In addition, while intra-subjective reasons are psychological states (representations) or facts, inter-subjective reasons are not. Finally, the fact of an agent having an intra-subjective reason can itself be an objective reason. Likewise, the fact of there being an inter-subjective reason can itself be an objective reason.

In speaking of subjective reasons, I will generally be speaking of "*pro tanto*" reasons rather than all things considered reasons. An intra-subjective reason is an EIR, i.e., a potentiation of an intentional action, where this is the fundamental way agents take facts to be. Intentional actions are "local" solutions in that they solve "local" practical problems. For instance, eating an apple solves a token need for nourishment. Local solutions are potential global solutions. Roughly, a global solution solves the global problem of prioritizing the local problems that are most important to one's overall flourishing in a context. So EIRs (potentiations of local solutions) are *pro tanto* reasons. The EIRs that are selected for performance are all things considered reasons.

Intentional Actions

One of the many difficult questions I faced in writing this dissertation is what to call the kind of action that's my focus. I refer to it as "intentional action" because my topic is intentionality. But I talk primarily about actions of which non-human animals are capable, and thus actions that do not constitutively involve *rationally self-consciously* acting for reasons. Some philosophers might want to reserve "intentional action" for this latter type of action, but I'll refer to it as "Rational Action."[19] I agree that it is important not to elide rational action with what I'm calling intentional action. As I see it, rational action is an important (philosophically probably the most

---

[19] Importantly, "Rational" here is not an honorific; it does not mark the type of action as rationally successful but rather as a *subspecies* of intentional action that is distinctive of rational beings/life forms.

26

important) subspecies of intentional action; the characteristics that make it interesting are the characteristics that more or less uniquely distinguish the subspecies. I have created the following table so that the reader may see how I propose to respect the continuities and discontinuities between different kinds of action and agency:

Table 1 Varities of Agency

Proto-Intentional Action/ Proto-Agency *"Biological Activity"*	*Solutions but not application-accountable solutions: - Modified through evolution not learning during individual's lifetime - Local solutions don't compete to be global solutions (Chapter 2)
Intentional Action/ Basic Agency *"Purposeful Action"*	*Application accountable solutions (Chapter 2, 4, 5) *Acting for reasons (intra-subjective reasons) (Chapter 3) but not acting on the basis of reasons.
Rational Action/ Rational Agency *"Human Agency"*	*Application- & agent- accountable solutions (Chapters 2, 4, Conclusion) - Modified automatically and intentionally - Requires self-conscious/rational capacities *Acting on the basis of reasons (inter-subjective reasons) (Chapter 3) *Note: "voluntary" and "involuntary" actions, to the extent that Hyman (2015) is right that these are ethical notions, get a hold here.

1.5   Chapter Summaries

Chapter 1

This chapter will mostly focus on undermining the descriptive thesis. After clarifying the thesis, I'll explore three potential sources of motivation or argument for the thesis. First, I'll consider a linguistic justification of the descriptive thesis. This discussion will allow me to distinguish between deflationary and substantive forms of endorsing the descriptive thesis. It will also allow me to begin to distinguish a representationalist conception of the intentional relation from commitments regarding the descriptive thesis. Second, I'll consider how a broadly Kantian conception of representation might seem to favor the descriptive thesis. On a broadly Kantian conception of representation, to represent something is to have that something available as a target of judgement; representations are essentially judgeable contents. And since it is widely supposed that judgement acts on (i.e., reasoning essentially involves) propositional contents, those who endorse this kind of representationalist view typically hold that representing something essentially involves that something in a propositional content. I will derive criteria for representation from the Kantian conception of representation and argue that they might be met by what I'll call *de actu* contents rather than propositional contents where *de actu* contents are involved in a fundamentally practical kind of reasoning or judgement. I'll next consider a potential argument for the descriptive thesis which turns on the idea that representations must be informative and that anything that's informative will conform with the descriptive thesis. This argument will be revealed to depend on the Derivation Thesis. I'll ultimately suggest that just as there's logical space for an essentially practical version of broadly Kantian representationalism, so there's logical space for an essentially practical notion of informative function. So, if informative function is the essence of

representation, some representations may have *de actu* contents. Finally, I'll argue that there's reason to think that original (fundamental) representations could not be as the descriptive thesis describes them.

Chapter 2

The primary target of this chapter is the derivation thesis. After arguing against theories of action that hold this thesis and in favor of theories of action that deny it, I'll motivate and introduce my problem-solving account of intentional action. According to this account, an intentional action is an application-accountable solution ("solution" for short) to a practical problem, where practical problems are defined in terms of need for flourishing.[20] On the proposed view, an intentional action is a type of process that, if fully actualized in its domain of application, brings about a state of affairs that constitutes the (non-accidental) satisfaction of a type of need for flourishing. A token of a type of intentional action is *appropriate* iff it is tokened within its domain of application. An intentional action's domain of application is the set of facts that are instances of the practical problem it functions to solve (need facts) together with the set of all facts that are instances of the type(s) of target involved in solving the relevant type of practical problem (target facts). Only if a solution is applied within its domain of application will it non-accidentally solve the practical problem it functions to solve. A type of solution to a practical problem is application- accountable iff it is sensitive to the (in)appropriateness of its tokens, i.e., to the success or failure of its applications. What determines a solution's dispositions to be applied are causal mechanisms I call

---

[20] *Needs for flourishing are not desires* but rather non-psychological intentional facts; they are a key component in what's ultimately an externalist (objectivist) account of reasons for action. That said, I am also concerned to give an account of subjective reasons-- the internalization of objective reasons. Indeed, in Chapter 3 I'll argue that this may be the key to providing an illuminating account of original intentionality (psychological representation). And in Chapter 5 I will provide an account of desires as aspects of EIRs (what I'll call "N-instructions") which correspond to need-input activations of "instructive dispositions," i.e. abilities for application-accountable solutions to (local) practical problems, or what I'll call "embodied practical concepts."

"application algorithms." Accordingly, the processes that are actions are tokens of a type of process that undergoes modification according to whether or not an agent's practical problems are being successfully solved, and hence according to whether a solution-type's application algorithms are properly sensitive to the relevant domains of application. A further condition on a solution being application-accountable is that this sensitivity has a global (holistic) dimension.[21]

This chapter will also introduce my "intentional action acorn" model of psychological representations. I develop this model as an initial (because largely metaphorical and insufficiently theoretically grounded) replacement for the propositional attitude model. According to this model, fundamentally, beliefs (and/or perceptions), desires (and/or emotions), and intentions are potentiations of intentional actions: they are related to completed intentional actions as acorns are related to fully developed oak trees. I'll motivate the intentional action acorn model as a way of providing a substantive account of psychological representations that will satisfy fans of the rich correspondence view and that coheres with teleological theories of action, such as my problem-solving account. Chapter 3 will provide theoretical grounding for intentional action acorns in the form of the practical epistemic access (PEA) analysis of representation while Chapters 4 and 5 will develop the intentional action acorn model so that it will ultimately become the embodied instructive representation (EIR) or "intentions-first" model. However, I'll continue to use the acorn metaphor wherever it seems useful.

---

[21] All of this will be explained in Chapters 2, 4, and 5.

Chapter 3

This chapter will rehash some of Chapter 1's criticisms of the descriptive thesis, but from a somewhat different angle. And where Chapter 1 just provided a glimpse of the PEA analysis, Chapter 3 will develop and directly defend this analysis.

The chapter will be framed around the problem of providing an illuminating analysis of original representations where these are thought to constitute an agent's subjective (motivating) reasons for action, i.e., her beliefs, desires, and intentions. The view I'll refer to as the "standard strategy" treats the standing-in and hence alethic-conditional form of derived representations as paradigmatic of all representations, including the original representations in terms of which it hopes to illuminate the nature of subjective reasons. Since the standard strategy does not appear to work, I'll argue for an alternative strategy: rather than trying to illuminate subjective reasons in terms of a prior conception of the essential features of representation, start by producing an analysis of subjective reasons and use that analysis to illuminate the nature of representation.

Since philosophers seem to assume that only psychological representations are subjective reasons, they're likely to think that the proposed strategy is a non-starter. I'll try to get to the bottom of this bias. I'll ultimately argue that what at first may look like evidence that derived representations are not subjective reasons might instead be evidence that original and derived representations correspond to two species of subjective reasons. I'll argue that both derived and original representations are essentially subjective and related to intentional actions so that derived representations may be understood as "inter-subjective reasons" *on the basis of which* an agent acts while original representations may be understood as "intra-subjective reasons" *out of which* an agent acts or which an agent *enacts*.

I'll then pursue the strategy I proposed of starting with an analysis of subjective reasons and using that as a basis for an analysis of representation. The practical epistemic access (PEA)

analysis will be introduced as an analysis of subjective reasons which I'll extend to representations. The function of providing direct-PEA will ground "intra-subjective reasons" *out of which* an agent acts, or what I'll call "embodied instructive representations" (or "EIRs"), while the function of providing indirect-PEA will ground "inter-subjective reasons" *on the basis of which* an agent acts, or what I'll call "surrogative representations" (or "SRs"). While the PEA analysis suggests that appropriateness is the fundamental norm for representation, I'll argue that the alethic success conditions the descriptive thesis considers essential to representation may be understood as distinctive aspects of the surrogative species of appropriateness conditions.

Chapter 4

This is something of a keystone chapter. It will bring together the positive proposals of the three preceding chapters and it will add a further piece to the picture.

I'll identify the notion of *de actu* contents from Chapter 1 with the notion of an "intentional action acorn" from Chapter 2 and with the notion of an embodied instructive representation (EIR) and hence intra-subjective reason (reason *out of which*) from Chapter 3. I'll also show that the definition of representation as PEA from Chapter 3 can be chained with the definition of intentional action as an application-accountable solution to a practical problem from Chapter 2. This will yield a unified account of intentionality which I'll have motivated from multiple angles in the preceding chapters.

The piece I'll add is an account of the "problem-solving dispositions" that realize EIRs. This will allow me to show that the notion of an intentional action acorn from Chapter 2 can be cashed-out non-metaphorically in terms of EIRs and to start to put some flesh on the bones of the conception, gestured at in Chapter 1, of practical judgement that acts on *de actu* contents. I'll also

illuminate the nature of EIR content-- what it means to be a *de actu* content or generative mode of presentation-- and I'll address some concerns about the representational status of EIRs.

Chapter 5

This chapter further fleshes-out the EIR account of psychological representation and supplies it with further motivation.

I'll begin by developing an intentions-first EIR psychology spelled out in terms of "embodied instructive attitudes" which are realized by different ways of activating instructive dispositions (IDs). I'll then argue that EIRs can subsume propositional attitudes. I'll distinguish cases in which EIRs involving surrogative representations (SRs) constitute "full-blooded" propositional attitudes from cases in which there are just EIRs relating subjects to SRs but that fall short of full-blooded propositional attitudes. I'll propose that full-blooded propositional attitudes (or "propositional attitudes proper") are constituted by clusters of IDs (and EIRs resulting from activations from IDs in these clusters) for broadly critical or rational activities. Such activities function to solve "rational" practical problems, which are often (if not always) proximity problems. Hence, they typically involve responding to and producing SRs. Rational practical problems may be grounded in something like a need for self-consciously acting (including epistemically acting) as one should-- a need for justification, understanding, or the ability to give an account to answer all varieties of "Why?" -questions.[22] Examples of such activities include: seeking and constructing explanations and testing and evaluating theories, and asking after the justification for explanations.

Finally, I'll argue that the proposed EIR psychology has advantages over the standard propositional attitude psychology. To show this, I'll argue that the standard propositional attitude

---

[22] A.k.a. Self-making and self-improving activities.

model faces two clusters of challenges. The first cluster concerns obstacles to illuminating psychological representations when the descriptive thesis, and thus the propositional attitude model, is assumed. These will be somewhat familiar from Chapters 1 and 3, but they'll contain new material. The second concerns limits on the explanatory power of propositional attitudes. I'll then argue that the EIR psychology can address both clusters of challenges. Hence the proposed EIR psychology appears to have some distinctive advantages.

Conclusion

In light of the preceding chapters, I'll be in a position to begin to address some lingering worries the reader might have for my account of representation and intentional action. One of these worries concerns the dispositional underpinnings of EIRs. The other concerns the extent to which my account of intentional action is adequate.

****

As these summaries make evident, most of my energy in this dissertation is focused on illuminating original representations in terms of EIRs (and hence *de actu* content, generative modes of presentation, and intentional action acorns). But of course, modelling derived representations as SRs has a number of implications especially in the context of the philosophy of language aimed at theorizing linguistic meaning. A full defense of the PEA analysis of representation would need to show that SRs can do the work of propositions and the concepts they involve. So, it is a weakness of this dissertation that it does not develop those implications.[23] While this is admittedly a weakness of this dissertation, I hope it's understandable. My approach in this

---

[23] I do not yet have anything close to an SR semantics for ordinary language. I am inclined to think it would in certain respects resemble Brandom's inferentialist semantics, but I have not even begun to try to work out a translation for SRs into that or any other semantics.

dissertation, which I'll to some extent defend in Chapter 3, is to start from the bottom, that is, to start by trying to understand fundamental (original) representations and then work my way up. Rather than constraining my theorizing about the nature of fundamental representation on the basis of existing ideas about public representation, my *modus operandi* is to start fresh with fundamental representation and to use that picture to constrain my theorizing about derived representation, such as linguistic representation. The project undertaken in this dissertation is nothing short of massive, so I hope I can be forgiven for being unable to see more of it through.

2.0 Chapter 1: Representation, Judgement, & Informative Function

As noted in the introduction, this dissertation is largely organized around rejecting and replacing the descriptive and derivation theses. Both theses are widely assumed by contemporary analytic philosophers. Many philosophers endorse both while some philosophers endorse the descriptive thesis but not the derivation thesis. I do not think it's an accident that many philosophers endorse both theses. In this and subsequent chapters I'll explore some ways in which commitment to one can lead to commitment to the other. But in this chapter, I'll mostly be concerned with the descriptive thesis.

Rejecting and replacing these theses does not entail denying that representations are ever as the descriptive thesis describes them, nor that representations are not at all or not at all substantially involved in the production of intentional action. What's denied is that *fundamentally*, representations are as the descriptive thesis describes them and that, *fundamentally*, the intentionality of action depends on the intentionality of mental representations. So, my primary target in this chapter is the view that the descriptive thesis is true of original representations, or intentional states, which I'll assume are the fundamental kind of psychological representation.

In §1 I'll clarify the descriptive thesis. In §2 I'll consider a linguistic justification for the descriptive thesis. This discussion will allow me to distinguish between deflationary and substantive forms of endorsing the descriptive thesis. As I explained in the Introductory Chapter, my focus in this dissertation is on substantive conceptions of intentional phenomena including psychological representations. This discussion will also allow me to begin to distinguish a

representationalist conception of the intentional relation from commitments regarding the descriptive thesis.

In §3 I'll consider a different kind of reason one might endorse the descriptive thesis. According to a broadly Kantian conception of representation, to represent something is to have that something available as a target of reasoning and judgement. And since it is widely supposed that judgement acts on (i.e., reasoning essentially involves) propositional contents, those who endorse this kind of representationalist view typically hold that representing something essentially involves that something in a propositional content. I will derive criteria for representation from the Kantian conception of representation and argue that they might be met by what I'll call *de actu* contents rather than propositional contents where *de actu* contents are involved in a fundamentally practical kind of reasoning or judgement.

In §4 I'll consider a potential argument for the descriptive thesis. This argument turns on the idea that representations must be informative and that anything that's informative will conform with the descriptive thesis. This argument will be revealed to depend on the Derivation Thesis. I'll ultimately suggest that just as there's logical space for an essentially practical version of broadly Kantian representationalism, so there's logical space for an essentially practical notion of informative function. So, if informative function is the essence of representation, some representations may have *de actu* contents. I'll conclude this section with a preview of the account of representation to be developed in the coming chapters.

Finally, in §5 I'll argue that there's reason to think that original (fundamental) representations could not be as the descriptive thesis describes them.

2.1 The Descriptive Thesis

The descriptive thesis has been almost universally assumed by analytic philosophers.[24] Importantly, this thesis is compatible with holding that description is insufficient for reference (or at least reference of a certain sort, e.g., singular reference). What it's committed to is that description *of a sort* is a necessary condition for reference.

The idea is that to represent something is to represent something *as being some way*.[25] Accordingly, maintaining the descriptive thesis is compatible with maintaining that there is direct reference of a sort. For instance, names may directly refer to their referents. But names are linguistic items that seem to presuppose that language users already have intentional access to the things they name. In the case of fundamental representation, e.g., mental representations like perceptual representations, the idea is that there are singular representations-- representations that directly refer in the sense of directly getting a hold on their referents such that the referents are constituents of their contents. But singular representations still represent their referents in some

---

[24] In saying that this thesis has been almost universally assumed, one should not take me to be saying that philosophers have only been thinking in terms of descriptive reference. After all, this would obviously be false-- a number of philosophers have argued for the need for and possibility of non-descriptive (non-*de dicto*) senses, i.e. for the need and possibility of *de re* senses. That they have done so, however, is consistent with my claim that philosophers have almost universally assumed the descriptive thesis, for these accounts still assume that directive reference (or de re senses) achieves alethic-conditional representation, or at least, they have developed their accounts of direct reference in terms of such representations. While I will not develop it specifically in these lights, my account of direct-PEA, i.e. EIRs/ generative modes of presentation of facts-- may be understood as an account of subjective (in that they only characterize a subject's understanding and senses are not propositions) *de re* senses. Some of my arguments for the PEA analysis of representation and the existence of EIRs involve criticisms of other analyses of representation that might be extended as criticisms of accounts of *de re* senses that assume the Descriptive Thesis. I'll say a little more about this in Chapter 4.

[25] Bare indication of the sort that might be constituted by some sort of correspondence (often causal) relations and that's often misleadingly described as "information" is not sufficient for reference. Reference (and anything properly called information) has to be defined in terms of at least being potentially exploited for a function involving a certain kind of user. Where the descriptive thesis is assumed, cashing out the relevant function (and users) inevitably leads to a richer notion of indication that involves description. Relatedly, it is typically assumed that intentional representations must be able to misrepresent. I do not think it is possible to make sense of misrepresenting where this means a representation is false, falsidical, unsatisfied, or inaccurate, without presupposing the kind of involvement of description I describe in this section. Alethic conditions get a hold on representations that represent their success conditions, and that's just what it is to have a descriptive content in my sense.

way. Hence, they canonically have the form "S is F" or "That F" where "S" is a subject-matter and "That" demonstratively refers to a subject matter in both cases by means of applying the concept, predicate, attribute, or property F. I will refer to representations of this form as "stand-ins." Names, labels, and concepts or attributes like "F" stand-in-for particulars, individuals, or kinds and in that way refer to them. For something to function as a name or label, it must be somehow attached to a referent. Linguistic names get attached to their referents by language users and the language conventions they establish. When it comes to fundamental representations, like perceptual representations, a subject (or her "mind") comes equipped with demonstrative concepts, or perceptual attributes or capacities or sensory categories that function to pick-out certain kinds of referents. The relevant concepts etc. enable a subject's perceptions and thoughts to stand-in-for, i.e., be about an object or individual (a referent). In other words, what results from the application or deployment of such concepts etc. are representations (or thoughts) that represent a referent as being some way, and hence representations of the form "S is F" or "That F."

The descriptive thesis entails that fundamental representations are of this form. This is because of the way it understands representational success. Alethic success conditions apply to stand-ins-- to things whose contents articulate the conditions under which they are true, veridical, accurate, or satisfied. While alethic success conditions do not necessarily entail propositional or conceptual contents proper, they do entail that representational contents are proposition-like. For instance, while many hold that perceptual contents are non-propositional (perhaps iconic) and thus non-conceptual, they nonetheless hold that they do something akin to asserting something about the world (Glüer 2009 p. 299, Cohen 2015 p.1514), i.e. something akin to putting forth

propositions.[26] For Burge (2010), broadly predicative structure is key: perceptual representations employ a *general* predicate-like "attributive" under which a *particular* subject-matter or "singular element" is picked-out (p. 83).[27] Nonetheless, for simplicity, I will mostly talk about propositional contents in this chapter.

I will ultimately argue that fundamental mental representations may have "*de actu* contents" rather than broadly propositional or attributive contents, where *de actu* contents representations have the form "ɸ-THAT-THUS" (where ɸ might be an act of producing a demonstrative or other type of judgement) rather than "S if F" or "That F." I'll arrive at this proposal by way of considering potential motivations for the descriptive thesis.

2.2   Representations & Language

One reason philosophers may endorse the descriptive thesis is that they tend to think of intentional relations as relations between subjects (or their attitudes) and propositions. This is

---

[26] Susanna Siegel (2011 p. 30) holds that perceptual experience is judgment-like by modelling its content on that of beliefs. See also Byrne (2021).

[27] Similarly, Nanay (2020, p. 1) claims that "representing something means attributing properties to this thing". As I see it, there are two ways to think of an attributive representation. First, there's the representation-- the sentence or judgement produced-- that says of something it names that it is so and so (that predicates something of something). For example, in the sentence "Bihui is small," "Bihui" names my cat and "is small" predicates smallness of her. So the whole sentence stands-in-for the state of affairs the truth condition of which is that Bihui (my cat) is small. Second, there's an agent's act of attribution: pointing at Bihui, I say, e.g., "she is small." In this case, the personal pronoun "she" paired with my pointing indexically stands-in-for Bihui, and I attribute the property of being small to her. The result of such an act is the production of a sentence consisting in a pronoun (with Bihui as the referent) and a predicate (is small) with the truth condition *that Bihui (the cat I'm pointing at) is small*. In both cases, what is produced is a representation that stands-in-for my cat and applies a description to her. It specifies or articulates the state of affairs that is its truth condition. In general, representations that involve naming, describing, picturing, modelling, or attributing properties to things named, described, modeled, or pointed to with some implicit categorization (e.g. as a cat) are representations that *stand-in-for* the states of affairs that make them successful-- true, veridical, accurate, or satisfied.

because sentences like "s believes that p" and "s desires that p" are sentences that ascribe intentional states to subjects in which the proposition-- the "that…." clause-- appears to specify the content-- the thing to which the subject is has some psychological attitude. There are, however, different ways of understanding what this comes to.

The deflationary view holds that the claim that intentional states relate subjects to propositions simply describes the nature of ascriptions of mental states without committing to any view about the underlying metaphysical nature of the referents of such ascriptions. In contrast, the substantive or inflationary view holds that the claim that intentional states relate subjects to propositions *does* describe the underlying metaphysical nature of the referents of intentional states ascriptions. I have no objection to the deflationary view. My target is the substantive/ inflationary view.

If one holds the inflationary view, one will endorse the descriptive thesis (unless one has an extremely unorthodox conception of propositions). So, what might justify adopting the inflationary view over the deflationary view?

One might appeal to the fact that if the intentional relation is fundamentally a relationship between subjects and propositions, this would explain why intentional state ascriptions can relate subjects to "non-existent objects." For example, in the case of illusion and hallucination or mistaken belief, we may ascribe experiential and belief states to subjects that appear to relate them to non-existent objects or states of affairs. We also ascribe states of desiring, wishing, intending, and predicting that seem to relate subjects to future and thus not yet existing states of affairs. Subjects can also have attitudes about theoretical posits like phlogiston and fictional characters like Superman which seem to lack any real referents. Call these "non-existent object cases." Because propositions are themselves thought to be entities (albeit abstract), they can arguably

uniformly account for what subjects are intentionally related to--- and thus what ascriptions of such states refer to-- in such cases.

However, it is not obvious that understanding intentional relations as relations to propositions provides the best explanation of "non-existent object" cases. First, we should distinguish between two ways of understanding the sense in which intentional relations are relations between subjects and propositions. On a "propositionalist" view, fundamentally, in an intentional relation, a subject is directly related to a proposition and only indirectly related to an ordinary object or state of affairs. In contrast, on at least some "representationalist" views, it's not the case that subjects are directly related to propositions and only indirectly related to ordinary objects and states of affairs. Propositions specify not what a subject is related to but rather how it is that a subject is related to some ordinary object or state of affairs. In other words, the representationalist view is a substantive view insofar as it takes it that propositional attitude ascriptions refer to psychological representations of a kind, but it is not necessarily committed to the descriptive thesis. The representationalist view seems to have two advantages over the propositionalist view.

First, it avoids the objections to propositionalism raised by proponents of the objectualist view. According to the objectualist view, intentional relations are fundamentally relations between subjects and ordinary objects. Objectualists argue that the propositionalist view is unintuitive: Intuitively, intentional state ascriptions attribute intentional relations obtaining between subjects and subject matters-- objects, facts, particulars and not, at least generally, representations. Indeed, even if the intentional relation were thought to entail that subjects represent the subject matters to which they are intentionally related, it wouldn't follow that subjects always represent their representations of those subject matters. Objectualists also argue that the propositionalist offers an

unparsimonious account of the many intentional state ascriptions that do not (in natural language) employ that-causes. For instance, ascriptions of hopes, fears, wants, moods are often of the form "Sally hopes/ is hoping for a raise," "Kevin is afraid of dogs," "Jess wants a dog," "Louise is anxious about her new job."[28] These sentences do not employ that-clauses, but on their face, they nonetheless describe intentional relations between subjects and objects. So, *prima facie*, the view that intentional relations are between subjects and objects rather than (generally) subjects and propositions representing objects offers a more unified account of intentional state ascriptions.[29]

Second, the representationalist view is, as I'm characterizing it, less committal. For while it claims that propositions specify the way in which subjects are intentionally in the sense of psychologically representationally related to facts (objects, states of affairs), the view does not necessarily have to commit to treating the relevant representations themselves as propositions. What the representationalist is committed to is that the intentional relation is a representational relation and that this relation may be specified in terms of a propositional content. But the nature of the representational relation itself might not involve any propositions. And for at least some of the examples of non-existent object cases, this may seem like a good thing. After all, in cases of illusion and mistaken perceptual beliefs, subjects do seem to be intentionally related to ordinary objects, albeit in a way that constitutes a misunderstanding or that otherwise embodies a kind of error.

Consider the attribution "S believed that X was F."   The so-called "non-existent object" here is an X with property F. Intentional states typically figure in explanations of intentional

---

[28] See the essays in Grzankowski & Montague's (2018) *Non-Propositional Intentionality*.
[29] Of course, one might try to rewrite such ascriptions using that-clauses, but at least in some cases the resulting sentences are awkward and unnatural and in any case, this way of trying to establish a unified propositional construal of intentional state ascriptions risks being *ad hoc*.

actions, including intentional acts of judging or committing to things being some way. So one could plausibly interpret the "S believed that X was F" when there is no X that is F to which S is actually related as characterizing how the world would have to be for S's actions, or the actions to which she was disposed, to be successful: it would have had to have been the case that S was related to an X with property F. Assuming that a subject's actions reflect the way she takes the world to be, one might further interpret this sentence as saying that S was actually intentionally related to some subject matter (object, fact, particular) M: S mistook M for an X with property F such that she was disposed to treat M as if it were an X with property F. A representationalist could potentially say that the representation relation consists in or at least essentially involves such a disposition, and that this kind of dispositional state can be specified by a proposition without being identified with a proposition or propositional attitude.[30]

On the other hand, the propositionalist may seem to have an advantage over the representationalist in cases in which it is less plausible that a subject is intentionally related to an ordinary object, e.g., future-directed states, theoretical posits, and fictional characters. But this seeming advantage might be superficial. After all, the representationalist can allow that subjects are sometimes related to representations but deny that this is all they're related to. While being intentionally related to something means representing it, that doesn't mean representing a representation. And even where one does represent a representation, representations need not be understood as seemingly primitive abstract objects. To be related to a proposition may essentially involve being related to a representational vehicle. And representations that lack concrete referents

---

[30] In a somewhat similar vein, Ryle (1949) argued that "many of the cardinal concepts in terms of which we describe specifically human behavior are dispositional concepts" (117).

in this world might have other kinds of referents, e.g., possible worlds, or perhaps logical implications of natural laws or causal powers, or functionally or theoretically descriptively specified referents, i.e., "the thing T that explains phenomenon P and that has features x, y, z." Theoretical posits like phlogiston might be like this. Such representations may refer to T because there is some fact that underlies P and the subjects that produced the representation had access (perhaps inferentially mediated) to P. Such a representation may refer to P even if P does not have features x, y, z, e.g.., even if the representation misdescribes P. Similarly, fictional characters might be understood as human artefacts or constructions built out of representations that ultimately refer to real individuals, places, and personality traits. Such a construction might have a hybrid way of referring: on the one hand, it refers to a kind of functional role in a story, and on the other hand, its parts might refer to facts its author once experienced. And, of course, in some cases, the representationalist can even allow that subjects are just related to representations (or propositions), whether one understands these as possible worlds or something else.

I don't mean to suggest that these are particularly good explanations, but they are potential explanations. The point of mentioning them is to show that the representationalist may have resources to account for the seemingly more difficult kinds of non-existent object cases. It's also worth noting that accounting for what a subject is directly related to in non-existent object cases by means of propositions is more than a little suspect. Afterall, the non-existent object cases seem problematic because they seem to involve a subject being related to objects, properties, or states of affairs that are not concretely instantiated in the actual world. But propositions are themselves abstract objects, and hence it's not obvious that they are concretely instantiated in the actual world. So arguably, the only virtue of the propositionalist explanation of non-existent object cases is that it unifies these phenomena.

What's more, it's not obvious that a unified explanation is, in this context, a real virtue. It may be better to analyze different sorts of cases differently. As we saw above, it might be better to understand illusions and mistaken perceptual beliefs as actually referring to concrete particulars, just badly. In addition, some non-existent object cases might involve relations to representations (concrete vehicles) that stand-in-for some past state of affairs by virtue of someone in the past who was directly perceptually related to the past state of affairs producing a type of sign to stand-in-for that state of affairs, and by virtue of that type being reproduced over the years in order to refer to that past state of affairs. Fictional characters and places might have a hybrid way of referring, and theoretical posits might refer on the basis of the intentions of their producers whose mental states may refer to general kinds of facts by virtue of repeated observations and perhaps inferential and experimental activities. Perhaps such activities enable rational creatures to directly access general natures or essences. Perhaps even (at least human) perception can do this to some extent. In short, how we understand non-existent object cases depends on broader ontological and epistemological commitments. So, it is not obviously a virtue that the propositionalist accounts not only for all the non-existent object cases but intentional relations *generally* in terms of relations to propositions. In contrast, the flexibility of the representationalist view may be a genuine virtue.

The representationalist can even account for the reason intentional relations can be uniformly specified by means of propositions even if the representational nature of such relations does not actually involve any propositions. The reason propositions can specify representational relations is, arguably, their subject-predicate structure. By involving both subject and predicate, or object and concept, or sense and referent-- not quite the same things but all in the relevant neighborhood-- specifying intentional relations in terms of relations to propositions at least seems to account for the fact that an intentional content-- that is, an fact (object, state of affairs) insofar

as it enters into an intentional relation--is not just some entity (e.g. the object referred to) but an entity *as it figures in an intentional relation*, i.e. an entity that is "joined to" a subject's conception. As it figures in an intentional relation, an entity is essentially grasped, understood, conceived, or "presented" in some way or under some "mode" or in terms of some "form." This idea is related to non-existent object cases insofar as such cases highlight the understanding or grasp the subject contributes to the intentional relation as well as the normative properties of that grasp: the nature of the grasp may be proper or improper, or may "fit" or "fail to fit" (e.g. be true or false of) the object to which it's applied (or on which it has a hold).[31]

While I have treated the representationalist view as relatively neutral on how to understand the nature of original mental representations, the vast majority of representationalists are committed to the descriptive thesis. What justifies this commitment? The next two sections will consider possible answers.

---

[31] Another broadly language-based reason one might hold that fundamental representational contents must be propositional is that one might subscribe to a broadly Gricean conception of the meaning of linguistic signs. According to a broadly Gricean view, linguistic signs derive their meanings from agents' intentions: an agent produces a meaningful sign via an intentional communicative act the aim of which is to bring it about that some audience acquire certain beliefs. So linguistic meanings, or the intentionality of public signs, derives from the intentionality of actions and the representational contents of mental states. So the broadly Gricean view holds that the contents of mental states like belief and the contents of linguistic signs are two sides of the same coin, only one side of the coin-- the mental state side-- is more fundamental. Accordingly, many philosophers distinguish between "original intentionality," which they understand as the representational powers of mental states, and "derived intentionality," which they understand as the representational properties of public signs. Since most philosophers believe that linguistic signs employ propositional contents, philosophers who embrace a broadly Gricean theory of meaning typically hold that the contents of subjects' mental states are propositions, or, if not strictly speaking propositions, a structurally analogous species of content (e.g., content that's constituted by an "attribution" of properties or attributes rather than predication of a concept, and that has constitutive veridicality, accuracy, or satisfaction conditions rather than truth conditions). However, even if linguistic content is propositional and even if such meaning does supervene on the contents of intentional states, it simply doesn't follow that the contents of intentional states are themselves propositional or essentially analogous to propositions. Of course, perhaps the intentional states that are essentially involved in the production or interpretation of public intentional signs essentially involve those signs, their propositional contents to boot, in their intentional contents. But that doesn't mean that the contents of those states are themselves fundamentally constituted by the propositional contents they're involved in producing or interpreting. And, of course, not all intentional states are concerned with the production or manipulation of intentional signs.

2.3   Representations & Judgement

One might argue that the truth of the descriptive thesis follows from a *broadly* Kantian (or Sellarsian) view of representation. On such a view, to represent something is to have that something available as a target of reasoning and judgement. And since it is widely supposed that judgement acts on (i.e., reasoning essentially involves) propositional contents or some form of standing-in contents, those who endorse this kind of representationalist view typically hold that representing something essentially involves involving that something in something like a propositional content.

I consider the view I'll develop in this dissertation broadly Kantian/ Sellarsian view. But it is no doubt revisionary. For I do not think that propositions (or representations with a broadly subject-predicate form) capture the fundamental way in which subjects cognitively grasp the world. When it comes to grasping things that are not directly perceptually available, I think it's plausible that we have to grasp propositions. And when it comes to discursive or broadly theoretical reasoning, I think it's plausible that propositional contents are essentially involved. However, when it comes to fundamental representation, I do not think propositions are the appropriate forms of content. I'll get to the reasons why in §5 below. The purpose of this section is just to argue that propositional (or attributive) contents are not the only available option.

As I see it, the key Kantian/ Sellarsian insight is that our knowledge or intentional grasp of particular objects is not "given" -- that it requires a contribution from the knowing/grasping subject. This insight may be broken down into three component ideas: (a) An agent's material, passive causal reactions to the world do not on their own constitute an agent's intentional grasp, appreciation, or knowledge of the world; (b) a "mind" comprehends not the matter of material

objects but rather their form; and (c) the comprehension of form is not a merely passive matter but is rather essentially active.

Why should we believe (a)? Well, it seems clear that we are passively causally affected by all kinds of things, and that being so affected is insufficient for being intentionally related to them. So, something more than being causally affected by an object must be required for one to be intentionally related to it, or for it to be a (potential) object of knowledge for one. This is also true of our sensory capacities: they are causally affected by objects in ways that are not reflected in our intentional grasp of the objects they relate us to. And while sensory capacities no doubt relate us to objects in part by virtue of the ways in which they passively-causally relate us to objects, the fact that they do not merely passively-causally relate us to objects (perhaps in ornate ways, like a river that causally processes water through a series of many gates and channels) must presumably be explained by something other than the passive causal relations they enable subjects to bear to objects.

Why should we believe (b)? One may understand (b) as explaining why the intentional relation is so interesting to philosophers, why "intentional inexistence" is thought by many to be key to understanding intentionality, and why philosophers sometimes associate genuine representation only with "distal" senses.[32] The common thread running across these ideas is that intentional access is a kind of formal rather than merely material (or material causal) access. An intentional grasp one has of a cup in seeing it over there on the table is different from the material grasp one has of it when one literally holds it. When one literally holds it there is, intuitively, direct active causal-contact between the agent and the cup. But when one merely sees the cup, one's

_____

[32] For a somewhat similar perspective, see Kalderon (2018).

actively oriented towards the cup, but one's eyes are not directly causally affected by the cup, but rather by the light reflecting off it (and lots of other things besides which visual systems must process for an agent to be in a position to target (discriminate) a particular cup). One is not yet literally grasping the cup in the sense one grasps it when one holds it in one's hands. The cup's matter and your matter have not yet made direct contact. Yet you have a kind of knowledge or cognitive acquaintance with the cup that positions you to make intentional (active) causal contact with the cup. Your knowledge of the cup-- your being in a position to interact with it intentionally-- consists in your grasping its form.[33]

Finally, why should we believe (c)? Well, we already established that passive causal contact with an object is a material, and thus not yet an intentional relation. And if mere passivity is insufficient for an intentional (formal) relation, it's natural to think that what must be added is the agent's activity: the formal relation to an object is something a subject is active in achieving. And this is one reason to think that a subject's activity must somehow figure into her grasp of the form of the objects to which she's intentionally related. Another reason may be found if we consider what it is for a subject to grasp something: the mere causes to which an agent is related pass into and in some sense *through* her matter, but they do not involve her as a subject or agent in any interesting way. Plausibly, this is because a subject or agent is not just her matter but also her form. So plausibly, for a subject to grasp the form of an object is for the subject's form as well as the subject's matter to become involved with that object (where her matter, like the object's

---

[33] In the same vein, the phenomenon of "Intentional Inexistence" concerns our capacity to represent things that do not exist, e.g. bigfoot. In representing bigfoot, we seem to be representing the form of a possible (but non-actual) material object. More generally, we can be intentionally directed at or related to objects that seem to be mere forms, or what are sometimes called "mere ideas" like fictional characters, hallucinations, etc. So intentional objects are plausibly forms of objects; where the objects exist, we are intentionally related to them via their forms; where the objects are merely possible, we're related to their possibility insofar as we are related to a possible (unactualized) form.

matter, is involved materially-causally). The question is then how to understand a subject's form. An option I find attractive has its roots in an Aristotelian tradition which aligns a thing's forms (in the sense of substance or essence) with its special (distinguishing) function(s)-- i.e., activities (or ways of existing qua the kind of thing it is). On such a view, an agent's form becomes involved with an object's form (shape) via an agent's activities, or the active capacities that are essential to the kind of agent it is.

To see these ideas at work where the descriptive thesis is assumed, we may consider them in connection with Kant's theory of empirical cognition, which famously privileges the propositional form of thought. On Kant's theory, a subject's receptive capacities spontaneously "inform" sensory matter: the intuitions impose spatial and temporal order or structure onto sensory matter. The result is a spatial-temporal representation of an object. Conceptual capacities then apply additional forms, namely those corresponding to the categories (quantity, quality, etc.), to this representation. The result is then a representation that refers spatio-temporally to some object and that describes that object as being in the extension of the concepts (categories) that have been applied to it. And this representation--itself the result of two sub-individual judgement-like processes-- constitutes the thought a subject grasps where in grasping a thought a subject is in a position to judge it-- to evaluate it as true or as false. Of course, to be in such a position, the thought must be evaluable for truth or falsity. And it is precisely because it is itself judgement-like in that it presents some object as being spatio-temporally ordered in some way and presents a spatio-temporally ordered object as instancing certain properties (corresponding to the categories). Representations of objects as being spatio-temporally ordered and as belonging to certain categories are akin to judgements in that "presenting as" means *presenting this spatio-temporal ordering and participation in these categories as true* of the object to which they've been applied,

i.e. the object to which they represent and thus refer. If the thoughts a subject grasps were not representations of this type, the thinking goes, the subject would not be in a position to judge them true or false; it is in virtue of their (broadly) predicative form that they can be so judged.

I take it that propositions manifest the following three essential features of representations:

GENERALITY: Understanding is in some sense more general than the object it understands: it identifies the object as falling under some category/ies, instantiating some property/ies, partaking in some kind, or as being an individual (self-identical) that may be encountered in variant contexts.[34]

ACTIVITY: Understanding is in some sense actively supplied by the subject and imposed on or applied to the object the subject grasps in its terms.

NORMATIVITY: Understanding succeeds where it "fits" the object to which it's applied or on which it's imposed; understanding fails where it fails to "fit" the object to which it's applied or on which it's imposed.

These properties are (more or less) the properties philosophers associate with concepts and their applications.[35] The predicative structure that characterizes propositional content is undoubtedly a way of capturing GENERALITY, ACTIVITY, and NORMATIVITY. But it may not be the only way.

Thoughts and/or acts of predication or attribution are no doubt an essential feature of distinctively human cognition. And as self-conscious beings, it is arguably essential that we not only produce acts or thoughts with such form, but also that acts or thoughts with such form are objects for us to have thoughts about-- i.e., to judge. I also think it's true that reasoning and

---

[34] Individuals may also fall under more general categories.

[35] There are other ways of understanding what properties are distinctive of concepts, e.g. Evans's (1982) generality constraint. Evans was concerned to identify the conditions for singular thought understood as the capacity to have true/false thoughts abo ut particular objects. My aim is to establish that true/false thoughts are not the only kind of thoughts. So while I agree with Evans that the capaciti es for such thoughts are special, and that the concepts they involve have unique properties which include the nature of their generality, what I want is a characterization of concepts that's generic enough to encompass all varieties of thought.

judgement of a certain kind-- namely theoretical or discursive reasoning and judgement-- essentially involves propositional contents. Indeed, I think that without propositional content, we could not engage in what I consider the intentional activity of reasoning about what to do (or what to believe). However, it doesn't follow that the fundamental form of thought must be propositional, but only that the fundamental form of thought (representation) must be capable of accommodating predication. A relation may be representational if GENERALITY, ACTIVITY, and NORMATIVITY are satisfied. In other words, it might be sufficient for representation if a subject grasps an object's form by relating its form to the subject's own form where a subject's form is in some sense an active (spontaneous) part of her. For Kant, this active form was the spontaneous capacity of judgement. But in the *First Critique* Kant is specifically concerned with *theoretical* judgement. If then, there's a more general capacity of judgement in which theoretical judgment participates as a (special) species,[36] there may be a more general kind of representational content that underwrites our ability to produce and employ propositional contents.

The most obvious candidate is *practical* judgement or reason. Philosophers, including Kant, traditionally distinguish practical from theoretical or speculative reason or judgement while maintaining an underlying "unity" and common principle. However, as Dancy (2018) points out, the habit has been to try to unify practical and theoretical or speculative reason and judgement by treating the practical as a special species of the theoretical/speculative. But this is not the only option. Dancy (2018) argues for what he calls the "primacy of the practical" according to which theoretical or speculative reason is a species of practical reason or judgment. Intuitively, there may well be something to this. After all, producing propositions (thoughts) and propositional judgments

---

[36] And one that's plausibly available to non-human animals.

is one of the things we do--- one of our practices, if you will-- but *prima facie* it is not the only thing we do. And there's also the fact that it seems highly doubtful that human infants and non-human animals are capable of producing genuine propositional judgements. Their intentional actions may nonetheless be manifestations of practical judgements (of a fundamental sort).

But what would these practical judgements be like? Well, non-human animals and human infants are clearly capable of recognizing and selecting between affordances in their environment in a way that's broadly rationally evaluable. This sort of activity may be an essential component of practical judgement. Indeed, Dancy (2018) characterizes practical reason as, in the good case, what "takes us from the situation as we see it to acting in the way that is most favored by that situation." (42) And according to Dancy, practical reason "does this directly" meaning that "it does not need to pass through an intermediate stage such as a belief (propositional judgement) that acting in this way is most favored." (42) For Dancy, practical reason concludes in action, not (at least not generally or in every case) in a belief-like judgement. Practical reason consists in "adducing" considerations and acting "in their light," but this does not mean that one acts in light of a belief (propositional judgement) that considerations favor that action.[37] So perhaps the fundamental kind of content could itself consists in actions, or potentiations of actions, in response to facts. I'll call such contents "*de actu* contents."   Of course, the descriptive thesis and the

---

[37] This is not to deny that practical reasoning and intentional action require one to have some understanding of the situation that informs what one is doing and some understanding of what one is doing. According to Dancy, the "relation of being informed'' is not to be taken as that of being accompanied by mental comment or preceded by mental cause. The "mentality," or mindedness, of intentional action is essential to it. In this way we can generate what one might think of as a content for an action; only intentional actions have such contents. It is in virtue of having such a content that the action can have the sort of appropriateness to a situation that is involved when the action is done in the light of the various considerations adduced. One can only act in such a light when one has some conception of what one is doing, a conception that informs and shapes one's response as it unfolds (23-24). I take it that Dancy thinks this conception of content is internal to the action-- not something that merely "accompanies" or "causes" it. If so, I believe that the view I will go on to develop in later chapters fits nicely with the way Dancy comprehends the required understanding.

derivation thesis both make it hard to see how *de actu* contents could really be contents. But the point of this chapter is to question these theses.

Now one might argue that recognizing and selecting between affordances doesn't count as acting in the light of how one sees a situation, or in light of the considerations one adduces, so it doesn't really count as practical judgement. But I don't think this is obvious. Actions that result from selecting between recognized affordances may constitute the primitive form of practical judgement; acting in the light of considerations adduced might be a sophisticated form of this.[38] At the very least, I think there's *prima facie* reason to consider the possibility that practical judgement is more general than theoretical judgement. Practical judgements might employ *de actu* contents that consist in the potentiations of actions where judging about such a content consists in determining whether or not the actions will be performed.

De actu contents would be a kind of affordance-based kind of recognition (or misrecognition) of considerations (facts) that favor or seem to favor performing that action. They'd have contents of the form "ϕ-THAT-THUS" (where ϕ might be an act of producing a demonstrative or other type of judgement) rather than the form "S is F" or "That F." *de actu* contents and the judgements involving them may characterize the cognitive capacities of both humans and animals. In humans (rational animals) such capacities may become inextricably bound up with rational capacities, chief among which might be the capacity for propositional thought and judgement.[39]

---

[38] I believe this is compatible with denying an additive theory of human rationality. My claim is that human rationality and non-human animal rationality may share a genus despite being importantly different species. But to characterize the genus we want to characterize what these species have in common. So we shouldn't suppose that characteristics of the distinctively human/rational species will necessarily figure into an analysis of the genus.

[39] I take this to be consistent with the idea that distinctively human practical judgement, or properly rational cognition, essentially involves, in all its activities, at least the partial exercise (potentiation of) theoretical judgements--judgments that essentially involve representations

So, if, as I've suggested, the broadly Kantian view holds that mental representations essentially constitute judgeable contents, there may be genuine non-propositional mental representational contents that consist in action potentiations. Such action potentiations, or *de actu* contents, would constitute something like an embodied recognition of affordances.[40] A fundamental form of practical judgement may act on *de actu* contents by selecting, rejecting, or otherwise ordering them for performance. On such a view, then, actions would quite literally be the conclusions of practical judgments. The next section will arrive at the same sort of proposal.

2.4   Representation, Informative Function, & Action

One might argue for the truth of the Descriptive Thesis as follows: P1) Representations must be informative. P2) Anything that is informative must describe or otherwise stand-in-for the state of affairs about which it is informative. P3) Since stand-ins have alethic success conditions, anything that is informative must have alethic success conditions. So, the description thesis is true.

Right away we must modify the argument since lots of things may be said to be informative that presumably are not representations. Indeed, anything called "natural information" -- which includes all kinds of correlations and correspondences, including merely causal correspondences-- may be said to be informative. But these things do not describe or otherwise stand-in-for the state

---

of one's own reasons for action as such, of one's own species-genus, i.e. participation in a human form, and that aim at a specifically human form of flourishing involving freedom/ exercise of own reason/ bringing oneself self into conformity with what the "laws" of reason require.

[40] See Gibson (1966, 1977, 1979).

of affairs about which they are informative. They may be said to be informative insofar as those relations may be exploited for some informative function (e.g., exploited by "interpreters").[41] This suggests the following teleosemantics inspired modification of the argument: rather than saying that representations must be informative in any old sense, say that representations must have the function of informing or of being informative.

But the assumption that all representations function to be informative is questionable. Some things that appear to be representations do not ultimately function to inform. For instance, I may utter what has the surface form of a declarative sentence-- a paradigmatic representation--, e.g. "That's a bear!" but that is in fact functioning more like a command-- "Run!". Commands and imperatives seem to be representations, but they do not function to inform the audience that something is the case, but rather to get the audience to do something, or perhaps to change the audience's preferences.[42]

However, even if a representation does not ultimately or primarily function to inform-- to stand-in-for and thereby represent a state of affairs-- one could argue that it must nonetheless involve an informing component. This is because, plausibly, the only way a linguistic imperative can (non-accidentally) achieve its non-informing functions is if it guides the audience by representing (standing-in for, describing, picturing, modelling, naming, tagging) what it is the audience is to do, or the state of affairs the audience is to bring about if the imperative is complied with, or to present a (possible) state of affairs so as to alter the audience's preferences. In other words, although linguistic imperatives may not *ultimately* have informing functions, they

---

[41] See e.g. Millikan (2017).

[42] See e.g. Starr (2020). Thanks to James Shaw for pushing me to be clear about this.

nonetheless have instrumental informing functions essentially. For if they did not have (and achieve) such instrumental informing functions, they could not achieve their ultimate non-informing functions. Perhaps, then, the argument for the view that anything that represents must stand-in-for what it represents is that this is the only way for a representation to have (or perform) an informing function, where having this function, at least instrumentally, is essential to anything that represents. But is it true that the only way for something to have an informing function is for it to stand-in-for its represented?

To answer this question, let's begin by asking: what is it to have an informing, or properly representational, function? According to the somewhat popular teleosemantic analyses,[43] it is for something A to be selected to bring about some effect B where this effect is successful iff A corresponds to a state of affairs C. For this definition to pick-out robust representations, the relevant effects are identified with broadly intentional behaviors. So, the idea is that A is a representation iff A was selected to bring about a behavior B where B is supposed to be appropriate (successful, fit) and where B is only (non-accidentally) appropriate when A corresponds to C. According to teleosemanticsts, when this happens, A stands-in-for C.

There are two things to notice about this analysis. First, this is a fundamentally practical analysis of representational function: something is a representation iff it functions to bring about an appropriate behavior. Second, notice the jump to the interpretation that A stands-in-for C. These things are worth noting because, as Tyler Burge (2010) and others (e.g., Hutto and Myin 2013, Springle 2019) have argued, teleosemantic analyses appeal to a practical function that does not (on its own) ground non-trivial truth, veridicality, or accuracy conditions. In other words, such a

---

[43] See e.g. Dretske (1988), Millikan (1984, 2004), Neander (1995, 2013), Papineau (1984, 1998), Matthen (1988, 2005).

function does not license inferring the existence of a stand-in, for there's no reason to think that something that performs such a practical function *must* attribute properties or attributes to particulars and therefore no reason to think that whatever performs this function has the properties specified by the descriptive thesis.

Now, the teleosemanticists might reply that their reductive analysis *is* successful-- they *are* entitled to infer that A stands-in-for C-- because the way in which the subject or agent registers information positions her to act intentionally. Such a reply assumes that information registration that guides action must function as a model or description of a state of affairs C in order to play its action-guiding role. [44] How else could information registration play the action-guiding role in virtue of which it is a representation? But there is another way.

Imagine that you and your friend are given the task of shaping a clay pot (like that sexy scene involving Patrick Swayze and Demi Moore in the movie Ghost (1990)). Only your friend's hands are allowed to directly touch the clay, but your friend is blindfolded. You're not blindfolded, but your hands are not allowed to directly touch the clay. Assuming that you can't change which of you are blindfolded and which of you can touch the clay, how shall you and your friend complete your task? I see two options:

> **Option I:** You describe what your friend is to do-- how she is to move her hands, in what direction, and when.

> **Option II:** You pull a Patrick Swayze-- you place your hands on top of your friend's hands and directly guide their motions.

---

[44] Burge (2010) denies this.

In this example, you are substituted for A, your friend's behavior for B, and the clay that you're to shape into a pot for C. The derivation thesis is at work in the assumption that the information that guides action must stand-in-for a state of affairs that causes it in order to guide appropriate action with respect to that state of affairs. This assumption entails that the only option is (I). But (II) is also an option. Information registration (or causal sensitivity) A can function to prepare and direct appropriate intentional responses B to C by *directly* priming and guiding particular responses when activated by C; inputs from C are not used to stand-in-for C but rather to directly guide an interaction with C. The derivation thesis makes it seem as if only (I) is an option, and thereby obscures the fact that, if what's essential to representation is that it plays an informing function, option (II) leaves room for representation, but of a different kind.[45]

Note also that options (I) and (II) are reminiscent of the two approaches to practical judgement discussed in the previous section. In line with (I), the more common view treats practical judgement as a species of theoretical judgement which operates on broadly propositional contents. Practical reason on this view is the process of arriving at a propositional attitude (belief or intention) that prescribes the way one is to act-- it's how Swayze arrives at the action description he recommends to Demi. In line with (II), Dancy (2018) treats practical judgement as an action and practical reason as what "takes us from the situation as we see it" (it might be better to say, "the situation as we take it in") directly to the preparation of an appropriately related action without moving through any intermediate stage belief or intention.

It seems, then, that we have arrived at two converging proposals: On the first, the essence of representation is to constitute a judgeable content. But the genus judgement may be

---

[45] See also Cummins (1996).

fundamentally practical, and the fundamental species of practical judgement might operate on *de actu* rather than propositional contents. On the second proposal, informative function is the essence of representation. But this function may be fundamentally practical, and at least a species of informative function might be played directly by actions (*de actu* contents) rather than stand-ins (propositional contents). In the chapters to come, I'll develop an analysis of representation that accommodates these converging proposals. As a preview:

If what is essential to representation is the function to inform, and if the function to inform is the function of positioning agents to produce situationally appropriate actions (responses to particulars), representational aboutness might be understood as a form of practical epistemic access (PEA) to facts:

> **PEA Def. Representation:** R is a representation of (about, directed at) a fact X iff R constitutes a form of PEA to X, where R constitutes a form of PEA to X iff, in having R, an agent is positioned to produce an intentional action (in a sense that extends to non-human animals) in response to X.

This is *The PEA Analysis of Representation*. It is an analysis of the *genus* representation. It says that anything that represents provides PEA; providing PEA is necessary and sufficient for representation. And mirroring options (I) and (II) above, there are two importantly different species of PEA:

> **Direct-PEA:** An agent has direct-PEA to X iff she's positioned to produce an intentional action that is a direct response to X, where a response is direct when its immediate target is X, that is, when the action consists in interacting with X itself.

> **Indirect-PEA:** Agent has indirect-PEA to X iff she's positioned to produce an intentional action that is an indirect response to X, where a response is indirect when its immediate target is something that takes the place of (stands-in-for) X-- something one acts on directly so as to act on X indirectly.

De actu contents are the kinds of contents that constitutively function to provide direct-PEA. The PEA analysis likewise accommodates the proposal that representations are essentially practically judgeable contents. On the PEA analysis, representation is a matter of having access to facts where that access consists in being positioned to act in response to those facts. If Dancy is right that practical judgement may select actions directly, then representations that provide direct-PEA, i.e., representations with *de actu* contents, can plausibly figure in such judgements. And at the same time, the analysis can accommodate the possibility that another species of representation (practical judgement) prepares or positions actions indirectly (Indirect-PEA) via an intermediary or stand-in (e.g., a proposition).

2.5   Fundamental Representation

So far, my objective has been to make room in logical space for a fundamentally practical, non-propositional way of understanding original representation (or the intentional relation). If I'm right that representations of this kind exist, and if they are related to intentional actions in the way I've suggested, then the descriptive and derivation theses are false. In this final section, I'll argue that there's further reason to think that the descriptive thesis is false.

The kinds of representations that are paradigmatically thought to have propositional contents are sentences. Sentences are things we produce for communicative purposes-- we make them, we hear or read them, and we interpret them. Plausibly, sentences come by their intentional properties and hence propositional contents via the intentional actions and social practices of

agents. If so, then the nature of their intentionality-- their propositional contents-- might be illuminated by appeal to the relevant actions and practices and the intentional powers of the agents.

So familiar representations with propositional contents are things we relate to as objects-- things we produce, sense, and act on (respond to). So, if the fundamental intentional relation involves propositions of the same sort, then presumably subjects are related to them as objects. This suggests that, if one holds that the intentional relation is a representational relation, and that the relevant representations are propositional, it will be difficult to avoid the propositionalist view according to which we are directly intentionally related to propositions. On the other hand, if one denies that we are related to the propositional contents that are involved in our fundamental intentional relationship to the world as objects, then these representations are importantly unlike the kind of representations-- sentences-- with which we began our theorizing.

And there appear to be several other important differences between the familiar kinds of propositional representations (sentences) and the kinds of propositional representations that could be involved in fundamental intentionality. First, if we are related to them as objects, then what are their vehicles? Sentences have concrete, sensible vehicles. Without those vehicles, we can't get at their contents. So how do we get at fundamental propositions? Do we spontaneously generate them? But then, where are they? Surely their vehicles are not neurologically realized, for then we'd have to have intentional access to our brains to get at them. Also, how do we generate them? Arguably, the capacity to mentally represent sentence-like entities is developmentally dependent on the capacity to employ sentences in a public language. What's more, how can we generate such contents when, in the case of language, they have an essentially social and conventional nature? Indeed, sentences seem to come by their intentional properties by virtue of the intentional and

social activities of agents. But if our fundamental intentional access to the world already involves propositions, then their intentional powers must not depend on prior intentional social interactions.

In short, if the representations that are constitutively involved in our fundamental intentional relation to the world are propositions, I'm not sure what they are, for they are so unlike the propositional representations with which I'm familiar, i.e., sentences (and perhaps pictures, maps, etc.) Fundamental propositional contents are either propositional contents subjects are not related to as objects, or propositional contents subjects are related to as objects somehow primitively-- before they're able to be intentionally related to anything else. For otherwise we're faced with a vicious regress: in order to be intentionally related to an object O, a subject must be intentionally related to a proposition P that represents O, but in order to be intentionally related to P, one must be intentionally related to a further proposition P' that represents P, and so on *ad infinitum*. So, if we're to avoid such a regress, the kinds of propositional contents that are constitutively involved in our fundamental intentional relation to the world must somehow be the sort of thing a subject can be primitively intentionally related to.

By my lights, the idea that such propositional contents are constitutively involved in our fundamental intentional relation to the world generates a lot of mystery. In order to explain in virtue of what subjects are intentionally related to objects, we now must explain in virtue of what subjects are, *ex hypothesi*, primitively intentionally related to an alien variety of propositions, and whence these propositions come by their intentional properties. Of course, one could argue that these propositional contents have their intentional contents primitively. But then it would arguably be more parsimonious to simply treat the intentional relation between subjects and ordinary objects as a primitive.

2.6 Conclusion

I have considered several possible motivations to endorse the descriptive thesis and found them all wanting. While I think the fundamental intentional relation is representational, the idea that the relevant representations are as the descriptive thesis describes them is neither the only option nor is it obviously a particularly good option if our aim is to illuminate original intentionality. The possibility that fundamental representations have *de actu* contents should be explored.

3.0 Chapter 2: Intentional Actions & Acorns

In this chapter, I will introduce my "problem-solving" account of intentional action together with a corresponding conception of psychological representations as "intentional action acorns." [46] I'll begin by clarifying the questions I take it an account of intentional action should (or would ideally) answer. (§1) In doing so, I'll distinguish between what I'll call "Type Question" and "Token Question." I'll also motivate the "connection desideratum" according to which an account of intentional action should explain the necessary connection between intentional states and intentional actions.

To situate the problem-solving account, I'll next discuss the two major theories of intentional action: The Causal Theory and The Teleological Theory. (§2) Roughly, the causal theory holds that intentional actions are physical events or processes that are caused (in the right way) -- in the "efficient," "Humean," or "producing" sense of cause, i.e., a sense of cause that's applicable to matter in general-- by intentional states, i.e., psychological representations like beliefs, desires, and intentions. Roughly, the teleological theory holds that intentional actions are teleological processes, where the teleological nature of such processes cannot be explained etiologically, i.e., in terms of how they're caused by things external to them.

Since the problem-solving account is a kind of teleological theory, I'll first motivate it by motivating the teleological theory more generally. (§3) I'll do this by arguing against the causal theory. I'll develop a layered critique of the theory aimed at showing that the causal theory cannot

---

[46] As noted in the introductory chapter, I'm treating rational action as a special species of intentional action.

answer the Type Question about intentional action. (§3.1) However, I'll consider several ways of salvaging the causal theory, and I'll ultimately suggest that it might answer the Token Question about *rational* action specifically. (§3.2)[47]

None of the issues that prevent the causal theory from answering the Type Question about intentional action apply to the teleological theory. However, I'll argue that existing versions of the teleological theory (at least those of which I'm aware) are likely to strike some philosophers-- including those I identify as my primary audience in the Introductory Chapter-- as less than fully satisfying. (§4) This is for two reasons. First, the existing accounts are explanationist: they account for intentional actions in terms of the distinctively teleological forms of explanations or concepts that apply to them rather than accounting for intentional actions in terms that would illuminate the reason distinctively teleological forms of explanations or concepts get a hold on them.(§4.1) Second, the existing accounts haven't been concerned to satisfy the connection desideratum: they don't articulate a substantive alternative to the propositional attitude model of psychological representations at work in the causal theory.(§4.2)

I'll introduce my problem-solving account of intentional action and my intentional acorn account of psychological representations to address the two ways in which existing teleological theories might be deemed less than fully satisfying. (§5) I'll begin by sketching the problem-solving account of action. (§5.1) The problem-solving account is a non-explanationist teleological theory of intentional action. According to the problem-solving account, intentional actions are

---

[47] I distinguished between intentional action and rational action in the introductory chapter. I'm treating rational action as a species of intentional action.

processes that themselves constitutively function to solve practical problems, where practical problems are defined in terms of primitive broadly biological facts I call "needs for flourishing."

Next, I'll introduce the notion of an intentional action acorn to provide an initial (because largely metaphorical) articulation of a substantive alternative to propositional attitudes (which I'll develop less metaphorically in subsequent chapters). (§5.2) On the proposed view, ascriptions of occurrent propositional attitudes (beliefs, desires, and intentions) *refer* to intentional action acorns but *do not accurately describe them*.[48] Like the propositional attitudes of the causal theory, intentional action acorns are physical particulars (or 'individuals' since they will often have more than a moment's duration) with intrinsic representational properties. However, in certain crucial respects, intentional action acorns are not well modeled by propositional attitudes.[49]

3.1 What Is To Be Explained

What is an intentional action? To guide us in answering this question, it is helpful to first ask: What is a characteristic mark of intentional action, or, what goes on when we ordinarily recognize intentional actions as such? Anscombe (1957) and many philosophers since have

---

[48] For simplicity I focus on occurrent propositional attitudes. In chapters 4 and 5 I'll explain how the intentional action acorn account extends to standing or "dispositional" propositional attitudes-- hint: the answer will involve what I call " instructive dispositions" which realize what I call " embodied instructive concepts."

[49] I am inclined to think that the relevant kind of processes should be metaphysically classified as continuants rather than occurrents. Perhaps there are different kinds of continuants, so the choice might not have to be between continuants and occurrents but instead different kinds of continuants. I plan to devote further thought to how to classify them. At the moment I do not feel prepared to defend any particular answer. (I'm intrigued by Stout's (1996) proposal to identify intentional actions with what he calls "Aristotelian Processes," but I' m not entirely sure I understand what these are on Stout's view. Some of his descriptions make Aristotelian Processes a natural fit with intentional action acorns, but less so with respect to other ways he describes such processes.)

suggested that a mark of intentional actions is that they are susceptible to certain kinds of questions which receive certain kinds of answers or explanations. The relevant sorts of questions are "Why?" questions, and the relevant sorts of answers are *reasons explanations*-- i.e., justifications, teleological explanations, or answers that cite normative (or rational or intentional) as opposed to merely causal types of reasons.

This, then, suggests an initial answer to our question: an intentional action is a happening that happens for reasons. But an intentional action is not just a happening that happens for reasons, but a happening that happens for *our* reasons; it isn't a mere happening, but rather something *we do*. The happenings that are our doings (intentional actions) happen because (in some sense(s) of "because") of the way they are related to our reasons. This, then, suggests that a theory of intentional action should answer the following sub-questions:

1. Fundamentally, what kind of happening is an intentional action? (e.g., an event or a process?)

2. What is it for that kind of happening to happen because of our reasons? That is, what grounds the intentional directedness of an intentional action?

3. What are the relevant reasons (and what makes them *our* reasons?)? That is, are they "objective" or "external" facts (normative reasons), and if so, how should these be understood? Or are they "subjective" or "internal" facts (motivating reasons), and if so, how should these be understood?

However, stepping back a moment, it appears that our initial question-- *What is an intentional action?* -- is ambiguous between:

The Type Question: What identifies an intentional action as such; what is constitutive of the *type*: Intentional Action? In other words, what essentially distinguishes the type of happening that constitutes an intentional action from other types of happenings?

&

The Token Question: What individuates some particular intentional action? In other words, what determines a token of the type of happening that constitutes an intentional action?

Once we distinguish these questions, our initial answer generates not one but two lists of sub-questions:[50]

The Type Sub-Questions:

1. What type of happening is an intentional action?

2. What grounds the intentional directedness of this type of happening?

3. What is the nature of the reasons that figure into the intentional directedness of this type of happening?

The Token Sub-Question:

1. What determines the intentional directedness of a token of this type of happening?

While I take it that theories of intentional action are aimed at answering the Type Question, the Token Question is in itself a philosophically interesting question. However, it is important to, on the one hand, keep these questions separate, and on the other hand, recognize the ways in which an answer to the Type Question(s) constrains the available answers the Token Question.

In addition to distinguishing between the Type and Token Questions about intentional action, it is important to distinguish between what I'm calling "intentional actions" and what I'm

---

[50] In a similar spirit, Setiya (2011) distinguishes the project of reducing intentional action to causes by psychological representations from the project of explaining how psychological representations can be efficient causes of intentional actions: " The target of the causal psychological theory is not doing f but [doing f] on the ground that p and its connection with doing f because one believes that p" (138).

calling "rational actions."[51] As I mentioned in the Introductory Chapter, I consider the latter a special species of the former. Just as there are Type and Token Questions about intentional actions, there are Type and Token Questions about rational actions. My focus will be on intentional actions. Arguably, the theories of intentional action I discuss are really theories of rational action. However, if intentional actions and rational actions are related as I have suggested, then a theory of rational action piggybacks on a theory of intentional action. So, I will treat them primarily as theories of intentional actions. However, if it seems important, I'll treat them as theories of rational actions.

Finally, I'll take it as a desideratum that the way an account of intentional action answers the Type and Token questions about intentional action should say something about intentional states as well. Specifically, they should illuminate why acting intentionally seems to entail that an agent has certain psychological states, i.e., intentions, beliefs, and desires.[52] Roughly, being intentionally directed at a reason means being (broadly) cognitively in possession of that reason. Hence, if one acts intentionally, then one acts for a reason, and if one acts for a reason, then one is cognitively in possession of that reason. Accordingly, if one acts intentionally, it follows that one is cognitively in possession of that reason. And when we are talking about an agent's cognitive possession of reasons, we seem to be talking about her intentional states. So, there's arguably a

---

[51] Importantly, "Rational" here is not an honorific; it does not mark the type of action as rationally successful but rather as a *subspecies* of intentional action that is distinctive of rational beings/life forms.

[52] Hursthouse (1991) argues that what she calls "arational actions" -- actions performed out of some kind of emotional response to something-- cannot be explained in terms of the sorts of belief-desire pairs Davidson thinks must figure into the rationalization and efficient causation of an intentional action. I think there are a number of ways to respond to her counterexamples, but there's enough going on in this dissertation that I've opted not to discuss them here.

necessary connection between intentional states, reasons, and intentional actions.[53]   Moreover,

there seems to be a causal dimension to this connection. As Davidson (1963) notes,

> A person can have a reason for an action, and perform the action, and yet this reason
> might not be the reason why he did it. Central to the relation between a reason and an action
> it explains is the idea that the agent performed the action *because* he had the reason. (1963,
> 9; italics mine)

The relevant notion of "because" seems to be in some sense causal.[54] Now this can be

challenged. Dancy (2000), for instance, offers an alternative interpretation:

> The most direct response to Davidson [is] that the difference between those reasons
> for which the agent did in fact act and those for which he might have acted but did not is
> not a difference in causal role at all. It is just the difference between the considerations in
> the light of which he acted and other considerations he took to favor acting as he did but
> which were not in fact ones in the light of which he decided to do it. (163)

But one may ask for an account of the difference between "the considerations in the light

of which he acted and other considerations he took to favor acting as he did but which were not in

fact ones in the light of which he decided to do it." If acting "in the light of a reason" is just a

matter of taking that reason to justify the action, then this can't account for the notion of "because"

to which Davidson draws our attention. As Setiya (2011) notes, one might deny that the difference

between having a justification for acting and acting on it can be given a further account, and also

deny that it needs one. One might, in other words, take the relation "in the light of" or "on the

ground that," as primitive (145).[55]

---

[53] In a somewhat similar vein, Setiya (2011) argues for a psychological-causal theory of acting for reasons on the basis of 1) the necessary truth of conditionals linking claims about agents' reasons to psychological explanations and 2) a metaphysical argument against brute necessities (138-139).

[54] Setiya (2011) argues that it has to be efficient causal, as in the same kind of causation that's going on in deviant causal chains. I will not argue against Setiya directly on this point. Instead, I'll present a different way in which to understand the causal nature of the connection between psychological representations and intentional actions. On the view I'll propose, psychological representations "cause" intentional actions in roughly the same sense that an acorn "causes" an oak tree. Efficient causes are separate from their effects in a way that acorns are not separate from oak trees. Deviant causation involves mere efficient causation.

[55] As Setiya (2011) notes, at one point in *Practical Reality*, Dancy comes close to acknowledging this, asking 'How should we explain the fact that, where the agent's reason for acting is that p, the agent must believe that p, if not by saying that the agent's reason for acting is

This means that what counts as satisfying the "Connection Desideratum" is itself a controversial issue. However, as I mentioned in the Introductory Chapter, I'll be privileging the rich correspondence view's approach to illuminating agential and representational intentionality. On the rich correspondence view, satisfying the connection desideratum means that if a substantive account of psychological representation is available, it must be developed. I'll also privilege the rich correspondence view in considering what counts as a complete or fully satisfying answer to the Type Question about intentional action.

3.2 Theories of Intentional Action

There are two major theories of intentional action, instances of which admit of variations most of which don't matter for present purposes.[56] One is The Causal Theory of Action. The other is The Teleological Theory of Action. Before discussing these theories in more detail, two preliminary clarifications are in order.

First, as I noted above, I take it that both theories are aimed primarily at answering the Type Question(s). The tasks of the discussion in §3 and §4 will be to consider whether or the extent to which these theories successfully answer the Type Question(s), but I will, to some extent, also

---

''really'' that he believes that p?' (Dancy 2000, 126) Dancy offers interpretations of 'A is doing f because he believes that p' on which it is not the specification of his reason for acting. Instead, it gives an enabling condition of acting on the ground that p, or is to be read 'paratactically': 'He is doing f because p, as he believes' (Dancy 2000, 127– 30). So Dancy's question is what we are saying when we purport to explain action in terms of belief; he does not attempt to show that what we are saying must be true when A is acting for a reason. So, as Setiya notes, on Dancy's account, "it remains mysterious why believing that p is a condition of acting on the ground that p." (34)

[56] I'll distinguish some different versions of the causal theory in §2 and §3 and I'll distinguish some different versions of the teleological theory in §2 and §4.

consider the extent to which they can be extrapolated or interpreted so as to successfully answer the Token Question. In the present section, I will only exposit the way they each propose to answer the Type Question(s).

Second, it is important not to be fooled by the names of these theories. Both assume that intentional actions have a teleological dimension as well as a material and hence physical or efficient causal dimension. The difference between these theories is how these dimensions figure into their accounts of intentional action. There are importantly different things a theory of action that calls itself "causal" or "teleological" can mean by those terms. For instance, one might adopt Aristotle's distinction between the four causes, and call each of the theories any one of those causes figures in as "causal." By "causal theory" I have in mind something much more specific -- I'll give it a definition momentarily. Likewise, a theory of action might call itself "teleological" just because it acknowledges that intentional actions have an essentially teleological dimension. However, as we'll see in a moment, a theory that attempts to externalize that dimension or to reduce it to something non-teleological does not count as teleological as I'll be using the term.

3.2.1 The Causal Theory of Action

First, what is a causal theory? Some theories of intentional action are advertised as causal that do not satisfy my criteria for a causal theory. The key criterion for counting as a causal theory, as I'll be using the term, is commitment to what the Introductory Chapter dubbed "the derivation thesis." According to this thesis, the intentionality of action causally derives from the intentionality of psychological states that stand outside of the intentional action. The essence of a causal theory

is thus that it grounds the intentionality of action in efficient causal relations between physical events or processes and intentional states.

Causal theories centrally involve a teleological element, and, as we'll see below, teleological theories centrally involve causal elements. The key difference between causal theories and teleological theories as I'll distinguish them is that the former but not the latter ground the teleological (end-directed) nature of intentional actions in their etiology. According to the causal theory, the reasons we cite in our teleological explanations of intentional action are the contents of an agent's beliefs, desires, and intentions that cause and/or causally guide and control the agent's intentional actions. On this view, intentional actions have a teleological nature because they causally derive from the agent's psychological states with such contents. It is because the psychological states that causally give rise to and/or shape intentional actions are themselves end-directed that intentional actions are end-directed. And it is because psychological states belong to subjects and constitute their grasp (representation) of objective reasons (facts) that intentional actions are essentially authored by agents and connected with *their* reasons. The causal theory does not merely hold that states of having certain pro-attitudes and corresponding means-end beliefs are among the causes of our actions. Rather, the causal theory holds that an agent does what she does, when her doing is an intentional action, by virtue of having certain pro-attitudes and beliefs, i.e., psychological states with (respectively) a conative and a cognitive nature with certain propositional contents. The propositional contents and directions of fit of these states determine

the specific character of the events or processes they cause when those events or processes constitute intentional actions.[57]

Accordingly, I consider it a commitment of the causal theory that the relevant mental states are propositional attitudes in a substantive sense: they are causally efficacious entities with semantic properties, where both their causal and semantic properties are essential to the way they produce intentional actions. The causal properties of propositional attitudes account for the way in which they give rise to intentional actions. The semantic properties of propositional attitudes are that in virtue of which the intentional actions they cause (control, guide) and cause to exist and/or to have a certain shape come by their rational (rationally evaluable) properties.

Two major issues distinguish different versions of the causal theory. The first concerns the nature of the intentional states which cause actions. The questions at issue include: Are intentions a distinctive kind of attitude or just belief-desire pairs? Are intentions more like knowledge (and if so, propositional knowledge or knowledge-how or both?) or beliefs or more like desires (pro attitudes) or all-out judgements "that doing such-and-such is all out desirable"? What is the nature of the content of an intention? -- is it "I shall φ", "I am φing", "I should φ", "I am trying to φ" or something else? Are there different kinds of intentions, e.g., prior intentions, intentions in action, motor intentions (e.g., Searle 1980)? If so, how are they related to each other? Are propositional attitudes exhaustive of the kinds of entities we need to call on to explain how intentional actions come by their shapes, or do we also need further entities, e.g., plans (Bratman 1987, 1999)?

---

[57] Fred Dretske (1988) brings out the role propositional attitudes are supposed to play on the causal theory by contrasting an intentional action with a non-intentional event that an intentional action might cause. He uses the example of a soprano singing an aria which causes a glass in the room to shatter. The contents of the soprano's propositional attitudes figure essentially into the explanation of her singing the aria. In contrast, the lyrical contents that are part of the singing play no essential role in the explanation of the glass breaking. All that's relevant to explaining the fact that the glass breaks are the acoustic (or "triggering causal") properties of the singing.

The second issue concerns the nature of the relevant causal relation(s). This is perhaps the biggest issue for the causal theory, because it's been known since the theory's conception that it's vulnerable to "deviant" or "wayward" causal chains. That is, it is possible to devise scenarios in which:

(i) The agent intends to φ,

(ii) What happens is in conformity with the agent's representation of what is to happen, &

(iii) That event was caused by the intention to φ, but where what happens is accidental rather than intentional.

Consider, for instance, the following example from Alfred Mele (1992):

> A philosopher intends to knock over his glass of water in order to distract his commentator. However, his intention so upsets him that his hand shakes uncontrollably, striking the glass and knocking it to the floor. (182)

In other words, a philosopher has the intention with the content "that I knock over my glass of water in order to distract my commentator." That very intention is causally responsible for effecting a physiological stress response which includes shaky hands, his shaky hands make it the case that he knocks over his glass (presumably thereby distracting his commentator). So, his intention is causally responsible for a state of affairs that happens to satisfy that intention's propositional content.

Deviant or wayward causal chain counterexamples like this show that an action φ may be caused by a propositional attitude with content "that I φ" (or some such) without thereby being an intentional action. As Davidson (1973) puts it,

> If the agent does x intentionally, then his doing x is caused by his attitudes that rationalize x. But since there may be wayward causal chains, we cannot say that if attitudes that would rationalize x cause an agent to do x, then he does x intentionally. (79)

So, the causal theory can't just define intentional actions as actions that are caused by an agent's propositional attitudes. Rather, it must qualify the relevant kind of causation: intentional actions are actions that are caused by an agent's propositional attitudes *in the right way.* Until there's a successful account of what "in the right way" amounts to, the causal theory does not supply sufficient conditions for, and therefore fails to provide a complete analysis of, intentional action.[58]

So, the major challenge for the causal theory is to show that the problem of causal deviance or wayward causal chains can be solved. A popular way of trying to do this draws on Bratman's notion of "conduct-controlling pro-attitudes" (1987). Where on Davidson's (1963, 1971,1980), Goldman's (1970), and Audi's (2001) theories an intentional action must be caused (in the right way) by an agent's intentional states (motivating reasons), on Bratman's, theory, what makes an action intentional isn't that it's caused by one's psychological representations, but rather that it is controlled and/or guided by one's psychological representations (see also Bishop 1989, Brand 1984, Mele 1992, 1997, Thalberg 1984). Of course, these options are not mutually exclusive: one might try to reduce intentional action in terms of both the causation by psychological representations and a continuous kind of causal guidance and control by psychological representations. I'll return to such proposals in the next section.

---

[58] Davidson is usually credited with putting forward the Ur causal theory. However, as I'm using the term, Davidson's theory was not, strictly speaking, a causal theory. Davidson was the first to recognize that his theory does not provide necessary and sufficient conditions, and he didn't care to change this: "What I despair of spelling out is the way in which attitudes must cause actions if they are to rationalize the action" (1973, 79). Davidson didn't mind having a non-circular solution to the problem of causal deviance because he wasn't trying to give a reductive account. In contrast, what I'm calling the causal theory is trying to give a reductive account.

3.2.2 Teleological Theories of Action

Teleological theories reject the possibility of explaining the means-end form of action in terms of external causes (however much "guidance" or "control" they exert). They take it that the essence of intentional action is the means-end form that's exposed by answers to Anscombe's "Why?" questions, and they take it that this form is intrinsic to intentional actions.

The strategy of most teleological theories is to undermine the premise that explanations of intentional actions in terms of beliefs, desires, and intentions (henceforth "reasons explanations") are efficient causal explanations in terms of such psychological representational states. Part of this strategy is to shift the focus: The causal theory tends to start with prior intentions-- intentions one seems to have before one begins acting-- and completed actions. When one begins theorizing in this way, it's not surprising that the picture of the relationship between intention and action that presents itself consists in a relation between *intentions* and *events* where intentions are psychological representations which can exist wholly apart from the actions they represent.[59] Shifting the focus to intentions-in-action, that is, to the intentions we ascribe to agents in the midst of acting, and, correspondingly, to actions-in-progress (actions primarily as *processes)*, rather than

---

[59] The causal theory traditionally treats actions as events, but it is open to causal theorists to treat actions instead as certain kinds of relations. For instance, Bach (1980) argues that actions are instances of relations of bringing about or making happen, and he takes it that such relations are to be analyzed as causal relations between events, and where the event that brings about the event that's brought about is an intentional psychological event. It is also open to causal theorists to treat actions as certain kinds of processes. On a causal theory that treats actions as events, what distinguishes the events that are intentional actions from other events is their etiology-- the intentional nature of the events that cause and/or control and guide them. Likewise, on a causal theory that treats actions as processes, what distinguishes the processes that are intentional actions from other sorts of processes is their etiology-- the intentional nature of the events that cause and/or control and guide them. Counterexamples involving deviant or wayward causal chains work by intercepting the etiology so that "the right thing happens but in the wrong way," that is, the propositional attitude represents a result which it then causes to be actualized, but in a way that prevents it from being intentional. So, treating an intentional action as a process in this sense does not protect the causal theory from counterexamples involving causal deviance.

completed actions (actions primarily as *events*), is a first step towards undermining the causal theoretic premise.[60]

So, for instance, following Anscombe (1957), George Wilson (1989) and Carl Ginet (1990) argue that reason explanations are grounded in an agent's intentions in action and that reasons explanations are not causal explanations. Their argument is based on an analysis of ascriptions of intentions in action. On their analysis, ascriptions of intentions in action are *de re* propositions that say *of* a particular act of ϕing that it was intended by its agent to Ψ (by means of ϕing). They argue that such propositions constitute non-causal reason explanations of why the agent Ψed. Wilson further analyzes the content of intentions in action as "The agent's act of ϕing was directed by him at [the objective] of Ψing." This analysis makes explicit the teleological character of ascriptions of intentions in action. Since intentional actions are essentially goal directed, one can explain such actions by mentioning a goal or purpose of the behavior. That is, one can give a teleological explanation of the action. And that, Wilson claims, is precisely what ascriptions of intentions in action do.

But this approach doesn't go far enough. Causal theorists can agree that justifying explanations are teleological explanations (explanations in terms of goals or purposive explanations) but argue that teleological explanations are themselves to be analyzed as causal explanations in which the agent's primary reason(s) (or pro-attitudes or desires, beliefs, and/or intentions) for (or in) ϕing figure as efficient causes (and/or guiding and controlling efficient causes) of the act of ϕing. But other teleological theories go further.

---

[60] The causal theory begins with prior intentions and extends that model to account for other kinds of intentions such as intentions in action. The teleological theory begins with intentions in actions and extends that model to account for other kinds of intentions such as prior intentions.

For example, like Wilson and Ginet, Brian McLaughlin (2012) argues that reasons explanations are teleological-- they essentially explain actions in terms of the purposes or goals for which they are performed-- and not a species of causal explanation. But unlike Wilson and Ginet, McLaughlin denies that teleological explanations are *grounded* in agents' intentions in acting. Intentions in action, after all, can seem to call out for a psychological explanation, and so likewise purposes and intentions. Instead of grounding teleological explanations in intentions, McGlaughlin grounds such explanations in what agents are trying to do. On his view, when we say that an agent ϕed in order to (or for the sake or purpose of) Ψ, what we're saying is that in ϕ-ing, the agent was trying to Ψ. And trying is not to be cashed out psychologically, e.g., in terms of an agent's intentions in acting. In other words, "in order to" does not translate to "by virtue of having an intention to Ψ or the goal of Ψing." Instead, if in ϕ-ing an agent was trying to Ψ, then ϕ-ing is identical with or a proper part of Ψing. So, on McLaughlin's view, reasons explanations are teleological explanations, and teleological explanations are constitutive-- they offer a redescription of the actions they explain-- rather than causal

Michael Thompson (2008) develops a somewhat similar but also importantly different view. Thompson holds that intentional actions belong to a *sui generis* kind of broadly biological activity that's marked by the applicability of a distinctive species of predicates which appear in what Thompson calls "natural-historical judgements." Examination of such judgements reveals that actions and reasons are connected formally (logically or teleologically) by "for the sake of" or "means-end" relations which hold first and foremost between the practical activities that ultimately characterize a "form of life." Thompson's view is similar to McLaughlin's in that Thompson holds that reasons explanations *ultimately* refer not to propositional attitudes (psychological representations) of intending, desiring, wanting etc., but rather to the larger actions

or activities of which they are a part. However, unlike McLaughlin, Thompson does not deny that teleological explanations appeal to propositional attitudes. Instead, Thompson offers a broadly behaviorist account of ascriptions of intentions, wants, desires, etc.[61]

Thompson allows for a sense in which propositional attitudes—in particular, "intentions" and "desires"—can be invoked in some rationalizations: the question, "why did you go to the store?" can sensibly be answered by "because I wanted to buy eggs." But on Thompson's view, these ascriptions amount to "sophisticated" descriptions of the agent's larger action or developing process (plus additional information). Such sophisticated descriptions have their home in what Thompson calls a "sophisticated" action theory and a corresponding practice of action explanation which is ultimately parasitic upon what he calls the "naïve action theory" and corresponding explanations. The naive theory rationalizes or explains actions not by means of descriptions of agents' psychological states but rather by locating actions in the "developing process" or context

---

[61] Perhaps Thompson leaves room for "mental states" as material causes of actions in the sense that brain states are part of the explanation of actions *qua* material processes. But they have no real role in explaining intentional action *as such*, i.e., at the intentional, rational, or logical level. A substantive account of propositional attitudes requires that they span (and somehow unite) both levels: they both rationalize and cause actions. One might also read Thompson as suggesting that reasons explanations provide what Anscombe calls an "enlarged description":

Since I have defined intentional action in terms of language-- the special question "Why?"-- it may seem surprising that I should introduce intention-dependent concepts with special reference to their application to animals, which have no language. Still, we certainly ascribe intention to animals. The reason is precisely that we describe what they do in a manner perfectly characteristic of the use of intention concepts: we describe what further they are doing in doing something (the latter description being more immediate, nearer to the merely physical): the cat is stalking a bird in crouching and slinking along with its eye fixed on the bird and its whiskers twitching. The e*nlarged description* of what the cat is doing is not all that characterizes it as an intention (for *enlarged descriptions* are possible of any event that has describable effects), but to this is added the cat's perception of the bird, and what it does if it catches it. The two features, knowledge and *enlarged description*, are quite characteristic of description of intention in acting. Just as we naturally say 'The cat thinks there is a mouse coming,' so we also naturally ask: Why is the cat crouching and slinking like that? and give the answer: It's stalking that bird; see, its eye is fixed on it. We do this, though the cat can utter no thoughts, and cannot give expression to any knowledge of its own action, or to any intentions either. (Intention 1957, §47, p. 86, italics mine)

So one might read Thompson as suggesting that reasons or rationalizing explanations answer "Why"-questions by locating actions in "developing processes" such as larger actions to which they stand in part-whole relations. The intentional action acorn account of subjective reasons I'll offer below has a similar way of understanding the relationship between the reasons that figure in rationalizing explanations and intentional action. For on my view, subjective reasons are stages in the developing process of an intentional action.

of agents' other actions and activities (ultimately, the activities that are characteristic of her form of life). In the naive theory, the question, "why are you cracking eggs?" is sufficiently rationalized with the answer "because I'm baking a cake" which makes no reference to the agent's intentions, wants, or desires but rather places the action of egg-breaking in a sense-making context.

Thompson draws upon the linguistic distinction between "perfectives" and "imperfectives" to argue that "sophisticated" rationalizations that involve "practical-psychical verbs," e.g., "wants," "intends," and "tries," function "to express certain forms of imperfective judgment" and to give information about an action in addition to what an action is for (131). This means we can understand explanations like "I'm cracking an egg because I want to bake a cake" and "I'm cracking an egg because I intend to bake a cake" as ways of saying "I'm cracking an egg because I'm breaking a cake." For Thompson, then, sophisticated rationalizations ultimately occupy the same "categorical space" as actions, "the space of kinesis, if you like, and not of stasis" (134). Actions are rationalized by actions and *processes* in the same categorial space as the actions that the agent (presumably) takes to be rationalized; they're not rationalized by psychological *states*. Thompson's discussion of "imperfective judgements" also positions him to respond to the objection that intentions, wants, desires and the like cannot be explained in terms of actions because they are not always accompanied by actions. This is because Thompson's examination of imperfective judgements shows that an agent may be acting even when there's no outward sign of it, for instance, when an action is interrupted. Indeed, an intentional action process might be interrupted such that it's never completed (perfected)-- what the sophisticated action theory would

call a "trying." And such a "trying" could occur without there ever having been an outward sign of the action process.[62]

A common feature of these various versions of the teleological approach is that they all focus on action explanation-- they ultimately hold that intentional actions have a special kind of form, and they go about illuminating this form by analyzing, albeit in different ways, the teleological explanations (or ascriptions of intentions or natural historical judgements) that make this form explicit, and arguing that such explanations are not reducible to efficient causal explanations involving propositional attitudes, or that the kinds of predicates that apply to intentional actions are of a fundamentally different kind from those involved in efficient causal explanations. So, I'll refer to these views as "explanationist." Depending on one's broader philosophical orientation, the explanationist nature of these views may or may not be an issue. Philosophers who identify with the therapeutic view discussed in the Introductory Chapter probably won't find it an issue. But for reasons I'll explain below (§4), philosophers who identify with the rich correspondence view are unlikely to be fully satisfied by explanationist views.

---

[62] The reason my discussion of Thompson's view delves into relatively more detail is that it is, in many ways, an inspiration for my problem-solving account of action and my intentional action acorn account of intentional states. So describing these details of his view is a way of preparing the reader for those views.

3.3 A Critique of The Causal Theory

I'll begin by developing a layered critique of causal theories -- one problem will be layered on top of another, and another on top of that (§3.1). I'll then consider ways in which the causal theory might be salvaged. I'll ultimately argue that it fails to do what it aims to do: it fails to answer the Type Question(s) for intentional actions (§3.2).

3.3.1 A Layer Cake of Criticism

The Causal Theory has some intuitive pull. Anscombe's question "Why?" provides a criterion for picking out intentional actions. This criterion reveals their teleological nature -- the fact that intentional actions are means (for ends). It also highlights the fact that we explain intentional actions by citing agents' reasons, or what they took to be their reasons -- their beliefs, desires, and intentions, i.e., representations of facts. This suggests that what it is for an intentional action to be an agent's means to her ends is for the action to be connected in some essential way with the agent's reasons, or with her representations of them, or, alternatively, that an agent's mental representations just are her reasons. The causal theory proposes to understand the way in which an action (event) is connected with an agent's reasons (or representations of them) as causal: an agent's representations of her reasons causally give rise to and/or guide an event, and this is what accounts for an action's means-end form.

I think the intuitive pull of the causal theory is connected with the intuition that the nature of intentional action presents something of a mystery to be solved. On the one hand, an intentional action seems to be an event or process. On the other hand, unlike other events or processes, an

intentional action is intentionally directed at an end; it has a means-end form. The seeming mystery is: how can an event or process be intentionally directed at an end since that is something that does not yet exist? Part of what is at issue here is skepticism about the possibility of "backward causation" -- not yet obtaining states of affairs causally influencing presently obtaining states of affairs. (Kim 1989) The causal theory suggests a way of accounting for the way in which an event or process can be directed at not yet existing states of affairs. First, it notes that intentional states (psychological representations) are thought to have the power to be about things that don't exist. So, it suggests that the way in which an intentional action is directed at an end can be explained in terms of a psychological representation of that end. When faced with the question, "how might a psychological representation of an end make it the case that an event or process is directed at that end?" the causal theory also has an initially plausible answer: The representation efficiently causes the event or process and guides it in a way that would satisfy the representation's content by making it the case that the end obtains.

I think, however, the initial plausibility of this explanation hangs on a problematic analogy. The analogy is between psychological representations efficiently causing processes or events to have means-end form and agents producing tools and other functional objects that have means-end form. It's true that humans and other animals can produce things that have means-end functions-- humans produce houses, birds produce nests, foxes produce dens, etc. And no doubt the production of such things has an efficient or Humean or productive causal aspect-- there's no way to make a house, nest, or den without physically interacting with the relevant building or burrowing materials. And taken on its own, the physical interaction is of the type that doesn't raise any philosophical eyebrows.

However, there is reason to doubt that anything but a creature already capable of intentional action could produce something that counts as a house, nest, or den. In other words, there is reason to doubt that the means-end or functional form of these items can be explained just in terms of their physical causal etiology-- the kind of physical causal chain of events that doesn't raise any philosophical eyebrows. It seems that something with means-end form can't just be something that an intentional state causes in any old way. Rather, something with a means-end form has to be something that an intentional state intentionally causes. What's crucial, in other words, is that things with means-end or functional forms are caused *by agents*. If so, then it seems that the causal theory is attributing agential powers to intentional states insofar as it's attributing to them the power to give things purposes. If intentional states really do have the power to give things purposes (functions and/or means-end forms), then they have the power of agents, and the very power they were invoked to explain. So, it appears that the causal theory is attempting to explain what it is for an agent to act in terms of the actions of her intentional states. In other words, the explanation appears to invoke a homunculus, and consequently faces a version of the fundamental problem facing any such explanation: if we try to account for the intentional actions of intentional states in the same manner that we've tried to account for an agent's intentional actions, we're off on a regress.

That an intentional action can't just be something that an intentional state causes in any old way is precisely the problem counterexamples to the causal theory involving deviant or wayward causal chains make explicit. For in such counterexamples, an intentional state causes something to happen, and even causes something to happen that seems to satisfy the state's content, but "in the wrong way." I have just suggested that the causal theory will not be able to deal with such counterexamples because only agents (intentional actors) can, in producing something, imbue it

with a means-end form. So, specifying the "right way" for an intentional state to causally bring about an action such that it is intentional will inevitably smuggle intentional agency into the *explanans*.

Consider, for instance, the proposal that the "right way" for intentional states to be causally related to an event or process that will thereby constitute an intentional action is for an intentional state to causally "control" and "guide" the production of the event or process (see e.g., Bratman 1987, Bishop 1989, Brand 1984, Mele 1992, 1997, Thalberg 1984, Mele & Moser 1994, Wu 2016). This isn't yet a theory until we have a way of specifying what guidance and control come to conceptually. Importantly, this means one cannot explain what the relevant kind of guidance and control come to in terms of some causal story about the underlying processing.[63] Just because we can describe an activity in causal terms doesn't mean that an activity is merely causal-- intentional action being the obvious case in point. In other words, in determining whether intentional action can be explained in terms of causation by intentional states, the empirical details of how guidance and control are causally realized are simply beside the point. What matters is whether guidance and control are genuinely just physical causal activities, or whether they are already intentional activities. And *prima facie*, the concepts of guidance and control are already intentional concepts. Supplying them with an empirical mechanical story does not change this. Guidance is towards a goal. And what we do intentionally we can do in a more or less controlled way. The notion of control gets a hold on a process or event already identified as an intentional action. And that we (or our mental states) guide our actions as they unfold just seems to be another way of saying that we are the ones who do them.

---

[63] Some recent attempts to deal with causal deviance (e.g., Wu 2016) neglect this point.

Of course, at first blush, the guidance and control proposal looks like it might provide a sufficient condition for intentional action. This is because, where one finds guidance and control of the relevant sort, one finds intentional action. But of course, this fact is also explained by the thesis that guidance and control are already intentional concepts--- if they are essentially attributes of already intentional actions. One reason to think that this latter explanation is the right one is the fact that it seems that one can perform an intentional action even when that performance does not exhibit much in the way of guidance and control. Patients with Parkinson's disease can and do still act intentionally despite significant reduction in their motor control capacity. In addition, different types of actions can involve very different kinds of guidance and control. For instance, what's involved in writing a paper is quite unlike what's involved in hitting a baseball and running to first base. Moreover, different token actions of the same type can involve very different kinds of guidance and control-- there are many different ways to eat lunch. These points suggest that a principled and informative purely causal or mechanical characterization of the kind of guidance and control action requires will be very difficult, if not impossible, to come by.

The causal theory, and the sorts of modifications of it just discussed, can seem plausible when one thinks of intentional actions as if they were fundamentally material events. For then what distinguishes such events could just be the fact that they have a certain kind of cause, namely, causes that represent certain ends, where these causes turn an ordinary material event or processes into a means to an end because they causally produce, control, guide and otherwise create the event/process in such a way that the represented end is realized. In other words, the causal theory can seem quite plausible if one thinks that actions are really just material events/processes plus something extra. This makes it sound as if there could be, e.g., intentional eatings and non-intentional eatings, or intentional greeting and non-intentional greetings, etc.

Now some might say that there are such things. The most obvious reason one might think that there are both, e.g., intentional eatings and non-intentional eatings, is that eating appears to involve a kind of physical activity that can happen intentionally or not. It is of course possible to "unintentionally" eat something, where this is distinct from, say, merely unintentionally swallowing something. In other words, we may say someone "unintentionally" ate something when the individual in question genuinely ate something. But what it means when we say this is not that eating was just a physical activity; it's still an intentional type of activity, still a means to an end. To say one ate something unintentionally or not on purpose or by accident is to say that the eating was done by mistake. In other words, we can act intentionally and in error; our intentional actions can manifest ignorance. One may be intentionally but absent-mindedly eating something, say, at a party, and accidentally eat something off of another person's plate.[64] This is a case where it would be appropriate for one to say "I'm sorry, I ate that unintentionally" by which one would mean that they were not aiming at eating off of that person's plate, not that they did not aim at eating something, and not that they were making eating motions without aiming at nourishment or satisfaction of a craving or some such. So, we can accommodate cases in which

---

[64] Intentional action is sometimes contrasted with habitual or automatized actions. But plenty of intentional actions are habitual and largely automatized. When we deny that an action done out of habit or an action done, as it were, automatically, is intentional, and when the action in question is a genuine action and not a spasm or some such, then what we are denying is not that the action is an intentional type of action. Rather, we are *admitting* or *confessing* that the action was done "without thinking" meaning without the appropriate oversight of reason (one's rational faculties). Had our reason been appropriately operative, we'd have realized that what we were doing was not appropriate. We were indeed aiming at something in doing what we were doing, but that was not the thing that should have been our dominant aim in the situation. And had we been "thinking," we would have realized this and acted differently.

One might think that actions that are done without properly exercising our rational capacities are not rational actions in my sense (i.e. a special species of intentional action). But this doesn't follow. The fact that we can notice and apologize for our "unthinking" actions is evidence of their rational nature. There is, of course, a sense in which such actions are not "rational," namely, the honorific sense, or the success-term sense. What we were doing was not the overall right or rational thing to do, and this is what we are admitting to when we say our actions are not intentional in the kind of case that's at hand.

we say we ate something "unintentionally" while maintaining the idea that eating is essentially an intentional kind of process or event.

This seems like the right result. *Prima facie*, eating is not merely a physical activity but rather an intentional or teleological physical activity. Eating constitutively aims at something, i.e., nourishment, or satisfaction of appetite or craving. If so, and if such end-directedness is the essence of intentional action, then there is no such thing as a non-intentional eating, but only non-intentional eating-like behavior.[65,66] So, it's far from obvious that eating should be identified with a mere pattern of activity that sometimes is and sometimes is not accompanied by an intention. Rather, eating, walking, talking, shopping, cutting wood, writing, etc. are all intentional types of activity. It is possible to make similar motions-- for instance, to pretend to eat, walk, talk, etc.-- but this is not the same as actually doing the relevant actions (indeed, when one pretends one is doing a different type of intentional action).

If this is right, then the causal theory is barking up the wrong tree. For the causal theory sets itself the task of explaining how a merely physical happening (event or process) comes to be an intentional type of happening-- the type of happening that's a means to an end. But I suspect many who are sympathetic with the causal theory won't yet be convinced. So, I'll now try to spell

---

[65] One might object that even if we accept that eating is an essentially teleological activity, it can happen unintentionally. However, I'm inclined to think that this objection confuses intentional action with rational action. Perhaps a person could eat in their sleep, where hunger is what drives them. Then I agree that we'd be inclined to say they were not acting intentionally. But what we mean is that they were not acting rationally (self-consciously) and thus intentionally in the way humans typically act intentionally. If it were an animal and the sleep activity was completely automatic and insensitive to other situational facts, it wouldn't be intentional either. But this is because being properly end-directed seems to entail sensitivity to whether that end is being fulfilled and to whether fulfilling that end now, as opposed to some other more urgent end, is the thing to do. If one eats compulsively, we say it is not an intentional action. What one is doing involves the same movements and interactions as eating does, but it isn't an eating. The fact that one is producing those movements may mean that certain of the same abilities are being exercised when one eats and when one merely seems to eat (when one "eat-motions"). But they are not exercises of the same capacity, namely the capacity for intentional action (though the relevant abilities may well be components of such capacities). Finally, it is of course possible to eat "unintentionally" as in the above example of eating off of someone else's plate at a party. But as noted, intentional actions can be mistaken and confused but still intentional.

[66] See also Anscombe (1989, p.111).

out in a bit more detail why causation by propositional attitudes cannot account for something's being a means.

First, what goes into being a means? There seem to be two components.[67] First, a means is a kind of (physically realized) process (method or mechanism involving certain sorts of entities and activities). Second, a means is for some end. In other words, M is a means iff:

(1) M is a method/ mechanism/ process that, if completed, brings about (or tends to bring about) X, and

(2) X is an end

Roughly, M is my means iff M is my method-- one I employ (or cause to exist), and M is my method iff x is my end. I mentioned in §1 that intentional actions essentially involve agents; an intentional action is someone's action (someone's doing). If intentional actions are means, means must be connected with agents, presumably by being connected with an agent's ends.

Now it might be thought that propositional attitudes can represent the movements (etc.) they produce as means. A first thing to note is that such a theory appears to be committed to our having the concept of a means (or intentional action) as such before we can intentionally act. After all, if we need to be able to represent something as a means to cause it to be a means, and if doing this is required to act intentionally, then it appears we must have the concept of a means before we can act intentionally. Now one might argue that one doesn't need to have a sophisticated concept of a means in order to represent something as a means, and perhaps this would help in the case of human children. But it's not obvious that it will help with human infants (unless the concept is

---

[67] This account of a means differs from Stout's (1996) account, which will be discussed below (§4). I think the two could be combined-- the main difference is that this account requires that a method that is a means brings about a state of affairs that is specifically an end. By my lights, then, this account of being a means is an improvement over Stout's.

innate and operational at birth, which is an empirical question), and it's even less obvious that it will help with non-human animals, unless we suppose that they have concepts of means, albeit thin ones. This is not to deny that non-human animals can in some sense appreciate things as means. The claim is rather that it's far from clear that non-human animals can *represent* things *as means*. Of course, this gets at a more general problem, namely, whether non-human animals can have propositional attitudes at all. So, for now I'll put this issue aside.

The issue I want to focus on is that efficient causation by propositional attitudes does not make something a means. Propositional attitudes may represent something they causally produce as a means. But representing something as a means doesn't make it a means. So, this is not sufficient. So now let's add causation. So, we have a representation of some movements as a means, and that representation causes those movements. Does that make the movements a means? I don't see why it should. Let us stipulate that a movement X with my hand is not intentional. Now say I have a mental representation of a movement X of my hand, and I also have a representation "X is a means," and together these cause X (my hand makes the relevant movement). Have I contradicted myself? Must X be an intentional action after all? I don't think so. It seems to me that all that has happened is that I have produced a movement that I also represent as a means. It's still the case that representing it as a means doesn't make it so. Just because that representation causally produces the movement that I represent as a means doesn't mean the movement is a means.

One might object that they *do* get the intuition that I contradicted myself. One must merely add that the fact that those mental representations cause X is no accident, because that is what those mental states function to do. But this suggests that mental states might need to be understood in terms of intentional actions, in which case causation by them can't very well ground intentional actions. This brings me to a further dimension of the problem with the causal theory.

According to causal theories, an intentional action is *my* intentional action because it is a way of satisfying the content of my desire that is caused to occur as a result of the corresponding desire. My desire produces activity that constitutes the means. First, I just want to note that this does not provide an account of the fact that the action is a means: it does not show that causation by desire is what makes the thing produced a means. Quite the contrary, it presupposes the notion of a means. But what I want to focus on now is the way the causal theory understands reasons, specifically what Davidson called "primary reasons," and in particular, desires.

I think the causal theory begins with the following seeming truth: for X to be my end is for me to desire X. But here we must be careful. One could say: for X to be my end is for me to desire X, because there's nothing more to desiring X than having X as my end. Then we might ask: what is it to have X as one's end? And one could answer that one has X as one's end iff one is disposed to try to bring it about that X happens. In other words, to have X as one's end is for one to be disposed to produce a means for bringing X about.

But the causal theory does not say that for X to be my end is for me to desire X because there's nothing more to desiring X than to have X as my end. Instead, the causal theory says that having a desire about X (desiring to make it the case that X obtains) just is what it is for one to have X as one's end. But now if we ask: "what is involved in having a desire *about* X?" it appears that the causal theory faces a serious problem. Propositional attitudes are individuated in part in accordance with their functional roles, and their functional roles are defined in terms of intentional actions. To have a desire about (representing) X is to have a state whose functional role is to bring it about that X given appropriately related beliefs, etc. In other words, a desire's functional role is defined in terms of causing something that is a means for bringing about an end X. So, we are back to the problem of presupposing intentional action in the analysis of intentional action because we

may not be able to give an account of the relevant kinds of mental states except in terms of intentional action.

At the very least, the causal theory runs into a serious intertheoretical tension: there's a tension between maintaining the causal theories of intentional action and the project of illuminating the nature of psychological states. This tension emerges in two ways: First, the causal theory generates a "no-circularity constraint" on accounts of psychological states: if an analysis of psychological states must appeal to intentional actions, then, if we assume a causal theory, both the account of psychological states and the account of intentional action will fail to be illuminating insofar as both will be circular. And there's good reason to think that an adequate account of psychological states will have to reference intentional action in specifying the nature of the kind of function that's constitutive of mental representations.

Second, the incompleteness of the causal theory that results from its susceptibility to deviant causal chains threatens to contaminate and infect analyses of psychological states. This is because there's reason to think that such analyses will not be able to specify the nature of the function that's constitutive of mental representations until they can specify the "right way" a putative mental state functions to cause intentional actions.[68]

The First Tension

The causal theory's analysis of intentional action reduces intentional actions to events or processes that are caused and/or causally guided and controlled by psychological representations. Consequently, the causal theory generates the following constraint on theories of psychological representation:

---

[68] These tensions are also discussed in Springle & Humphreys (2021).

The No-Circularity (NC) Constraint: An account of psychological representations cannot make (ineliminable) reference to intentional actions, on pain of circularity.

As a result of this constraint, the causal theory is incompatible with theories of what constitutes psychological states that reference intentional actions, for the causal theory renders substantive analyses of what constitutes psychological states that reference intentional action viciously circular. What's more, if an analysis of what constitutes psychological representations must make essential reference to intentional actions, then the causal theory will itself turn out to be viciously circular, since, in referring to psychological states, it will indirectly refer to intentional actions.

This is bad news because there's good reason to think that an account of psychological states must appeal to intentional actions. Consider, for instance, behaviorist accounts of psychological states which analyze psychological states in terms of behavioral dispositions. Of course, not everything that might be considered a behavior (sweating, reflexes, etc.) is plausibly a mental state (at least not in a robust sense-- see previous chapter), so behaviorist theories of psychological states will conflict with the NC constraint.[69] Functionalist theories of psychological states likewise analyze psychological states in terms of their causal effects (together with their causal inputs). At least in many cases, the relevant causal effects are intentional actions, or at least the potentiation of action. Even where the effects are productions or modifications of other psychological states, these might be considered mental or psychological types of intentional actions.[70] So functionalism is likewise likely to violate the NC constraint.

---

[69] Hornsby (1986 p. 104) makes more or less this same point.

[70] There are different views on the nature of mental actions. For a glimpse of some of the more recent debate: Strawson (2003) argues that few things are really intentional actions.. However, Buckareff (2005) argues that Strawson's account is problematically restrictive while Proust (2001) defends the explanatory role of mental actions and offers a definition of a mental act. In terms of what sorts of actions mental actions

Of course, most philosophers these days are representationalists. So perhaps the NC constraint just gives representational theories another advantage over dispositional and functionalist theories. That will depend, of course, on whether philosophers can analyze mental representations in a way that doesn't make essential appeal to intentional actions. Alas, there's reason to think this isn't possible.

There's reason to think that an adequate analysis of mental representations will have to appeal to and specify a distinctively representational function and there's reason to think that the relevant function will have to mention specifically intentional actions. Consider, for instance, Dretske's teleosemantics account of mental representation which appeals to causal or information ("tracking") relations together with biological functions. According to Dretske, something functions to represent iff it has the biological function of producing certain effects. Relatedly, all functions make possible malfunction, but not all malfunction is misrepresentation. And, while all functions produce effects, not all effects are relevant to representational functions.[71] This applies to biological functions as well: not just any old biologically fit effects will do if the analysis of representational function is to exclude, for instance, the cellular activity in one's gut. So just as the behaviorist needed to restrict the class of behaviors, teleosemantics must restrict the class of effects. Thus, in distinguishing effects of representations from non-representational, causal activity, the teleosemanticist will also plausibly have to appeal to intentional actions. If so, then the NC constraint will again be violated.

---

include, McCall (1987) argues that deciding is an action and Wu 2013, to some extent following William James, argues conscious mental action is cognitive attention. Finally, Hieronymi (2006) argues that mental actions-- specifically the formation of beliefs and intentions-- is importantly different from bodily action. The problem-solving account of intentional action extends to mental acts of reasoning, deciding what to believe, etc. However, I do not have a worked-out theory of mental actions.

[71] See also Lynn Rudder Baker (2010) for an argument that Dretske's account of attitudes is circular.

The Second Tension

As we have seen, according to the causal theory, intentional actions must be analyzed in terms of mental representations causing them *in the right way*. For counterexamples involving deviant causal chains show that psychological states can cause actions that are not (intuitively) intentional. But susceptibility to such counterexamples seems to be a structural feature of the causal theory. Thus, if an account of mental representation assumes that mental representations are the constitutive causes of intentional actions, it will likely be unable to completely specify the nature of representational function. So, the analysis of mental representation will also be incomplete. To see how an account of mental representation might inherit the problem of having to specify the "right way" for psychological states to cause intentional actions, it is helpful to review the theoretical progression that culminated in teleosemantics/ teleo-functionalism:

Prior to teleo-functional/teleosemantic accounts of representation, philosophers were offering purely "informational," or "tracking" accounts that analyzed the representational relation as some kind of causal relation; representations were essentially effects -- passive causal responses to stimuli. Dretske initially developed and endorsed such a view (e.g., Dretske 1988). But causal effects or passive causal responses are not in general about their cause; they do not represent or refer to or stand in any kind of interesting intentional or robustly representational relationship to their causes. In addition, nothing about a merely passive causal relation is normative. So causal relations cannot, on their own, explain the robustly representational (or intentional) relation.

So Dretske's idea was to augment the causal (informational) account: since causal relations on their own are not sufficient for mental representation (bracketing the question of whether they

are even necessary),[72] how about we add a functional component to them? Dretske added the "function to indicate" or to represent, which he spelled- out in terms of a causal response to a stimulus that itself biologically functions to cause an organism to respond to the stimulus in an appropriate or fit way.

Now, as I just explained, the class of biologically fit responses to stimuli is too inclusive. Plausibly, the class of responses that will allow the analysis to pick-out all and only what are intuitively mental representations will be identical to the class of responses that constitute intentional actions. So plausibly, Dretske would have to modify his analysis thus: the causal responses to stimuli that constitute representations are those that function to cause an organism to produce appropriate (or fit) intentional actions. But, if psychological states are causes of intentional actions, there's reason to think they will be in principle capable of causing actions that satisfy their contents in deviant ways. So, if a mental state is a state that causally responds to a stimulus and constitutively functions to cause appropriate intentional actions, there's a question of which way of causing the action counts as satisfying the function, i.e., which of the ways of causing an action is "the right way." If the action's causal effects in the deviant case are the same as in the non-deviant case, the action produced may still be practically appropriate, but the state will not fulfil its function of causing an appropriate or fit *intentional* action. So, in order to explain representational success and failure, Dretske will have to specify what is the right way to cause an action such that it is not only practically appropriate (as this may be accidental) but intentional.

---

[72] Millikan argues that the reference relation or indication is grounded in natural correspondences which may be causal, logical or mathematical, etc. see e.g., Millikan (2004, 2017).

In short, Dretske's solution lands him with almost the mirror image of his initial problem: he began with the problem of needing to supplement the analysis of passive causes to specify the right ones, namely, those that are plausibly mental representations, and now he needs to supplement his supplement to specify the right productive causal relation between the passive causal state and a practically appropriate or fit effect (action), namely, the kind that's constitutive of intentional actions.

Now, behaviorism, functionalism, and teleosemantics representationalism are not the only analyses of substantive psychological representations in the offing. But given the apparently necessary connection between intentional states and intentional actions, the fact that they all suggest that the connection may have to be explained in a way that treats intentional actions as more basic than intentional states is reason to think that the causal theory will run into trouble with the NC constraint.[73]

To sum up: Causal theories acknowledge that intentional actions have a means-end form (revealed by answers to why questions), and they want an account that grounds this form. But they treat that form as though it were something added to an action that doesn't otherwise have it. I argued that actions, proper, already have such forms and that causation by propositional attitudes probably can't ground means-end form. In other words, there's good reason to think that the causal theory can't answer the Type Question(s) for intentional action.

---

[73] Springle & Humphreys (2021) argue that intentionalist theories of representation will also have to appeal to intentional action.

3.3.2 Salvaging The Causal Theory

Perhaps we could repurpose the causal theory to answer the Token Question. Recall that, roughly, where the Type Question asks, "what makes something an intentional action?", the Token Question asks, "what makes this intentional action the intentional action it is?", that is, "how is a token intentional action determined?" If the causal theory answers the Token Question and the teleological theory answers type questions, then they are not directly at odds with each other.

I think it is unlikely that the causal theory can answer the Token Question for intentional actions-- not if the teleological theory is going to answer the Type Question. This is because how one answers the Type Question has a bearing on how one can coherently answer the Token Question.

Consider: the token causal theory would say that intentional action tokens are determined by the intentional states (propositional attitudes) that cause them; a token action is intentionally directed in the particular way it is by virtue of its etiology. Teleological theories deny that psychological representations ground the end-directed nature of intentional actions. Anything that is an intentional action of eating is intrinsically intentionally directed at a type of end. So, any token intentional action of eating is already intentionally directed and causation by intentional states isn't what makes this so. In other words, the fact that token intentional actions are intentionally directed has already been explained. What hasn't yet been explained is the *particularity* of that intentional directedness, that is, the fact that *this* particular action of eating is directed at satisfying my hunger or need for food *now*, and the fact that this particular action of eating is directed at *that* particular apple. So, the question is whether the token causal theory can explain what determines the particularity of the intentional directedness of token intentional

101

actions. If it can, then intentional states (psychological representations), e.g., a mental representation of myself as hungry and of something as an apple, are what determines this kind of particularity.

I think the causal theory cannot explain what determines the particularity of the intentional directedness of token intentional actions because psychological representations seem to be neither necessary nor sufficient to determine this kind of particularity.

First, psychological representations seem unnecessary to determine particularity. (To be clear, I'm talking about fundamental cases. Representations might be needed when we are intentionally directed at e.g., merely possible states of affairs.) Since the action token is necessarily the kind of thing that's intentionally directed, it seems all we need to add to get the particularity of intentional directedness is ordinary physical causation; the intentional action provides the intentional form, and we just need matter, or patterns of causal relations in matter, to fill-out that form. On this view, an action is *intentionally* directed at particulars because it's a token of a type of action; an action is intentionally directed *at certain particulars* (and it is the particular action that it is) because of the way ordinary physical causes determine that token.

Consider again an act of eating. Plausibly, acts of eating are intentionally directed at providing nourishment by means of an interaction with something nourishing that leads to its being digested. So, any token eating is a digestion-positioning or initiating interaction with (directed at) a target where that interaction (directed at that target) is directed at providing the agent with nourishment or satisfying their need for nourishment or hunger. Now the idea is that if one has a capacity to eat, merely causal impacts from the world could be processed so as to determine a token eating by causally connecting that capacity with a particular instance of hunger (now) and with a particular target (this apple). One might think of a capacity for an intentional action as a

kind of schematic mechanism (or mechanism schema) for a type of process that teleologically connects certain variables. It is schematic in the sense that it determines a token of that type of process when its variables are assigned values, and efficient causation is what assigns these values. If this makes sense, causation by intentional states is unnecessary.[74] Action types supply form, then all you need for a particular action is causal impacts from the world to provide the matter.

Psychological representations also seem insufficient to explain the particularity of the intentional directedness of a token action. After all, representing, e.g., an apple, doesn't make it the case that one is directed at a particular apple unless the representation has a singular content. A different kind of representation-- one that could exist without the causal relation-- wouldn't account for particularity. But the singularity (particularity) of singular contents is typically explained in terms of perceptual relations where such relations are supposed to include an important causal element. Perceptual capacities (conceptual or non-conceptual) yield perceptual states that, in the good case, refer to and are thus intentionally directed at particulars by virtue of a causal or informational link (Evans (1982), Peacocke (1998), Dretske (1969, 1981). Causal activations of such capacities are sufficient to determine token singular contents. Such contents are directly linked to the world by causal/informational relations, and hence without any intermediary representations. And the idea that's been proposed is that intentional actions themselves could work like this. An ability for a type of intentional action could be like a demonstrative concept in that all that's required to token an instance of that type of action is that the ability be causally activated. So, a particular intentional action could be intentionally directed

---

[74] If this doesn't make sense, it may help to recall the Patrick Swayze/ Ghost example from Chapter 1.

at the world just on the basis of such a causal or informational link without any need for intermediate representations.[75] [76]

One might try to argue that we need causation by psychological representations to account for the fact that intentional actions can be directed at the wrong kinds of things. For example, say I inappropriately pet the arm of the couch as if it were a cat. What explains this, the argument goes, is the fact that I misrepresented the couch arm as a cat. My petting isn't really intentionally directed at the arm of the couch, but rather a representation of a cat, after all, there isn't really any cat there. However, first, it's not obvious that this is a good explanation. *Prima facie*, the action is directed at the couch-- that is the thing I'm petting, albeit mistakenly.[77] Second, we can explain inappropriate actions like this in terms of efficient causal determination of a type of action that stems from the wrong kind of causal source (the wrong kind of target). Petting is a type of action that's appropriate for cats and not for couch arms. Nonetheless, a token act of petting may be causally determined by sensory causal impacts originating from a couch arm. Then the act of petting will be intentionally (and causally) directed at the couch arm. I mistook the couch arm for

---

[75] To be clear, special kinds of concepts (e.g. demonstrative concepts) may be essentially involved in singular content. But those concepts are directly applied to particulars-- they do not depend on further representations to secure their referents.

[76] Arguably, Dancy (2018, 2020) would say that intentional actions, *qua* conclusions of practical reason, may work in this way. (See discussion in Chapter 2).

[77] One might say that I was petting the couch "unintentionally," but this does not mean that the action was not intentional. Pettings are intentional actions. That petting was "unintentional" in the sense that it was a mistaken intention. To put the point another way: what I do intentionally may be what I do practically-knowingly; a successful intentional action may be thought of as a successful exercise of my capacity for practical knowledge. So a successful intentional action may itself be thought of as (an expression of) practical knowledge. Knowledge is an achievement, so an unsuccessful intentional action would not be a case of practical knowledge precisely because it was unsuccessful. In that sense the intentional action could not be thought of as an expression of practical knowledge and in that sense it's unintentional. But there's also a clear sense in which it's nonetheless intentional: if asked "why are you petting the couch?" you wouldn't say "I didn't know I was doing that," but rather " O! I thought it was a cat!" In other words, Anscombe's question "Why?" has application, and the knowledge criterion is at least partially satisfied. It's possible to act intentionally but stupidly or confusedly. I for one seem to do it all the time! Also, mistaken actions like petting the arm of a sofa are unlikely to occur when one is really paying attention to what one is doing, and what we are inclined to say we do fully intentionally are things to which we are attending.

a cat just in the sense that I acted on it as if it were a cat, that is, in a way that would have been appropriate if it were a cat.

So, I don't think the causal theory can be salvaged as an account answering the Token Question about intentional action because an answer to the Token Question needs to cohere with an answer to the Type Question. The causal theory doesn't answer the Type Question and the way it would answer the Token Question doesn't cohere with the way the teleological theory answers the Type Question. But perhaps the causal theory can be salvaged in a different way.

Perhaps the causal theory should be understood as answering the Type Question not in regard to intentional action in general but rather in regard to rational action specifically, where rational actions are a special species of intentional or purposeful actions. The causal theory could use Anscombe's (1963) "Knowledge Criterion" to distinguish specifically rational actions such that, in order to count as acting rationally, an agent must know what she is doing. More precisely, one's φ-ing rationally entails knowing that one is φ-ing; it is a necessary (though not sufficient) condition for rational action. If, for instance, I am asked why I'm shaking my leg and I respond that I wasn't aware of doing so, then shaking my leg wasn't a rational action. Some have argued that the knowledge criterion is too strong and should be weakened to a belief criterion (Setiya 2008). In other words, to count at doing φ rationally intentionally, one has to believe that one is φ-ing. Either way, the causal theory might be understood as claiming that *rational* actions are intentional actions that are caused (in the right way) by propositional attitudes. Among the propositional attitudes that must somehow cause rational actions are explicit beliefs (or knowledge) about what we are doing or what we are going to do, and which explain how it is that we are self-conscious of what we are doing (and thus satisfy Anscombe's knowledge criterion).

Alas, I do not think the causal theory can answer the Type Question for rational actions either. Just as intentional actions are not actions plus intentionality, so rational actions are not intentional actions plus rationality; just as intentionality is not some special feature bestowed on actions from without by a special mental state of intending, so rationality is not some special feature bestowed on intentional action from without by, e.g., a special kind of belief about what one is doing. Given what we saw above about actions being kinds with constitutive teleology, it stands to reason that rational actions are distinguished from merely intentional actions by virtue of a distinctive kind of intrinsic telos. For instance, (part of) the aim of rationality may be to act only in ways that cohere with a rational standard.

On the other hand, if rational activity is essentially a matter of reflecting on our reasons for acting as such and then acting in the way that's deemed most appropriate by that reflection, then plausibly, when rational actions are determined, an agent's representations of reasons must have done some causal work; they must have been inputs to the rational process leading to the action. This suggests that the causal theory might be right (at least roughly) about how token rational actions are determined. If so, then the causal theory might be salvaged as a theory aimed at answering the Token Question(s) for rational actions.

3.4 Towards A Different Kind of Teleological Theory

Teleological theories don't have the problems causal theories do. Still, they might not be entirely satisfying-- at least not to folks in the rich-correspondence camp (see Introductory Chapter). This is for two reasons. The first has to do with the kind of account extant teleological

theories provide of the means-end or teleological form of intentional actions. The second has to do with the lack of an account they provide of intentional states.

As we saw above, extant teleological theories of action tend to be explanationist: they analyze the means-end or teleological form of intentional actions in terms of the kinds of explanations or concepts that have application to them. But arguably, such an account does not address the seeming "mysteriousness" of intentional actions. The mysteriousness of intentional actions consists in the fact that they're intentionally directed at something that does not yet exist. The causal theory is initially attractive because it seems to suggest an answer to this. However, as we saw in §3.1, that answer doesn't work. In a sense, the causal theory tries to solve one mystery by means of another. After all, the intentional directedness of psychological representations presents just as much of a mystery as does the intentional directedness of actions. And to make matters worse, as we saw in §3.1, there's reason to think that solving the mystery about the intentional directedness of representations may require us to first solve the mystery about intentional actions.

However, folks who identify with the therapeutic view (see Introductory Chapter) are likely to reply that the idea that an explanationist account is failing to dispel some mystery is confused. For there never was any real mystery. On the other hand, one who identifies with the reductive view (see Introductory Chapter) will think there is a mystery, and that solving it will involve reducing action and representation to something natural and thus non-intentional. For reasons discussed in the Introductory Chapter, however, I'm privileging the rich correspondence view. That view agrees with the therapeutic view that intentionality is not suspicious, so it does not need to be reduced to something non-intentional. But that doesn't mean there's nothing mysterious about intentionality. One who identifies with the rich correspondence view thinks that

107

intentionality is perfectly kosher, but that doesn't mean that an explanationist account of intentional action is satisfactory.

The rich correspondence view wants to articulate rich intentional concepts in terms of which to understand intentional phenomena and how they're realized by their physical realizers. It wants to articulate such concepts both for representational intentionality and agential intentionality. And as I've now mentioned a few times, there's reason to think that an account of representational intentionality will have to ride piggyback on an account of agential intentionality. So, one reason fans of the rich correspondence view are unlikely to be satisfied with an explanationist account of intentional action is that such an account won't provide the resources from which to construct a substantive account of psychological representation. In connection with this, it's plausible that psychological intentionality is more fundamental than linguistic intentionality--that we won't be able to understand linguistic intentionality except against an independent background of psychological intentionality. So, if explanation-- a linguistic phenomenon-- shows up in an account of intentional action, and if intentional action is likely to figure in an account of psychological representations, and if, in turn, we need psychological representation to account for linguistic representation, then there's, evidently, a significant risk of going in a circle. And even putting aspirations of giving an account of representational intentionality aside, explanationist accounts will seem unsatisfying to philosophers who want to

give an account that articulates criteria to determine whether a teleological explanation in fact has application to a happening.[78,79]

As Rowland Stout (1996) notes, after introducing his own quasi-explanationist teleological account of intentional action,

> ...there is something rather indirect about using explanation in the account at all. To say that something is explainable is to say that some preconditions of its being explained are satisfied. These preconditions are likely to involve the right sort of causal relation holding. So, in the end, the account could be formulated in terms of this causal relation, making no mention of explanation. (8)

Now, Stout's "teleological theory" is unusual-- this is why I omitted it from the discussion of the teleological theory in §2. On the one hand, his theory is explanationist, as he initially analyzes intentional action by saying that "activity constitutes intentional action in virtue of being explain*able* in terms of a teleological justification of it" (7). On the other hand, he argues that teleological explanations are a kind of causal explanation.

Importantly, Stout's is not a causal theory as I've defined it because he denies that efficient causation by psychological representations is what makes something teleologically explainable. He argues that teleological explanation is a special kind of casual explanation: to give a teleological explanation of something is to explain it by pointing out that it is a means to some end. And Stout takes a broadly realist view of causal explanations. He holds that "behind correct causal explanations should be causal processes which are grounded in some real underlying nature; a

---

[78] And since there's reason to think that intentional actions are not the only phenomena appropriately explained teleologically, they want to be able to give criteria that distinguish intentional actions from other kinds of teleological phenomena and likewise intentional action-type teleological explanations from other types of teleological explanations. See also Roth (2000). According to Roth, reasons explanations may be irreducibly teleological and at the same time cite primary reasons as efficient causes at the same time. As Wilson (2016) notes, arguably, certain explanations in biology, in particular those involving homeostatic feedback, have both causal and teleological force.

[79] The knowledge criterion will not work to pick-out non-human animal actions.

correct explanation of something should describe a process that actually results in that thing." (3) So, although Stout's account of intentional action begins as explanationist, he takes significant steps towards formulating it in terms of the real underlying nature of the causal phenomena that ground teleological explanations.

Larry Wright's (1973, 1976) account of teleological explanation gives the beginning of such an account. Roughly, Wright's account has it that x happens for the sake of y iff (i) x tends to bring about y, and (ii) x occurs because it tends to bring about y. So, x is teleologically explainable in terms of y iff (i) and (ii) hold. However, this account of teleological explanation applies to biological activities rather broadly; it does not uniquely pick-out intentional actions. So, Stout has to refine the account. He argues that only "strongly" teleological explanations apply to actions, whereas "weakly" teleological explanations apply more broadly. Unfortunately, Stout's account of strongly teleological explanations is, at least by my lights, difficult to understand.

According to Stout, ϕ is an intentional action iff ϕ is the immediate result of causal process which is governed by a method of justification which is normative and embodies a means-end sensitivity (155). So, among other things, Stout needs an account of what means-sensitivity is and he needs an account of what a means to an end is. According to Stout, for some event m to be a means to some end e, two things must be the case: (i) there must be some way for e to occur which involves m, and (ii) m must be an available activity (156). Note that Stout's account of what a means to an end is appears to be purely efficient causal. Many causal processes and mechanisms satisfy his definition. Ends seem to disappear.

To the extent that ends enter the picture, it seems that teleological explanation enters with them and thus enters into the explanation of the nature of the causal process that's supposed to ground teleological explanations. For by "method of justification," Stout means a method of

deriving a description of the world (16-18). Now, Stout distinguishes between methods of theoretical and practical justification. Beliefs can be thought of as justified or not relative to a given method of theoretical justification whereas a method of practical justification is a normative method that says what ought to be done. And Stout doesn't want the method of justification to be merely incidental to the causal process (means) that it describes. Instead, Stout says, the method of justification "must govern the process" where this means that "The process must result in whatever the method recommends. If the method came up with a different result, the process would have to adapt to that, rather than vice versa." (156) Unfortunately, I am not entirely sure what this means.

Stout's account is, clearly, complex. To the extent that I understand it, I'm largely sympathetic, and I expect fans of the rich correspondence view would be too. And since I am not entirely sure what he means when he says a method of justification must govern the causal processes it describes, I'm not sure how much of my disagreement with him is substantive as opposed to terminological. In any case, if there is something wrong with his account, it's that means are just causal processes that can be understood without reference to ends, and they get connected with ends via methods of justification which seem a lot like teleological explanations that somehow causally govern them. By my lights, this looks too much like a causal process running parallel to an explanatory practice with mysterious causal strings connecting them. So, in the next section, I'll develop a thoroughly non-explanationist account of intentional action. It will focus on analyzing the means-end form of intentional actions. As I see it, you cannot say what a means is without reference to an end. So, means can't just be physical processes (nor can they just be what Stout calls "Aristotelian Processes" (50-62) insofar as the process of an apple decaying counts as such a process). And while I agree that ends do not fundamentally get involved by taking

the form of psychological representations (e.g., primary reasons), I do think they need to get involved at the same level as the physical processes that, in relation to them, constitute means. So, I'll suggest that ends are just basic normative physically realized facts-- what I'll call "needs for flourishing" or "practical problems."[80]

First, however, I must explain the second reason extant teleological theories might not be entirely satisfying. This reason has to do with satisfying the Connection Desideratum.

The reader might recall that the reason a complete account of intentional action should say something about intentional states is that it seems undeniable that if someone acts intentionally, then they had an intention, a belief, and a desire. In other words, acting intentionally does seem to entail having certain psychological states. And while acting intentionally does not entail that those states are that in virtue of which an action is intentional, intentional actions are essentially intentionally directed at reasons, being intentionally directed at a reason seems to entail being cognitively in possession of those reasons, and when we are talking about an agent's cognitive possession of reasons, we seem to be talking about her intentional states. So there seems to be a necessary connection between intentional states, reasons, and intentional actions. And, as I noted in §1, this connection seems to be causal in some sense. The causal theory attempts to explain this connection-- via the derivation thesis, i.e., by means of efficient causation-- but I've argued that that explanation doesn't work. Since teleological theories do not appeal to intentional states to account for intentional actions, it's open to them to account for intentional states in terms of intentional actions.

---

[80] I recognize not everyone will like this move, quite possibly including some of my fellow rich correspondence fans. But I really don't see what's wrong with it. In an age in which the reductive view holds a lot of sway, it's hardly surprising that plopping norms in nature in the way I've proposed will be met with skepticism. But hell, philosophers are putting consciousness or qualia in nature. So why not put things we and ordinary people care about there too (or even instead!).

There are different ways for teleological theories to do this, and they involve different kinds of revisions to the conception of intentional states operative in the causal theory.

First, a teleological theory may take an ascriptionist approach to intentional states. Such an approach makes no commitments to intentional states *qua* substantial (causal) mental entities. So, it may agree with the causal theory that intentional states-- beliefs, desires, and intentions-- are propositional attitudes, but unlike the causal theory, it holds that propositional attitudes are in no philosophically interesting sense "things in the head" that are causally involved in the production of action.

The ascriptionist approach goes some way towards meeting the connection desideratum in that it explains the conceptual connection between intentional actions and intentional states. As Stout (1996)-- taking his lead from McDowell (1978 p. 15) and Nagel (1970 p. 29-30)-- puts it,

> The fact that there is a conceptual connection between having a desire and having a reason for one's action need not be explained by the fact that one has a reason for one's action in virtue of having the desire. It may be explained by the fact that one has the desire in virtue of having a reason for one's action; that is, in virtue of the reason explaining or potentially explaining one's action. (33-34)

This shows that it is possible to give a non-causal interpretation of the psychological relationship between intentional states and intentional actions.

Indeed, since the teleological theory considers the end-directedness of actions an intrinsic, formal, or logical feature, it doesn't need psychological representatives to represent ends and casually direct the events that thereby are intentional actions at them. So, unlike the causal theory, strictly speaking, the teleological theory can make do with a merely ascriptionist conception of psychological representations. Moreover, in explaining why the causal theory doesn't provide an answer to the Token Question (§3.2) for intentional actions that's compatible with a teleological theoretic answer to the Type Question, I argued that once we have the form of intentional actions

on the table, we don't need to appeal to causation by intentional statues to account for the particularity of the directedness of token intentional actions. In the light of all this, it might seem as if there's no room for substantive intentional states in a teleological theory of intentional action. And those who go in for the therapeutic view will likely also think there isn't any need.

However, it's hasty to conclude that there isn't room for substantive intentional states in (or alongside) a teleological theory of action. Those who go in for the rich correspondence view want a substantive, non-ascriptionist account of intentional states that's compatible with the teleological theory. From the perspective of the rich correspondence view, then, what the above considerations show is the need to reevaluate how substantial psychological representations fit into (or alongside) a teleological theory of action. And this requires a significant shift away from the propositional attitude models of representational intentionality.

As we saw in §3.2, propositional attitudes, substantively conceived, do not seem to fit with the teleological theory. This is because the teleological theory denies the derivation thesis -- it denies that intentional actions are constitutively caused by psychological representations. But propositional attitudes, substantively conceived, are causes of intentional actions. As such, they are external to intentional actions. Moreover, on the causal theory, psychological representations, and thus propositional attitudes, have intrinsic intentionality, and intentional actions derive their intentionality from them. But on the teleological theory, intentional actions do not derive their intentionality from psychological representations. Indeed, rather than accounting for the intentionality of action in terms of the intentionality of psychological representations, some teleological theorists propose to ground (at least some) intentional states/psychological representations in or in relation to intentional actions (e.g., Stout 1996, Thompson 2008).

In addition, propositions stand-in-for what they represent. It's hard to see how or why psychological representations should stand-in-for anything if they are essentially understood as elements or aspects of intentional actions and directly share in their intentionality, not unless intentional actions stand-in for that at which they're intentionally directed. Now some intentional actions do something like standing-in for things. For instance, one can intentionally stand-in for someone in line or serve as a model or represent one's business or university or represent someone in the court of law. And of course, intentional actions of producing stand-ins-- pictures, models, sentences, etc. -- involve stand-ins. But these are special cases. In general, intentional actions are not acts of standing-in or producing things that stand-in. And even in these special cases, the standing-in is something one actively does, not something one *is*. Even though we do say a person is a representative, what we mean is that they are acting as one; the action is one of standing-in, it isn't itself a stand-in.

In contrast, propositions are contents that stand-in-for states of affairs. They are about their contents by virtue of standing-in-for them. They are contents that specify their satisfaction conditions; they present something as being the case and succeed if or when that is the case. Propositions can be parts of actions in that actions can be responses to them and in that actions can produce or employ them. But this just shows that propositions are external to actions--they're things that go into or come out of them. And this is consistent with the causal theory's conception of intentional states as external causes of actions. It appears, then, that a substantive account of psychological representations that will cohere with the teleological theory will need to reconceive what it is for psychological states to represent.

Just as Stout's (1996) appears to provide at least the beginnings of a teleological theory of action that will satisfy the fans of the rich correspondence view, he also provides at least the

beginnings of an alternative to the propositional attitude model of mental representation in the form of his "Teleological Behaviorist" account of beliefs and intentions. Stout defines beliefs and intentions in several ways, the first of which, he notes, provides a reductive account of belief and intention only "in the sense that there are no mental states *mentioned* on the right-hand side of the biconditional" (178). However, as Stout notes, these definitions employ concepts on the right-hand side of the biconditional, including that of a "conception of rationality," that arguably depend on mental state concepts.[81] So the account is not reductive "in any strong sense" (178). Stout considers some ways of reworking the definitions and arrives at the following "reductive behaviorist" account of belief:

> A thing believes that P if and only if there is a way of construing rationality (externalistically and as embodying working assumptions) so that it is disposed to produce that activity which would be rational on the assumption that P. (180)

But this definition mentions a way of construing rationality externalistically and as embodying working assumptions, so it is not strongly reductive either. Indeed, Stout notes that this definition is almost identical to his teleological behaviorist definition of belief:

> A mechanism is associated with the belief that P if and only if its working is governed by a method of practical justification which works on the assumption that P. (180)

The teleological behaviorist account fails to be thoroughly reductive because it mentions a method of practical justification (or a way of construing rationality). As Stout notes, what he has

---

[81] Stout seems to think that a subject cannot appear in the *definiens* of a reductive account. I think this is not obvious.

provided is an "account of mental intentional states in terms of intentional concepts that apply to language" and Stout considers this "a reduction of sorts" (180).

Now, the rich correspondence view does not seek a reduction of intentional concepts to exclusively non-intentional concepts. However, as I mentioned above, there's reason to think that psychological intentionality is more fundamental than linguistic intentionality--that we won't be able to understand linguistic intentionality except against an independent background of psychological intentionality. So, if intentional concepts that apply to language show up in an account of psychological representation and if we can't understand linguistic intentionality except against a background of psychological intentionality, then we again run the risk of going in a circle. In addition, although Stout's analysis invokes causal processes and mechanisms, it still has an ascriptionist ring to it. So, although it's an excellent start, I think Stout's teleological behaviorism is unlikely to satisfy the connection desideratum (by the lights of the rich correspondence view).

To get closer to satisfying the connection desideratum, in the next section, I'll sketch an intentional action acorn account of psychological representations. Like Stout's teleological behaviorist account, it will understand psychological intentional states essentially in terms of intentional actions. But unlike his account, it will not appeal to explanation, language, conceptions, assumptions, etc.[82] Still, it will only provide the *beginnings* of a substantive rather than ascriptionist account of psychological representation that's coherent with a teleological theory of intentional action. It will fall short of a proper analysis of psychological representation in two ways. First, it won't say anything about the nature of representation more generally. Second, it

---

[82] It will appeal to agents, or life forms, and in that sense subjects. But we can understand life-forms not in the first place as things that act intentionally but rather as things that essentially have needs for flourishing and (environment-dependent) capacities to meet them.

relies on a metaphor. My objective in introducing the intentional action acorn account is to make relatively intuitive -- by means of this metaphor-- a radically different way of thinking about the nature of psychological representations and their relationship to intentional actions. This will allow me to characterize important features of such representations that clearly distinguish them from propositional attitudes. In this way, the intentional action acorn account will provide the beginnings of an articulation of a substantive alternative to propositional attitudes. Later chapters will attempt to go the rest of the way.

3.5 A Problem-Solving Account & Intentional Action Acorns

I'll begin with an exposition of the problem-solving account of intentional action. (§5.1) After running through the basics, I'll address some worries a reader might have for the view. After that I'll turn to intentional action acorns. (§5.2)

3.5.1 Means as Application-Accountable Solutions to Practical Problems

The purpose of this section is to attempt a non-explanationist account of the nature of a means (to an end), i.e., the constitutive form of an intentional action. In other words, it's an attempt to answer the Type Question(s). We already have a name for this form: means (or means to an end). Here's how I propose to analyze it:

Def. Means-End Form: A has a means-end form iff A functions to be a token application-accountable solution to a practical problem.

The point of this definition is to say that a means isn't just a physical process that happens to be explained or described in special ways, but rather a solution, and that's why it has to be explained or described in a special way. Solutions are special because they have to be defined in terms of practical problems. As we'll see in a moment, a practical problem isn't just any kind of physical fact, but rather a special kind of fact-- a fact about a need for flourishing, or what we might call a "need fact." I consider such facts irreducible, i.e., primitive. A means, then, is a mechanism that functions to be a solution, meaning that the actualization of such a mechanism brings about a state of affairs that constitutes a solution to a practical problem. So, if an intentional action is essentially something that has a means-end form, we can define intentional action in terms of the way I've just defined that form as follows:

Def. Intentional Action: A is an intentional action iff A functions to be a token application-accountable solution to a practical problem.

To unpack the definition:

First, a **practical problem** is a type of "need for flourishing," where such needs are objective facts about life forms.[83] Different forms of life (or species of life form) have different

---

[83] Practical problems apply to whole organisms. An organism's sub-parts function to promote its flourishing by figuring as proper parts of its actions (solutions to practical problems). Digestion, for example, is part of an intentional action of eating; it is a kind of sub-action. Like some other sub-actions, it is not as sensitive to global needs as actions proper. But digesting is something we do. We control it by eating and even by eating certain things in a certain order and at a certain rate. Digestion has its own-sub-activities which take place automatically, but this may not be so different form the automatic sensory-motor adjustments that go into, e.g., picking an apple or climbing a tree. (Thanks to James Shaw for encouraging me to make this explicit).

types of needs or conditions for flourishing some of which are more and some less basic, some needs may contingently derive from other needs (e.g., the need to work to get food and shelter), and needs may be coarser or finer grained. Needs for flourishing are directed at forms of interaction between life forms and types of "targets" in the world such as berries, oceans, and stop signs, for the sake of bringing about a mode of flourishing. For example, a life-form needs the metabolic resources-- some amount of water, minerals, carbohydrates, etc.-- to maintain and energize its body. The need for nourishment is a need for a kind of interaction (eating) with certain kinds of targets (edible, nutritious substances). The need for nourishment is also a quite general type of problem. One may define more specific subtypes in terms of more specific types of targets (more specific types of sources of specific sorts of nutrients) and interactions (the types of processes of acquisition required for different types of sources, e.g., grazing, hunting, or shopping). Nutrition and protection from threats are basic types of needs for biological flourishing. Friendship may also be a basic need for flourishing for some forms of life, e.g., humans.

Next, a **solution** is a type of process of interaction between a life-form and a type of target (or type of combination of targets) that is generated by the life-form and that is a (non-accidental) mechanism for satisfying a type of need for flourishing (i.e., solving a type of practical problem). Examples include anything from eating, drinking and evading predators to making friends to producing objects such as paintings and shoes.[84]

There are different kinds of practical problems and corresponding solutions: A **local solution** solves a type of local problem; it answers: *What do I do with respect to some local*

---

[84] One might say: that's just to say that it's a means. I don't disagree. But what this definition aims to bring out is that there are mechanisms that reliably (non-accidentally) lead from A to B, but that doesn't make them a means of getting to B, but only a potential means. Such a mechanism becomes a means only by virtue of being a solution to a practical problem.

problem P? Local solutions are concerned to solve local problems, like a particular need for nourishment. Every token local solution is a candidate global solution. A **global solution** solves a type of global problem by selecting the right local solutions at the right times; it answers: *Which of some set of the local problems I'm facing should I try to solve now to promote my overall flourishing?* or: *Which of these available local solutions should I prioritize, i.e., perform (now)?* [85] Global solutions are concerned to solve global problems by selecting and ordering local solutions for performance according to what balance of local needs are most pressing with respect to an agent's overall flourishing in a context.

Finally, there are also **sub-solutions** (and sub-sub-...+ solutions), which solve sub-+problems—instrumental problems that arise in the course of solving a larger practical problem. Roughly: a (super-)solution tells you *what to do;* sub$^+$-solutions tell you *how to do it.* Sub-solutions are essentially involved in the performance or actualization of a solution; they function to solve sub-practical problems that arise in the course of performing (actualizing) an overarching solution to an overarching practical problem.

Solutions (and therefore intentional actions) are fundamentally evaluable in terms of a norm of appropriateness. A local solution is appropriate iff, were it performed to completion, it would non-accidentally solve an instance of the type of local problem for which it is a solution. A local solution is inappropriate iff it fails to "practically fit" the context in which it is applied (potentiated), i.e., if it is applied outside its type's *domain of application* which consists of instances of the types of needs and targets which define the problem-type the solution-type it

---

[85] A global problem is defined by the logico-practical relations (compatibility, incompatibility, complementarity, and sub means-end or realization relations) that obtain between one's available local solutions in a context. An appropriate solution to a global problem consists in prioritizing and selecting for performance the action(s) that will most promote one's overall flourishing in a context.

instantiates solves. A global solution is inappropriate iff it consists in the selection of a local solution that would not, were it performed to completion in the context, non-accidentally promote the subject's *overall* flourishing (and thus constitute an appropriate global solution in the context).

Importantly, the (in)appropriateness of a solution (intentional action) is not to be confused with the success or failure of its performance or execution (though the latter may redound on sub-solutions). After all, an appropriate action might be formed but never be performed, either at all or to completion (it may be interrupted, abandoned, conditions may not be favorable). And, of course, inappropriate actions can be performed perfectly to completion, but that doesn't make them appropriate.[86]

Finally, an intentional action is not merely a solution, but an application-accountable solution. Application- accountable solutions are sensitive to the appropriateness of their instances; their application algorithms (which will be described in more detail in Chapter 4) are disposed to be modified in response to whether or not the solution's tokens practically fit the situations to which they are applied, i.e., whether or not they manifest appropriate solutions.

Now, I claimed that solutions, and thus intentional actions on this account, are fundamentally evaluable in terms of a norm of local appropriateness. But intentional actions are evaluable in terms of different sorts of standards and norms, so one might wonder what reason there is to privilege local appropriateness. To see why, let's begin by distinguishing between four standards or norms in terms of which intentional actions can be evaluated.

---

[86] As we'll see below, appropriateness should also be distinguished from the standard of *optimality*; an appropriate solution need not be an optimal solution.

First, an action can be (in)appropriate with respect to a local end (*local appropriateness*). Second, an action can be (in)appropriate with respect to the end of overall flourishing in a situation (*global appropriateness*). Third, an intentional action can be evaluated in terms of whether or the extent to which it is successfully performed (*executional success*). Notice that an appropriate action might be initiated but never performed, either at all or to completion (they may be interrupted, abandoned, conditions may not be favorable), and inappropriate actions can be performed perfectly to completion, but that doesn't make them appropriate. So, appropriateness is a distinct standard from the standard of executional success. Finally, an intentional action may be evaluated as more or less *optimal*. For instance, eating some bread will suffice to satisfy my hunger, but eating some vegetables and hummus would too and would keep me full for longer and has a better nutritional profile. So, both actions are appropriate, but the latter is more optimal. Local appropriateness is an all or nothing, non-comparative notion,[87] while optimality is a graded and essentially comparative notion.

On the problem-solving account, the core kind of success for an intentional action is local appropriateness. This is because 1) global appropriateness entails local appropriateness (but the entailment doesn't work in the other direction), 2) a perfectly executed inappropriate action is practically worthless, and 3) optimality is a measure for comparing different appropriate actions. In other words, an appropriate action is an action that is the right kind of means for a situation. But

---

[87] An overarching action simply is an appropriate type of action even if some of its sub-acts are inappropriate. Global appropriateness is semi-comparative but not in the same sense as in optimality. What's globally appropriate is a matter of which of the possible actions at your disposal will most promote your flourishing overall. Thus we evaluate an action as globally appropriate only after it has been selected among others where that selection process involves comparing it to others along a graded scale (the scale of overall possible flourishing). However, the action is evaluable for global appropriateness only once it has been selected, and at that point whether it is globally appropriate is simply either true or false.

it may not be the uniquely right kind of means. Optimality introduces criteria for selecting among different means to the same end, and hence different appropriate actions.

One might object that the problem-solving analysis of intentional action counts too many things as intentional actions. After all, many things look like local solutions that intuitively are not intentional actions, such as reflexes. But an intentional action functions to be a local solution, which means it must also function to be a candidate global solution. While reflexes and the like solve practical problems, they are *not* candidate global solutions: whether they occur is not sensitive to their (in)compatibilities with and other logico-practical relations to other needs for flourishing. In addition, they are *not* application-accountable solutions. Take, for example, sweating as a reflexive response to stress. Sweating is something we do because it functions to contribute to solving our practical problems. But sweating is not something we intentionally do. This is because we cannot "control it."

On the proposed view, intentional actions are "controlled" insofar as they are sensitive to an agent's global needs and insofar as they are sensitive to their context of application. One way in which the latter kind of sensitivity manifests is via the particular manner a token solution takes-- a kind of "means-end" sensitivity. For instance, a token eating solution will be distinguished from other tokens of that type in part by virtue of the particularities of its target: eating an apple and eating a soup involve different sorts of movements. In other words, different token eatings will involve different targets and correspondingly different manners. A second way in which application accountability manifests is in the disposition of a type of solution to be applied. An EAT solution might be disposed to be applied to anything three-dimensional and apple shaped. But after being applied to a wax apple, i.e., inappropriately-- outside of its domain of application--the solution's dispositions should change such that it is less likely to be applied to a wax apple in

the future. This is a kind of sensitivity to local success and failure. A third way accountability manifests is in the way the solution type is disposed to interact with other solution types. For instance, in some situations, an eating solution solves the most pressing practical problem (e.g., when one badly needs nourishment and is unaware of any present danger), but in others, a fleeing solution solves the most pressing practical problem (e.g., when one is aware of present danger). If one's eating solution is to serve one's global flourishing, it should not be prioritized in the case of present danger. This is a kind of holistic logico-practical sensitivity. Finally, application-accountability also manifests when an action performance is underway via online (dynamic) modifications of sub-+ solutions.

Sweating is a kind of solution that is insensitive to an agent's global needs and that does not manifest the kind of sensitivity required for application accountability. Therefore, sweating is not an application-accountable type of local solution and hence not an intentional action. On the other hand, sweating is a solution, and so to that extent it might be counted as a means. But intentional actions are a means in a richer sense because, *qua* application accountable solutions, they are actively accountable to their ends, both local and global.

A different objection one might have concerns the fact that the proposed account defines intentional actions in terms of solutions and solutions in terms of needs for flourishing. The worry is that this means it won't be able to account for actions that go against one's flourishing (e.g., committing suicide, taking drugs)? In response, intentional actions that go against our flourishing are inappropriate. This results from a practical failure of fit[88]-- the wrong kind of solution is applied to solve a practical problem. The problem-solving account can also distinguish between "natural"

---

[88] The notion of practical fit(ness) is developed in more detail in Chapter 3.

and "artefactual" solutions. Natural solutions are mechanisms that non-accidentally solve practical problems that an agent employs because they are such mechanisms. The explanation of actions that go against one's flourishing is an explanation in terms of natural solutions. Artefactual solutions introduce a further way in which actions can go against one's flourishing. Artefactual solutions are solutions that are intentionally devised by agents to be application-accountable solutions. So, they function to be such solutions. But not everything that's made to be a solution actually is a solution. Some artefactual solutions are just should-be-solutions: they are not actually mechanisms that solve the practical problems they're designed to solve. They are still intentional actions, but they are fundamentally defective intentional actions. Such artefactual solutions may systematically go against one's flourishing. Since they are application-accountable, it is possible for agents to notice their defectiveness and in light of this, drop or redesign them.

Another worry one might have for the problem-solving account is that it fails to account for rational actions, i.e., distinctively human actions. I don't disagree. The problem-solving analysis is an analysis of the genus Intentional Action, and, as I mentioned in the introductory chapter, I consider rational action and agency a special species of intentional action and agency. But to get a complete analysis of this special species of intentional action, we need to add something. Rational actions are not just application-accountable solutions. They are, rather, application-and-agent-accountable solutions. In other words, to account for rational actions, we need to add a further dimension of accountability, namely, agent-accountability. My aim in this dissertation is not to provide an account of rational action. Accordingly, I don't have an account of what agent-accountability comes to. Indeed, there might be different aspects of agent-accountability each of which will require its own account. But roughly, an intentional action is agent-accountable iff agents constitute part of its application algorithm. In other words, if an

agent's intentional actions can be directed at influencing the well-functioning of a solution. (In Chapter 4 this will be cashed-out in terms of agents intentionally intervening on their own dispositions, skills, and habits.)

It may not be possible to define rational actions without mentioning mental representations as such, or at least without mentioning capacities for mental representations as such. But that's no matter-- the problem-solving account of intentional action is aimed at the wider genus. Its job is not to individuate the species. And the problem-solving account defines intentional actions without mentioning representations (as such). This means we can define psychological representations in terms of intentional actions. This, as noted, has been the strategy of other advocates of the teleological theory. But I want to go beyond noting that we can ground intentional states (psychological representations) in intentional actions. I want to articulate a substantive account of psychological representations on which they have properties that cohere with the teleological theory. As noted in the previous section, propositional attitudes, conceived of substantively, do not seem to have the right properties. So, in the next section, I'll sketch an account of psychological representations that construes them as "intentional action acorns" and show how these differ from propositional attitudes.

3.5.2 Intentional Action Acorns

Suppose that intentional actions, like living organisms, are the sorts of things that develop.[89] Specifically, let us suppose that intentional actions are like acorns. In other words, they begin as seeds the "formal cause" of which is their structuring potential. For the acorn, this is the structuring potential to become an oak tree; an acorn contains the oak tree's DNA "instructions." For the intentional action acorn, this may be the structuring potential to become a completed action, i.e., solution to a practical problem. An intentional action acorn contains DNA-like instructions for an intentional action. Just as the acorn is what will, if enabled, become the oak tree, so the intentional action acorn is what will, if enabled, become the completed intentional action. The oak tree for which an acorn is potential is not yet visible, but that doesn't mean it isn't there. It isn't yet there in the sense that the acorn's potential hasn't yet been actualized, but the potential is as actual as the acorn. Might we then say the same of intentional actions? Might our craniums be a sort of intentional-action-acorn-shell so that the seeds of action lie precisely where we're inclined to "locate" our psychological states?

I propose that ascriptions of psychological representations may refer to the "acorns" of intentional actions.[90,91] Beliefs and desires are to be understood as aspects and/or partial formations

---

[89] McDowell (2011) develops a somewhat similar idea but restricted to accounting for intentions where he takes the intentions in action (as opposed to prior intentions) as basic: " An intention for the future stands to the acting one engages in when one starts to execute it, or, equivalently, to the intention in action that it becomes at that point, as a caterpillar stands to the butterfly it becomes in metamorphosis." (16)

[90] A process of this short is activated by facts (states of affairs) and is itself that through which an agent is intentionally related to the activating objective reasons. An initiating event-- the first phase of an intentional action process-- is state-like in the following three ways. First, it seems self-contained insofar as it is the "seed" or "embryo" of the entire process; it contains the DNA-like instructions for and thus determines the shape or form that constitutes a potentiated intentional action. Second, when it occurs, an agent enters what we might consider the state of being in a particular form of intentional contact with the activating facts. Third, an initiating phase is often distinct in form from subsequent phases (we may compare the intention to walk with the final act in a process of walking).

[91] I'll develop these ideas in the next chapter via the notion of a PEA-Function.

of intentions, and in forming an intention, an agent potentiates a token solution, i.e., intentional action. In other words, in forming the intention to walk, I've initiated a walking-- I have given life, so to speak, in the form of a thought or idea-- the intentional action acorn--- to a particular walking.[92] Intentional action acorns are the necessary initiating phases in the developmental progression of an intentional action process where each stage in such a progression contains potential-- think DNA-- for the full action.[93]

Intentional action acorns will plausibly explain the observation that agents can mentally represent objective reasons without manifesting any corresponding outward actions. After all, acorns necessarily precede oak trees: you can't have an oak tree without an acorn, and many acorns (alas, the vast majority) never actualize their potential to become oak trees. What's more, the development of any particular oak tree out of an acorn depends on the circumstances of the individual -- some acorns will take longer than others to germinate, some oak tree saplings will spend many decades in a juvenile phase while others will reach adulthood more quickly. In addition, just as multiple acorns may compete with each other for the resources required to grow, multiple intentional-action-acorns may compete with, or otherwise impact, each other's probabilities of actualization, extent of actualization, as well as certain details of actualization. For instance, just as an oak that grows alongside other trees will grow tall and develop a large canopy

---

[92] The reason mental states like willings and intentions seem to be separate from and to lie behind non-mental activities and to be classified as some special sorts of "mental activities", e.g. "thinking activities," might come down to this: to give life t o a thinking activity is to give life to the thought that is the first act in that action- process. To give life to a non-thinking activity is also to give life to the thought that is the first act in an action- process. But where the activity is thinking, the process begins and ends with acts of th e same form: an act of pure thinking, i.e. thinking to oneself, without any outward activity, is (where complex/ divisible) made up of smaller acts of pure thinkin g. In other words, it's thoughts all the way down. In contrast, where the activity is something other than pure thinking, e.g. walking, the first act is not of the same form as the final act: an act of walking begins with an act of thinking, but ends, as it were, with one's feet on the ground.

[93] Since the acorn contains the DNA of the action, and is not external to the action-- as an external cause-- but internal in the sense of being a phase in the self-same process, there's a real sense in which it persists through the action.

so that it can get enough sunlight, whereas an oak that grows in a field by itself may not get quite as tall and may balance its canopy with more branches along its trunk, so an intentional action of preparing a meal will take more or less time, include a larger or smaller number of steps, etc. depending on the other intentional action acorns, i.e. other activities, with which one is simultaneously occupied. Intentional action acorns may also account for the source of the idea that psychological representations are essentially mental *states* with *prior* intentionality: qua intentional action acorn, a psychological representation is the first phase of any intentional action process and seems self-contained insofar as it's the "acorn" of the entire process; it contains the DNA-like instructions for, and in this sense determines, an intentional action.

If psychological representations are intentional action acorns, then psychological representations are not propositional attitudes.

First, the contents of psychological representations *qua* intentional action acorns are potentiated intentional actions which are themselves intentionally directed at objects (facts). They have what I'll call *de actu* contents. Such contents present their objects under what I call "generative modes of presentation." In being in such representational states, agents are not directed at propositional contents; they are not directed at objects *qua* the contents of propositions. Rather, agents are directed at the objects at which their intentional actions are directed under the mode of presentation of those actions, i.e., under generative modes of presentation. (I'll give a fuller account of *de actu* contents and generative modes of presentation in Chapter 4.)

Second, where on the propositional attitude model, psychological representations are separate from--causally and thus externally related-- intentional actions, on the intentional action acorn model, psychological representations are phases in the means-end (application-accountable solution) processes that constitute intentional actions. A psychological representation is related to

130

the performance of an intentional action as an acorn is related to the oak tree into which it develops, if it does indeed develop.

Finally, third, on the propositional attitude model, psychological representations have constitutive alethic success conditions. In contrast, on the intentional action acorn model, psychological representations don't stand-in-for anything, and therefore don't have constitutive alethic success conditions. Instead, psychological representations have constitutive appropriateness conditions: they succeed when they constitute/ are constituted by the potentiation of appropriate intentional actions (solutions to practical problems). This last point bears elaboration.

Since intentional action acorns are constitutively potentiations of intentional actions, the success of an intentional action acorn is essentially bound up with the success of the action for which it is a potential, just as the success of an acorn is essentially bound up with the success of an oak tree. And the idea that an acorn succeeds only if it fully develops into an oak tree seems wrong. The vast majority of acorns never mature into full-blown oak trees, but that doesn't seem to entail that the vast majority of acorns are defective. An acorn may fail to mature because of a lack of nutrients in the soil in which it finds itself, or a drought, or an animal might eat it, or it might rot. Whereas an acorn would be defective if none of these things happened to it and it failed to grow into a tree, or if, due to some genetic mutation, it grew into a tree that was unable to photosynthesize or take water through its roots.[94] Analogously, an intentional-action-acorn is defective not if the intentional action for which it's a potential merely fails to actualize, but rather

---

[94] The topic of genetic difference and defect is important and delicate. Plenty of genetic differences are *not* defects. Indeed, diversity is probably essential to animal and perhaps especially human-animal flourishing. I believe we should reject the view that differences that are often labeled "disabilities" are defects.

if, were it to actualize, it would fail to achieve its end, where a case of this would be its failing to mature even under ideal conditions.[95]

Now, I take it that the success conditions of intentional action acorns should correspond to what is considered a defect—a problem that is internal to the intentional action acorn, rather than external or accidental to it. In the case of the oak acorn, this is the difference between failing to ever find itself in suitable soil (an external problem) on the one hand, and failing to be appropriately sensitive to suitable soil by, e.g., failing to put out a root (an internal problem), on the other. In light of this, how should we understand the success conditions of mental representations understood as intentional action acorns? Since intentional action acorns are potentiations of intentional actions, their success conditions are appropriateness conditions.[96]

We can also see why intentional action acorns have appropriateness conditions by combining the intentional action acorn account with the problem-solving account. According to the problem-solving account, the roots of intentional actions are practical problems, i.e., needs for flourishing. Needs for flourishing are not themselves desires, i.e., states of organisms, but rather broadly biological facts *about* individual life forms. Needs are, nonetheless, intentional facts in that they are intrinsically directed at what they are needs for and in this sense, they have satisfaction conditions. Those satisfaction conditions are appropriate solutions; an appropriate solution is something that will, if actualized, non-accidentally satisfy the relevant need. Of course, the need will only actually be satisfied if an appropriate solution is actualized. Thus, the conditions under which an appropriate solution is actualized constitute a need's satisfaction conditions. But again,

---

[95] An important dis-analogy with acorns is that the intentions persist while the action is still in process; that for which the intention is a potential can evolve in the process of the actualization of the potential.

[96] Thanks to John McDowell for encouraging me to be clear about this.

needs themselves are not psychological representations. They're not actions, either. They're just their intentional roots.

On the intentional action acorn view, psychological representations are the acorns of (potentiations of) token solutions. The acorn's job isn't to already be a tree. Likewise, the intentional action acorn's job isn't to already be an actualized solution. Rather, the acorn is a "good" acorn if it is the potentiation of an (healthy) oak tree. Likewise, the intentional action acorn is a "good" intentional action acorn if it is the potentiation of an *appropriate* solution (action). Whether the solution for which the intentional action acorn is the potentiation is actualized, that is, performed, depends on further factors beyond the control of the acorn. It is the performance phase of the action that's responsible for its actualization (and only if the action has been selected to be performed). So, it is the performance phase of the action that corresponds to a need's satisfaction conditions.[97]

The intentional action acorn, in contrast, determines whether those satisfaction conditions are the right ones because it determines the action to be performed. If the action is inappropriate, the satisfaction conditions of the performance don't line-up with the satisfaction conditions of the need. Only when an acorn is the potentiation of an *appropriate* solution will those satisfaction conditions line-up. So, an intentional action acorn, or psychological representation, needs to be appropriate if it is to determine the right kind of satisfaction condition for the performance (actualization) of the action of which it is the potentiation.

---

[97] There's a sense in which the acorn phase and the performance phase are not mutually exclusive: they can and typically do temporally overlap. The acorn phase is the first phase in the action in that it is the potentiation of that action, so it precedes performance. But intentional action acorns may also persist through the course of the action's execution. When this happens, the acorn itself dynamically evolves as the performance unfolds. These ideas will be discussed in more concrete terms in Chapter 4.

Intentional actions have different kinds of success conditions-- local and global appropriateness conditions as well as satisfaction (execution or performance) conditions. Intentional action acorns, and thus (fundamental) psychological representations, only share in the action's local appropriateness conditions. Whether an acorn's local appropriateness conditions are met determines whether the intentional action it determines has the *right kind* of satisfaction conditions, i.e., whether the satisfaction (executional success) conditions for the performance of the solution match the satisfaction conditions of the relevant local need. And, finally, whether an acorn that is selected to be performed meets the situational global appropriateness conditions determines whether the satisfaction conditions of the action performance it determines will be of the right kind, i.e., whether they will match with the satisfaction conditions of situational global need.

The intentional action acorn model provides the beginnings of a substantive alternative to the propositional attitude model; it begins to articulate a substantive account of psychological representations that's consistent with the teleological theory of action. The model is only the beginning of an account because intentional action acorns are yet a metaphor. In Chapter 4, I'll demonstrate that this metaphor can be non-metaphorically cashed out. So, by the end of the dissertation, we'll hopefully have satisfied the connection desideratum in a way acceptable to fans of the rich-correspondence view. For now, I think we've made a good start.

3.6 Conclusion

I have argued for the problem-solving account of intentional action first by way of arguing for the type of theory of which it is a version, namely, the teleological theory. I argued for the teleological theory insofar as I argued against the causal theory. I then motivated the problem-solving account of intentional action, together with the intentional action acorn account of psychological representation, in light of two ways in which other teleological accounts may strike some philosophers as unsatisfying. The problem-solving account of intentional action together with the intentional action acorn account of psychological representations (beliefs, desires, and intentions) answers the Type and Token Questions as follows:

Answers to Type Question/ Sub-Questions

1. A process that functions to be an application accountable solution to a practical problem.

2. The intentional directedness of intentional actions is grounded in their constitutive (intrinsic) means-end form, and this form is analyzed in terms of the notion of an application-accountable solution to a practical problem.[98]

3. Practical problems are defined in terms of need-for-flourishing facts and environmental facts (normative/ objective reasons);[99] a solution is intrinsically teleologically directed at the types of facts that define the type of practical problem it solves. Need-for-flourishing facts are objective/ normative, but they are also facts *about us*, and in this way *our* reasons. At the same time, the potentiation of an intentional action (i.e., an intentional action acorn) constitutes the way in which a subject grasps need and target facts.[100,101] So the potentiations

---

[98] Not necessarily instrumental means; might be a means that simply realizes a type of end.

[99] These are objective reasons in the sense that they are simply facts, not representations of facts. They are of course subject-dependent facts since they are facts about subjects' needs.

[100] In Chapter 4, I'll explain that an ability for a type of solution is a kind of embodied practical concept or standing intention and constitutes our capacity to be in cognitive possession of objective reasons for action, i.e. they constitute our capacity to have subjective reasons for action.

[101] Rational actions are distinctive in part because subjects have their reasons in a special way-- they are able to conceive of their reasons as such. This idea will be discussed in Chapter 3 under the heading of " inter-subjective reasons."

of intentional actions themselves are subjective or motivating reasons -- they are a subject's grasp of facts as favoring certain actions.

Answer to Token Question:

1. Token solutions are intentionally directed by virtue of being tokens of the type of solution they are, but what in particular they are intentionally directed at depends on eliciting material causal facts and events. Intentional action acorns are potentiations of intentional actions that are directed at facts and constitute a subject's fundamental practical grasp of those facts. If an action/ acorn is appropriate, it will practically fit the facts at which it's directed. So, in the good case, an intentional action is directed at objective or normative reasons for action understood as facts and its potentiation constitutes a subjective or motivating reason, i.e. a subject's grasp of an objective reason (fact).

4.0 Chapter 3: Shifting the Paradigm: The PEA Analysis of Subjective Reasons &

Representations

Philosophers are concerned to understand different types of representations—linguistic representations, scientific representations, artistic representations, fictional representations, iconic representations, and a variety of mental representations. The exact relationship between these types isn't always clear: perhaps scientific and artistic representations are individuated by pragmatic features, while the distinction between iconic and linguistic representations is more a matter of format. This somewhat complicates the first claim I want to make, which is that philosophers are concerned to understand a variety of *species* of the *genus* representation, for one might wonder what sort of criteria—pragmatic, formatting, or something else-- are relevant to individuating the species. But we needn't settle this issue here, for my first claim is simply that philosophers seem to acknowledge a genus, *Representation*, whose instances may be categorized, perhaps in different equally legitimate ways, under theoretically useful species headings. In other words, I take it that philosophers are often concerned to give accounts of the representational species of interest to them which typically involves identifying the properties that individuate that species.

Now, any time a philosopher is faced with the task of accounting for a *species* of some phenomenon, she faces a challenge. For an account of a species of X must illuminate the way in which that species realizes the essential features of X, where these features may be realized rather differently in other species. On the other hand, we seem to infer our theories of the essential features of the genus X precisely by means of considering the nature of its instances. One way

philosophers try to meet this challenge is by analytically extracting the core features of paradigm cases of the phenomenon of interest and positing these as the essential features of the genus. [102]

The problem is that our recognitional abilities are fallible. Of course, we're fallible at just about everything we do. So, what's the big deal? Typically, when we err, it isn't long before we find out that we have erred—the world has a way of letting us know. The reason our fallibility is worrisome in the case at hand is that philosophical errors have a way of disguising themselves. In particular, if we pick the wrong paradigm—say, an unrepresentative species of representation-- we're liable to get confused about the genus. And if our explicit knowledge of the genus is confused, this may further interfere with the implicit knowledge on which our recognitional capacities depend. In other words, once we're confused about the genus, we may fail to properly recognize some of its species whose properties might have alerted us to our error. Where our paradigm obscures or fails to illuminate other species of the genus, those species are liable to strike us as relatively "mysterious," "strange," or "obscure" as compared to our paradigm, and we may consider this all the more reason to stick with our paradigm rather than investigate what these other species might have to tell us about the nature of the genus. And, of course, if we misunderstand the genus, then we're liable to misunderstand the very species we took as our paradigm. In other words, confusion in this context is iterative.

---

[102] Here too they face a challenge, for this strategy presupposes that we're already in a position to recognize the paradigm cases of the relevant phenomena and, at least with some thought-experimental prompting, their core features. One might find this presupposition problematic because it seems somewhat reminiscent of Plato's doctrine of anamnesis or recollection. For it may seem that if we are already in a position to recognize paradigm cases and their core features, we must already have knowledge of the genus. However, this is not the issue I wish to press. It may be that we already draw on a kind of knowledge of the genus, but the kind of knowledge we draw on is different in kind from the kind philosophical activity seeks to produce. One might say that the kind of knowledge we draw on is "implicit" and the task of the philosopher is to make that knowledge explicit. In any case, this is not my concern.

The example I have just given isn't intended as a mere example—I believe that this is what has actually happened: analytic philosophers have selected the wrong paradigm for theorizing about representational phenomena and this has led to the sort of snowballing confusion I just described. The problematic paradigm is what Chapter 1 called "The Descriptive Thesis." Those who endorse this thesis treat as paradigmatic of all representations a feature that belongs only to a species, namely, descriptive success conditions, i.e., truth, veridicality, accuracy, or satisfaction conditions. This paradigm has obscured philosophical understanding of all species of representation—including its own species—but perhaps none more saliently than mental representations.

When it comes to most species of representation, few philosophers seriously question whether their instances actually exist. In contrast, a significant number of philosophers are skeptical that mental representations are really things in the world rather than merely useful fictions.[103] This is so despite the fact that mental representations appear to play an ineliminable role in action explanations, i.e. explanations that appeal to "originally intentional" mental states like belief, desire, intention, perception, memory, and knowledge.[104] This is because such states have that peculiar mental property of "intentionality" or "directedness": a mind is a conscious, first-personal point of view on or grasp of facts where this grasp directly informs those actions caused by an agent we call intentional.[105] An agent's subjective reasons-- her beliefs, desires, intentions, perceptions, etc.-- correspond to aspects of her point of view, i.e. the way in which she's directed at and understands facts such that she's positioned to produce certain actions in

---

[103] See e.g. Dennett (1994, Chapter 4). And some even doubt their usefulness. See e.g., Hutto & Myin (2013, 2017).

[104] For some of the early writing on the distinction between original and derived intentionality, see Haugeland (1981), Searle (1980, 1983, 1992), Fodor (1987). See Dennett (1987) for dissent.

[105] I will be bracketing the topic of consciousness in this dissertation.

response to them. Intentionality is also a property that's essentially exhibited by "derived" representations, e.g., meaningful signs, symbols, and sentences, as these are thought to be essentially directed at or "about" the things they represent (their objects or referents).

In light of the descriptive thesis, the standard strategy has been to treat subjective reasons as a special species of representation where it's assumed that the essential intentional or semantic feature of any representation lies in its truth, veridicality, accuracy, or satisfaction conditions.

> **The Standard Strategy:** To assume the descriptive thesis and try to account for subjective reasons as a special (fundamental or original) species of representation.

So far, the standard strategy hasn't worked: philosophers have failed to adequately illuminate the nature of mental representations. And this failure contributes to skepticism about the reality of mental representations.[106]

So where does the standard strategy go wrong and what might a better strategy look like? The aim of this chapter is to suggest an answer.

I will argue that the standard strategy gets matters precisely backwards: The genus *Subjective Reason* is not a species of the genus *Representation*; rather *Subjective Reason* and *Representation* are one and the same genus. But how could philosophers get so turned around? I'll argue that the descriptive thesis is largely to blame. Starting with an analysis of subjective reasons reveals that descriptive success conditions are constitutive features of a derivative species of representation but not of the fundamental species of representation.

---

[106] See e.g., Hutto & Myin (2013, 2017).

To start out, in §1 I'll argue that, like subjective reasons, representations are essentially subjective or "perspectival." That is, anything that is a representation is a representation for someone (or some group of subjects); a representation is an interpretation of or perspective on some subject-matter or facts, and interpretations or perspectives belong to subjects.[107] So the first-person point of view and the third-person points of view are both points of view; the "view from nowhere" is really a view from everywhere, or at least, a view available from different individual perspectival locations. Therefore, the subjective nature of subjective reasons does not bar derived representations that are "non-perspectival" in the sense of being third-personal from being subjective and hence subjective reasons. Rather, I'll suggest that first- and third-personal representational points of view correspond to two species of subjective reasons for action: roughly, first-personal reasons and third-personal reasons. I'll then suggest that the propositional attitude model of subjective reasons is problematic because it projects the properties of third-personal reasons onto first-personal reasons.

---

[107] There's a debate in the philosophy of language (growing out of Perry, Lewis, by way of Castañeda and Frege) over whether we need to introduce *de se* senses for a semantics of certain kinds of indexicals and mental state attributions. *De se* contents are introduced to account for the intensional difference between the first-personal way a subject is related to a propositional about herself and the third-personal way a subject is related to a proposition about herself. For example, when Jane thinks "I am thirsty" her thought is about herself and she thinks of herself in the first-person. In contrast, when Jane sees a video of herself inhaling a glass of water not recognizing that she is looking at herself and thinks the thought "that person is thirsty" she is thinking a thought about herself-- a thought with the same referent as when she thinks " I am thirsty" but in a third-personal way. When Jane thinks "that person is thirsty," she has a thought about herself anyone else could have. In contrast, when Jane thinks "I am thirsty," she has a thought about herself no one else could have-- if Greg thinks "I'm thirsty" his thought about Greg (himself), not Jane. I am not interested in first-personal thoughts in quite this sense. The debate about *de se* contents concerns ways in which subjects are related to *propositions* that refer to themselves and different ways in which propositions can refer to their thinkers/ speakers. In other words, that debate presupposes that the objects of thought are propositions and is framed in terms of how propositions refer and how subjects relate to them. I am interested in what is sometimes called *objectual thought*, where the objects at which thought is directed are sometimes but not always propositions (see Chapter 1). And I am interested in the most basic, fundamental kind of objectual thought-- not necessarily true/ false thoughts. If one restricts the term "thought" to things that can be true or false, then my topic isn't really thoughts but rather proto-thoughts or the representations that constitute our fundamental cognitive grasp of things. I'm interested in a phenomenon that is more basic than *de se* thought. It is possible that my findings will shed light on the nature of *de se* thought; I haven't spent enough time thinking about this matter to form any opinions.

I'll refine these claims in §2 by positing two species of subjective reasons. First, there are intra-subjective reasons --the sorts of subjective reasons *out of which* agents act or which agents *enact*-- that correspond to original representations. Both human and non-human animals can act for such reasons. For instance, when Ruch the cat sees a treat and eats it, she takes the treat as something to be eaten. Her taking the treat as something to be eaten is the intra-subjective reason out of which she acts.

Second, there are inter-subjective reasons-- the sorts of subjective reasons *on the basis of which* agents act-- i.e., representations that can be cited as reasons for acting in some way.[108] These are a kind of derived representation. As far as we know, only humans act for such reasons. For instance, Ruth sees a piece of cake and takes it that it is a piece of cake and that it is edible and hence something that can be eaten by someone like her. In this way, she takes it that the cake is a kind of fact that can serve as a kind of inter-subjective reason for acting. Inter-subjective reasons are "objective" in the sense that a) they are public or explicit interpretations of facts and b) they are interpretations of facts as reasons for *someone* or *anyone* (perhaps in a certain position or context) to act in some way. They are not, however, objective reasons in the sense I'll use the term.

---

[108] Some cases fall in-between these two species of subjective reasons as I'm describing them. Or, rather, I'm focusing on the inter-subjective that are distinctive of rational animals. I think only rational animals can really act on the basis of reasons as such, but there's a sense in which non-rational animals can act in response to a kind of inter-subjective reasons. For instance, I take it that non-human animals like vervet monkeys can have intra-subjective reasons. Vervet monkeys can also produce calls to solve proximity problems. For instance, they can produce certain calls to alert other monkeys in their group to the presence of different kinds of predators. When they do this, they produce surrogative representations-- their call goes proxy for the predator. Monkeys that hear the calls can then indirectly respond to the predator by responding to the call. Such SRs are objectified interpretations-- the producer interprets a situation as one in which there is a predator to be avoided and therefore as a reason to produce a warning call, and the call functions to convey this interpretation in that it functions to get the other monkeys to respond to the situation as if there were a predator to be avoided. However, there's a sense in which the monkeys simply interpret the call as if they were seeing the predator themselves and responding to it. When they understand the call, they have an intra-subjective reason to respond to it in a way that constitutes indirectly responding to the predator (if there is one). But they do not respond to the call as itself a kind of reason for action. So it is not a reason *on the basis of which* the monkeys act. As far as I know, only rational creatures can respond to SRs as intra-subjective reasons proper. I will get at more or less this same idea in Chapter 5 when I discuss what distinguishes propositional attitudes proper from other kinds of embodied instructive attitudes that involve surrogative representations.

Objective reasons are simply facts that favor certain kinds of responses for certain kinds of agents. They are not interpretations of facts. For instance, the presence of a treat is an objective reason for Ruch to eat it if Ruch is hungry.[109]

I'll suggest that the propositional attitude model may be understood as a strategy for accounting for the common element that unifies these two species of subjective reasons. This strategy identifies representation as the common element and understands representation on the model of derived representation. So, I'll suggest that the standard strategy tries to understand intra-subjective reasons on the model of inter-subjective reasons. This, then, is another way of diagnosing the source of the problems with the propositional attitude model discussed in §1.

In §3 I 'll explore a "neglected" alternative strategy: rather than starting with an account of representation and explaining subjective reasons in terms of that account, start with an account of subjective reasons and explain representations in terms of that account instead. In other words, start with an account of intra-subjective reasons and use that to understand inter-subjective reasons. To this end, I'll sketch a theory of objective reasons and subjective reasons for action. Roughly, objective reasons are facts that favor a certain kind of action for a certain kind of agent (perhaps in a certain kind of context), and subjective reasons are the ways agents take or interpret facts (and hence objective reasons) where an agent may be mistaken or may misinterpret the relevant facts. (§3.1) I'll also consider how to make sense of three important properties of representations for intra-subjective reasons, namely: "aboutness" (or "directedness"), "fitness" (success conditions), and "function." (§3.2) This will lay the ground for the introducing the "Practical-Epistemic

---

[109] In addition, objective reasons are not necessarily overriding reasons. In Chapter 2 I'll distinguish between local and global practical problems. Objective reasons, as I'll discuss them in this chapter, correspond to local problems (or local needs for flourishing). Overriding reasons (or all things considered reasons) correspond to global problems (or global needs for flourishing).

Access" (PEA) *functional* analysis of subjective reasons and representations. (§4) I have already introduced this analysis in the Introductory Chapter and in Chapter 1, so I won't reiterate it until §4.

I'll conclude by highlighting some important features of the PEA analysis including the "surprising" implication that alethic conditions are not a constitutive feature of representations generally.

4.1 Representations & Subjectivity

It may strike the reader as strange to propose that derived representations are a species of subjective reason. After all, philosophers are used to thinking of subjective reasons as psychological states-- representations, as it were, "inside the head." But derived representations are paradigmatically not in the head-- they're things like sentences, pictures, and maps. In suggesting that derived representations are a species of subjective reason, I'm not suggesting that they are actually inside the head. Instead, I'm suggesting that an important species of subjective reasons exist outside of the head. But some philosophers might resist this idea. So, let us begin by considering why this is-- why do philosophers tend to think of subjective reasons as things in subjects' heads or minds? Is there a good philosophical reason?

4.1.1 A "Perspectival" View: Representations as Interpretations

The best sort of philosophical reason one might hope for here would be one based in an airtight theory of subjective reasons. Such a theory would allow us to compare non-mental or derived representations to the criteria for subjective reasons and thereby pinpoint precisely what properties exclude non-mental representations from the category. But philosophers do not seem to be relying on any such explicit theory--- not even a rough and ready one. Rather, philosophers seem to be relying on an intuitive sense that subjective reasons are *subjective* and therefore part of a subject, whereas derived representations exist out in the world and are not essentially connected with any subject. However, even if derived representations are not essentially connected with any single subject, it does not follow that they are not essentially connected with subjects, and so it does not follow that they are not, in a crucial sense, subjective.

There's good reason to think that representations as such cannot exist apart from subjects. A representation must be a representation *for* someone.[110] This point is sometimes put in terms of "interpretation," but it's important to be very clear about what this means. For often what we interpret already has a meaning; it's already a representation. If the idea of interpretation brings to mind examples in which what is interpreted already has a meaning, we may be encouraged to "reify" representations--- we may be inclined to think that representations exist apart from subjects. But of course, the representations that exist for us to interpret, e.g., sentences, pictures, models, and maps, are representations that have been produced by and for subjects. Such representations are produced for an audience to interpret, so if they already have meanings those

---

[110] See e.g. Dennett (1978), Pierce (1883, 1977, 1998), Millikan (1984), Ramsey (2007).

meanings plausibly have their source in the subjects doing the producing. Just as one must possess something to give it to someone else, it's plausible that subjects must be in possession of the interpretations they publicly produce. These original meanings are themselves subjects' interpretations of facts. A subject's interpretations of facts are, in some sense, in her. They're connected with her dispositions for action. But subjects can also take their interpretations out of their heads, as it were-- they may "objectify" their interpretations of facts by making them public. In other words, they may take their first-personal interpretations of facts and turn them into third-personally available objects of interpretation.

We might imagine a primitive human, Eve, discovering that apples are edible, hence acquiring a first-person interpretation of apples as things she can eat. And we may imagine that Eve brings an apple home to her tribe and eats it in front of them. This demonstration is a way of sharing her interpretation, where the goal of sharing one's interpretation is to create something like common knowledge-- something like common dispositions to respond to similar facts. Those who see Eve's demonstration can interpret it as meaning that they can eat apples as well. Of course, for this to work, humans already had to be disposed to interpret demonstrations like this in a certain way-- they had to be disposed to see things that other people do as things that they can do and vice versa. Humans might have some such dispositions innately. So, sharing interpretations or perspectives plausibly depends on preexisting common dispositions to respond to similar sorts of facts and at the same time consists in creating new dispositions of this kind. By objectifying and thus sharing their interpretations, subjects position other subjects to acquire similar dispositions to interpret facts in similar ways. In this way, shared interpretations plausibly supervene on individuals' interpretations. But that does not mean that the shared interpretations are not themselves interpretations.

146

Sharing interpretations also makes it possible to borrow interpretations. In our toy example, Eve acquired an interpretation of apples as edible first-hand and shared that interpretation with the rest of her tribe. As a result, the others in her tribe were able to acquire that interpretation second-hand: they acquired an interpretation of her interpretation in that they interpreted her demonstration. They interpreted Eve's demonstration in the sense that they acquired dispositions to eat apples and to demonstrate to others that they can eat apples. Such shared dispositions plausibly play a number of social and individual functions. Languages may provide sophisticated and efficient ways of sharing interpretations or perspectives on facts. Language users must have shared dispositions to employ language so as to successfully convey contextually relevant information, i.e., contextually relevant interpretations of relevant facts. Conveying information by means of a language may at root be a sophisticated way of sharing interpretations of facts, i.e., of creating common dispositions to respond to similar facts.

The suggestion, then, is that all representations are "subjective" in that they are essentially subjects' interpretations of or perspectives on subject-matters. This "perspectival" way of understanding representations, in turn, suggests a way of characterizing some important differences between original and derived representations. Derived representations may correspond to "objectified" interpretations or perspectives-- individual or group perspectives that are expressed, produced, or otherwise presented to subjects as a subject-matter for interpretation for the sake of creating shared interpretations or perspectives. And shared interpretations or perspectives may consist in common (or otherwise socially coordinated) dispositions to respond to similar facts. Fundamental perspectives or interpretations may correspond to original (psychological) representations, where these may consist in an individual's dispositions to respond to facts. And if an individual belongs to a social form of life, it may have dispositions to share its

147

perspectives/interpretations. In other words, certain forms of life may be disposed to objectify their perspectives.[111]

4.1.2 Putting Propositions in Perspective

The standard strategy assumes that subjective reasons (original representations) are propositional attitudes, or something very similar. Call this the "propositional attitude model" of original representations. The "perspectival" way of understanding the difference between original and derived representations discussed above suggests that the propositional attitude model is ill suited to provide an account of subjective reasons, i.e., original representations.

As we saw in Chapter 1, propositions are entities with truth-values (or alethic conditions), and insofar as they are familiar to us, they are entities to which we are related as objects, i.e., as subject matters. And to the extent that the explanatory value of the propositional attitude model stems from the fact that propositional contents are theoretically familiar entities, there's reason to treat the propositions of propositional attitudes as having the familiar sorts of properties. Accordingly, the representations on which propositional attitudes are modeled are presumably derived representations. Such representations demonstrate, stand-in-for, model, or describe interpretations of facts, in this way making them public. They represent-- that is, their content gives or specifies or in a sense articulates-- their truth, veridicality, accuracy, or satisfaction conditions. The standard strategy-- via the propositional model-- generalizes this property of derived

---

[111] There's obviously a connection between my notion of objectifying perspectives and expressive forms of communication. Exploring this connection is a topic for future research.

representations to all representations. It treats original representations as agents' "attitudes" towards propositions.

But the propositional attitude model alienates subjects from their fundamental interpretations. Propositional attitudes make a subject's fundamental interpretation of facts into subject-matters to which subjects are related in the manner of a private audience, for propositions are the sorts of things that are presented to a subject for interpretation. So, the propositions of propositional attitudes must not, fundamentally, be the interpretation of the subject who takes attitudes towards them, since they are things that must be interpreted via some more fundamental interpretation. Agents do not produce these representations so much as bump into them. They're interpretations an agent might adopt or entertain-- towards which an agent might take some attitude; they convey a perspective to a subject by themselves serving as subject-matter of interpretation.

So, if subjects are fundamentally intentionally related to propositions, two questions arise. First, how is it that subjects are related to propositions? In what sense do they interpret them? If the answer is that subjects fundamentally interpret propositions by taking attitudes towards them, and if attitudes could be objectulaist-- that is, attitudes towards facts rather than propositions representing facts-- then this suggests that attitudes, rather than propositions, are that in which original intentionality consists. Second, since subjects must interpret propositions (perhaps via their attitudes), this raises the question: whose interpretations are those propositions? The standard strategy will surely insist that they are the subject's own interpretations, but it's not clear how they can do this if subjects are related to propositions as subject-matters. Perhaps a proponent of the standard strategy will suggest that the propositions of propositional attitudes are produced by subjects, albeit not intentionally or consciously. They are produced by subjects in the sense of

being produced by a subject's subsystems. But a subject's subsystems are not themselves subjects. And if a proposition is an objectified interpretation, it's not clear how anything other than a subject can produce them.[112] Consequently, the idea that a subject's subsystems produce the propositions of propositional attitudes seems to portray subsystems as a kind of homunculus residing in a subject that produces propositions and privately presents them to the subject. The idea seems to be that since the subsystem (homunculus) is in the subject, and presents the proposition to that subject, the proposition constitutes the subject's interpretation.

On such a picture, what explains a subject A's original representations -- her first-personal interpretations of facts--is some inner-homuncular subject B which produces and presents propositions to A. There are two problems with this. First, A's first-person interpretation seems to go not just with the proposition but with the attitude she takes towards it. Propositions present things as being some way, but A does not necessarily take things to be that way. Perhaps the proponent of the standard strategy will suggest that B (the homunculus) takes things to be some way. But then it's not clear how B's interpretation can be identified with A's interpretation. Second, B's ability to produce a proposition (to have an interpretation) is mysterious and therefore doesn't seem to be of much help in illuminating what it is for A to first-personally represent the world. On the other hand, if the proponent of the standard strategy attempts to illuminate how it is that B produces propositions, a regress threatens.[113] After all, if we explain A's fundamental

---

[112] Of course, there might be machines that, as it were, "spit out" propositions. But that's because subjects produce and use machines for this purpose. The fact that the things they spit out are propositions depends essentially on the fact that subjects produce the machine for this purpose and interpret its outputs in this way.

[113] It doesn't help to try to stop the regress by treating sub-subject-level mechanisms, i.e. systems that are not (in any non-equivocating sense) themselves subjects (interpreters, things that can have interpretations) as the things producing the interpretations the subject privately interprets. For how could sub-subject-level mechanisms produce interpretations proper if they are not properly subjects?

intentional access to the world in terms of B's intentional access to the world, either we take B's access as an unilluminated primitive, or we seem forced to explain B's access in terms of the access of some further homunculus C, and so on *ad infinitum*.

The considerations with which this section began suggest that original representations just are agents' interpretations of facts. While they can be, at least for humans, the subject-matter of subjects' interpretations, this is not what they are fundamentally. Fundamental (original) psychological representations (subjective reasons) are not themselves subject-matters but rather ways of having subject-matters i.e., ways in which an agent is fundamentally subjectively related to a subject-matter. In light of this, the standard strategy's adoption of the propositional model of original representations or subjective reasons appears to *alienate* subjects from their fundamental interpretations or perspectives. This alienation is the result of treating all representations on the model of derived representations. For although these too are subjects' perspectives, they're objectified perspectives which are distinct from and parasitic on fundamental perspectives.

4.1.3 Taking Stock

This section set out to address the assumption that only mental representations are subjective reasons. I began by considering a bit of reasoning that might seem to support this assumption: subjective reasons are subjective, and only mental representations are subjective. I've tried to show that the latter premise is dubious: all representations are subjective in that they are subjects' interpretations of subject- matters (facts). Derived representations are interpretations of subject-matters that subjects are positioned to interpret; they are "objectified" interpretations, or interpretations as interpretable subject-matters. Their "objective" nature may encourage the

151

philosophical tendency to alienate representations from their essential relation to subjects, but this tendency should be resisted as it appears to lead to nothing but trouble.

Here, then, is the picture so far: Representations intentionally relate subjects to subject-matters. An original representation is a subject's state of being intentionally related to, i.e., her state of interpreting, some subject-matter, while a derived representation is a subject-matter R that intentionally relates a subject to some further subject-matter X only if she interprets R. In other words, R is an objectified interpretation of X; an agent interprets X by way of interpreting R. And an agent interprets R just as she would interpret X directly, *viz*, by means of a fundamental interpretation-- the representation that constitutes her intentional relation to (interpretation of) R (or an immediate X). An agent is not intentionally related to subject matters by virtue of interpreting her fundamental interpretations; she is not, fundamentally, related to her fundamental interpretation as a subject-matter she interprets. This way of understanding the nature of an intentional or representational attitude stops the regress that threatens as a result of the standard strategy's adoption of the propositional model which, in turn, is motivated by its acceptance of the descriptive thesis.

4.2 Subjective Reasons *out-of-which* & Subjective Reasons *on-the-basis-of-which*

The fact that philosophers tend to hold that subjective reasons are distinctively *psychological* (or original) representations suggests that an analysis of subjective reasons might shed significant light on an analysis of mental representation and representation more generally. But, as noted, philosophers typically approach these issues from the opposite direction: they

presuppose a theory of representation that assumes the truth of the descriptive thesis and put it to work in accounts of the nature of subjective reasons. The general worry we've been addressing is that subjective reasons can't account for representation because not all representations are mental representations and only mental representations are subjective reasons. For if matters are as this worry suggests, an analysis of original representation that's based on an analysis of subjective reasons is likely to lack generality or unity-- it's unlikely to reveal the properties in virtue of which original representations belong to the same genus as derived representations. The last section suggested that this worry may be ill-founded: subjective reasons may include derived representations. I appealed to the relationship between representations, interpretations or perspectives, and subjects, to argue that all representations-- not just mental representations-- are subjective. So, if being subjective is a criterion on being a subjective reason, all representations meet this criterion.

But one may object that I equivocate. For one may suggest that subjective reasons must be subjective in the sense of being the private perspectives belonging to the individuals whose actions they determine or guide. After all, we noted above that original representations are often thought to constitute an individual's perspective on the world: original representations are not *what* we represent but *how* or *that through which* we represent. What we originally represent just is the world from our perspective or point of view. And *how* we originally represent *what* we originally represent, i.e., our subjective reasons, directly determines how we act, or at least how we intend to act, in response to the world. And subjective reasons directly determine the actions only of the agents whose individual perspective they constitute. They are to that extent "private"; they do not seem to be the kinds of subjective reasons different individuals could literally share. Accordingly,

we may say that original representations are subjective in the sense of being *intra-subjective reasons*, and that they're the reasons *out of* (o*r from*) *which* a particular agent acts.

But it seems to me that there's a species of subjective reason that isn't private in this way-- an *intra-subjective* species of subjective reason that (a) different individuals can literally share, and that (b) isn't as directly connected to the determination of an agent's actions. To wit, it seems possible for agents to have not only *intra-subjective reasons out of which* they act, but also *inter-subjective reasons on the basis of which* or *on which* they act.

Regarding (a): It seems possible for agents that are capable of cooperative action and communication to *share* an interpretation, understanding, or perspective. I don't just mean that it's possible for distinct agents to be related to the same subject-matters and have their own versions of the same general type of perspective on them, as when we are all visually related to my cats, and we're all dying to play with them. I mean, rather, that it seems possible for two agents to share exactly the same self-identical subjective reasons.

Consider, for instance, the earlier example of Eve. Eve's demonstration can itself serve as a reason for others in her tribe to interpret apples as edible. Insofar as the demonstration leads them to interpret apples as edible, they acquire certain dispositions to respond to apples, but they only acquire these dispositions if they are already disposed to interpret Eve's demonstration as relevant to their own response to apples. They have to be able to make the move from seeing that Eve can eat apples to themselves adopting the habit of eating apples and of reproducing Eve's demonstration so as to share that interpretation with others. In this way, they take Eve's demonstration-- her objectification of her interpretation-- as a reason for adopting that interpretation.

For another example, take the fact that it is raining. We may represent this fact by means of a model or description (proposition) which we may use to tell people that it is raining who cannot see or hear this fact for themselves. Two (or more) people could represent this fact in an objective or third-personal sort of way, e.g., via the proposition "it is raining," by virtue of hearing the same weather report. And two or more people may endorse this proposition, where this seems to amount to treating it as a reason to act as if the state of affairs it describes obtains, i.e., to act as if it is true that it is raining. The proposition thus serves as a reason (together with some relevant need or desire) to, e.g., put on a raincoat or rain boots, to grab an umbrella, to reschedule that picnic in the park, to take the bus rather than walk to work, to close the windows, to take down the laundry and bring it inside, to let your friends, family, colleagues or coworkers know that it is raining, etc.. In this way, to endorse the proposition "it is raining" is to adopt a third-personal (descriptive) interpretation as one's own. In so doing, an agent comes to have an *inter-subjective reason*-- a shared interpretation-- as her reason, i.e., as the (or a) reason *on the basis of which* she acts. [114]

Regarding (b): I mentioned that when an agent endorses a proposition (an intersubjective reason), she's disposed to treat the proposition as a reason to act as if the state of affairs it describes obtains. But there are many ways of being disposed to act as if the state of affairs it describes

---

[114] Inter-subjective reasons are interpretations that stand-in-for the kinds of facts to which our actions are responses, such as the fact that it is raining. But they may also be interpretations that stand-in-for facts involving the ends at which our actions may be directed, or for the sake of which we might act. Such propositions may describe or model facts in terms of their implications for our flourishing, or in terms of how an agent's flourishing is or would be impacted by some state of affairs. The facts described may involve agents performing particular actions and thereby producing certain consequences that, perhaps to a certain extent, do or do not promote the relevant ends. For instance, the sentence "getting rained on would not be good for you-- you might get sick" may be a descriptive interpretation of a causal fact involving the rain and your flourishing (your health). We might think of it as representing a "nearby possible world" in which getting rained on causes your health to suffer. Similarly, the sentence "You should not go out into the rain without a jacket" may be understood as representing a nearby possible world in which you go into the rain without a jacket and consequently suffer. We may also understand an order or command such as "don't go out there without a jacket!" as representing a nearby possible world in which you go outside without a jacket and consequently suffer.

obtains. Consider again the proposition "it is raining." What it means to be disposed to act as if it is true that it is raining depends on the ways in which one is predisposed to adjust one's plans and activities to the fact that it is raining. If you were planning a picnic, you might reschedule it. If you were planning to go for a walk, you might take an umbrella. Or if it occurs to you suddenly that you'd like to go for a walk, you might take an umbrella or put on a rain jacket. But we might also imagine an agent for whom rain is alien-- an agent who has never experienced rain first-hand, and who therefore has never learned how to adjust his plans and activities to the fact that it is raining where such adjustments involve things like bringing an umbrella. This agent may nonetheless have learned to make various broadly intellectual adjustments. For such an agent, then, endorsing the proposition "it is raining" may simply involve being disposed to disagree with someone who denies that proposition or asserts its contradiction, perhaps in slightly different terms, e.g. "there's no water falling from the sky." In this case, endorsing the proposition doesn't involve being disposed to grab an umbrella or rain jacket before going outside-- and not because this agent doesn't care about getting wet, but rather because he does not yet associate rain with this kind of response; he lacks the intersubjective perspective on rain that would dispose him to e.g., put on a rain jacket before going outside. Intuitively, this Alien does not "grasp" the proposition "it is raining" the way ordinary humans do in that the Alien only shares a subset of the sorts of dispositions to encountering that proposition ordinary humans do. On the other hand, to the extent that the Alien shares that subset of dispositions, it seems intuitively wrong to say that it has no grasp of the proposition.

Likewise, an agent may endorse the proposition "the right response to X is Φ-ing" in the sense of being disposed to agree with it, deny its contradiction, etc., but lack the relevant intra-subjective perspective that disposes her to Φ upon confronting an X. This is because inter-

subjective reasons are "detached": they provide an agent with indirect access to the facts for which they stand-in.[115] What one responds to directly are the sentences or propositions that purport to represent a fact, so it's possible to be in a position to do things with a proposition-- e.g. reason abstractly with it-- without being in a position to interact directly with the facts the proposition represents. In a word, meeting with the proposition "it is raining" is not the same as meeting directly with the rain. In addition, to act *on the basis of* an inter-subjective plausibly requires that one act *out of* appropriately related intra-subjective reasons where these are one's first-personal ways of understanding situations that directly determine how one responds to them. So for instance, one puts on a rain jacket on the basis of the inter-subjective reasons "it is raining" by virtue of acting out of the intra-subjective reason-- one's fundamental interpretation of the sentence "it is raining" -- if one's fundamental interpretation of that sentence is what immediately directs one's action of putting on a rain jacket.

If this is right, then there's a further important difference between intra- and inter-subjective reasons: intra-subjective reasons correspond to skills that involve recognitional abilities and that have the power to move us to act directly in response to facts, while inter-subjective reasons create abilities for indirect responses that both piggyback-on and extend the reach of intra-subjective reasons. And if this is right, then the fact that derived representations are not as directly involved in the determination of action doesn't entail that they are not subjective reasons: it may entail that they are not intra-subjective reasons, but intra-subjective reasons may not be the only species of subjective reasons. Nor does the fact that inter-subjective reasons may piggyback- on intra-subjective reasons entail that they aren't subjective reasons. After all, many philosophers

---

[115] See also Millikan (1995, 2004), Sterelny (2003).

think that "derived" representations piggyback-on original representations in something close to the way just described, and they don't take this to entail that only original representations are *really* or *authentically* representations. Quite the contrary: challenges concerning representational authenticity typically target original representation and not derived representation.

Now, if there are indeed both intra-subjective and inter-subjective reasons, we're faced with the question: what is it, exactly, that they have in common?

I think we may understand the propositional attitude model of mental states as providing one version of an answer: what's common is the nature of the interpretations-- i.e., representational contents-- they involve. More slowly: the standard strategy takes the inter-subjective reasons on the basis of which agents act as its theoretical basis and tries to adapt its theory of this species of subjective reason to illuminate the intra- subjective reasons out of which we act. This strategy attempts to unify these two species of subjective reasons in terms of a common core-theory of representational content that assumes the truth of the descriptive thesis. It thus attributes the same kind of representational contents-- namely, propositions or something roughly analogous -- to both the subjective reasons out of which we act (intra-subjective reasons) and the subjective reasons on the basis of which we act (inter-subjective reasons).

Perhaps philosophers have an implicit intuitive sense of the inseparability of the notion of subjective reason and representation. And perhaps this explains the impulse to use representation to unify these different species of subjective reasons. But since philosophers already had a theory of representation based on descriptive sentences, and since they associate such representations less with subjectivity and thus less with subjective reasons, they didn't really consider the possibility that an analysis of representation might depend on a prior analysis of subjective reasons. So, the connection between subjective reasons and representations was only considered for the private,

paradigmatically mental variety of subjective reasons (intra-subjective reasons, original representations), and philosophers adopted the standard strategy of trying to adapt propositional model of representation (inter-subjective reasons) to account for them.

But there's a neglected alternative way of resolving the puzzle of how it is that subjective reasons can wear-- at least for human agents-- both a private and a public face: posit that subjective reasons and representations generally are two sides of a single coin or *genus* that admits of importantly different *species* corresponding to these two faces. If this explanation is correct, it should be possible to provide an analysis of the genus in a way that illuminates both the private and the public species. The neglected strategy thus suggests the following procedure:

1) Identify the genus *Representation* with the genus *Subjective Reason*.

2) Develop an analysis of the latter that's designed to accommodate both the mental and nonmental species of subjective reasons.

3) Base one's theory of representational content on the analysis of subjective reasons developed in step (2).

Since philosophers associate fundamental mental representations more closely with subjective reasons than they do non-mental (non-fundamental, public) representations, the proposal is effectively to invert the standard strategy. For I've suggested that the standard strategy starts with a theory of representation that treats the derived variety as paradigmatic and tries to understand the original, psychological variety in those terms. The neglected alternative strategy instead begins by theorizing original representations or fundamental interpretations-- treating these

as paradigmatic-- and adapting that theory to account for derived (non-fundamental and/or non-mental) representations.[116]

Given the problems with the standard strategy I touched on in §2 above, it seems to me high time that philosophers explore this neglected strategy. So that is what I propose to do.

4.3 A Neglected Strategy

This section and the next will pursue the "Neglected Strategy" introduced at the end of the last section. This section will focus on the second step of implementing that strategy by working towards an analysis of subjective reasons and intra-subjective reasons that will be able to accommodate the species of representation to which (I've suggested) they correspond: original and derived representations respectively. To this end, I'll begin by describing three kinds of reasons for action and their relationship to one another. First, I'll describe *Objective Reasons*, which are facts about an agent's needs for flourishing and the targets with which she needs to interact to satisfy those needs. Second, I'll describe two kinds of subjective reasons, both of which are modes of understanding, grasping, or interpreting objective reasons: intra- subjective reasons-- the embodied kind of subjective reasons *out of which* agents act-- and inter-subjective reasons-- the extended, objectified, or detached kind of subjective reasons o*n the basis of whic*h agents act

---

[116] There's a sense in which I'm keeping with a broadly Gricean tradition of ultimately analyzing derived representations in terms of original representations. However, that tradition makes assumptions about the nature of original representations that are based on extrapolating presupposed essential features of derived representations. The way the Gricean approach to representation is executed within the context of the standard strategy puts derived representations (or inter-subjective reasons) first in its theorizing insofar as it assumes that the descriptive thesis will be true of original representations (intra-subjective reasons).

(§3.1). I'll then build towards an analysis of subjective reasons that begins with intra-subjective reasons. Since I ultimately want to apply this analysis to representation, I'll build the analysis around making sense of aboutness, fitness or success conditions, and the constitutive function of intra-subjective reasons (§3. 2).[117]

4.3.1 Objective & Subjective Reasons

Action explanations appeal to the reasons *for which* or *because of which* an agent acts. Such reasons appear to be of at least two types.[118]

First, an agent's actions are for the objective facts *at* or *with respect to which* her actions are *directed*. These facts are *Objective Reasons*. Objective reasons may be divided into two types. The first type includes facts about an agent's needs (broadly considered) for flourishing. I'll refer to these as *needs*. The second type includes facts about the world that are relevant to meeting such needs. I'll refer to these as *targets*.

Objective reasons are reasons *for which* an agent acts in that her actions are teleologically and efficient-causally directed at the satisfaction of her needs and at the targets with which she must interact in a certain way to satisfy her needs. Objective reasons are reasons *because of which* an agent acts in that her actions are responses to need and target facts. For example, an agent needs nourishment to flourish; when low on nutrients, one needs to ingest some. One's being low on nutrients (at some particular time) is a fact about one. Generally, needs are *for* types of interactions,

---

e.g., eating, between whole agents and environmental targets, e.g., food. The existence or occurrence of particulars with which an agent can interact to satisfy her needs, e.g., the particulars that can nourish her, are facts, and there are relational facts which we might think of as "practical potentials" or "need-satisfaction-interaction-potentials" whose *relata* consist in agents and targets. In short, objective reasons are facts that favor certain kinds of responses for certain kinds of agents. They are not interpretations of facts. For instance, the presence of a treat is an objective reason for Ruch to eat it if Ruch is hungry.[119]

But facts are objective reasons *for* agents only insofar as agents are in a position to intentionally respond to them as such, i.e., only insofar as her actions are teleologically and efficiently-causall*y* directed at them. So, in addition to objective reasons, there are *Subjective Reasons*. Subjective reasons are the way an agent takes, understands, interprets, appreciates, represents, objective reasons, or to her perspective of them. Objective reasons determine how an agent intentionally acts via an agent's subjective reasons; an agent's subjective reasons, together with action actualization- or execution-relevant facts about the situation in which they occur, directly determine--or one might say *direct*--her intentional actions.[120,121]   So, for example, Ruch the cat sees a treat and eats it, she takes the treat as something to be eaten. Her taking the treat as something to be eaten is the intra-subjective reason out of which she acts.

Now, as we saw in the last section, there seem to be two varieties of subjective reasons: intra-subjective reasons *out of which* we act, or which we *enact* or *act-out* and inter-subjective

---

[119] Objective reasons are not necessarily overriding reasons. In Chapter 2 I'll distinguish between local and global practical problems. Objective reasons, as I'll discuss them in this chapter, correspond to local problems (or local needs for flourishing). Overriding reasons (or all things considered reasons) correspond to global problems (or global needs for flourishing).

[120] E.g., enabling conditions.

[121] That an agent has some subjective reason or other is a relational fact about her. The relevant *relata* are agents and objective reasons (needs and targets). To illuminate the nature of a subjective reason is thus to illuminate the nature of this relation.

reasons, i.e., the objectified, extended, detached subjective reasons on *the basis of which* agents act.

Intra- subjective reasons position agents to directly respond to objective reasons. They're constituted by exercises of agents' embodied *abilities* to act in ways that are directed at objective reasons. Intra-subjective reasons are had by both human and non-human animal agents. They're an agent's *fundamental understanding* of the facts that correspond to her needs and targets (objective reasons). Their vehicles are internal to agents such that they directly causally shape what an agent does.

Inter- subjective reasons position agents to respond to objective reasons by standing-in-for them. In addition, inter-subjective reasons can, at least for human agents, stand-in-for intra-subjective reasons as well as other inter-subjective reasons.[122] Their vehicles are external to agents (though they may be "internalized" or internally simulated) and consequently do not directly causally shape what an agent does.

For example, Ruth sees a piece of cake and takes it that it is a piece of cake and that it is edible and hence something that can be eaten by someone like her. In this way, she takes it that the cake is a kind of fact that can serve as a kind of inter-subjective reason for acting. Ruth may, at the same time, have an intra-subjective reason to eat the cake in the same way Ruch has an intra-subjective reason to eat-- Ruth may have an inclination to eat the cake. But Ruth, unlike Ruch, can also take the fact that she has that intra-subjective reason (that she is inclined to eat the cake) as an inter-subjective reason: Ruth can appreciate that she wants to eat the cake just as she can appreciate

---

[122] An intra-subjective reason can only access (represent) another intra-subjective reason if the latter is objectified, i.e., only if the latter is made into an object to which she can respond or with which she can intentionally interact. Inter-subjective reasons are the means by which an agent understands her understanding where inter-subjective reasons are objectifications of an agents' understanding, i.e., entities that stand-in-for and enable intentional interactions with understanding.

that you or I want to eat some cake as a reason *one* might eat some cake. That is, she can appreciate her desire for the cake as a type of reason one might have for eating some cake, and in this way, she can appreciate herself as sharing in an intra-subjective kind of reason. She can also appreciate that the cake has a lot of sugar and fat and little other nutritional value, so she can take that as an intra-subjective reason -- a reason for one, such as her-- not to eat the cake. If Ruth decides that she has better reason not to eat the cake than to eat it, then, hopefully, she will not enact her intra-subjective reason-- her inclination to eat the cake-- and will instead enact her intra-subjective reason to resist acting in the way she deems she has best reason to act (or her intra-subjective reason to act in accordance with what reason demands). In this way, Ruth may decide which of her intra-subjective reasons to actually enact on the basis of what she thinks she has the most inter-subjective reason to do.

As noted in the introduction, inter-subjective reasons are "objective" in the sense that a) they are public or explicit interpretations of facts and b) they are interpretations of facts as reasons for *someone* or *anyone* (perhaps in a certain position or context) to act in some way. They are not, however, objective reasons as I've defined them.

4.3.2 Towards a Theory of Subjective Reasons

I ultimately want to identify subjective reasons with representations; to provide a theory of both at once. And where the standard strategy takes inter-subjective reasons as its paradigm (or so I've suggested), I'll be building my theory around intra-subjective reasons. So, in this section I'll work my way towards a theory of subjective reasons by thinking about how to understand three

key features of representation-- aboutness, fitness or success conditions, and function-- in the context of intra-subjective reasons.

Aboutness

What are intra-subjective reasons about? I've suggested that intra-subjective reasons are an agent's fundamental access to objective reasons. So, I propose to refer to the relevant kind of access intra-subjective reasons provide as "Practical-Epistemic Access" or "PEA" for short. The idea is that intra-subjective reasons position agents to intentionally respond to objective reasons and in so doing constitute an agent's immediate or direct "practical-epistemic" grasp of objective reasons. Both inter-subjective reasons and intra-subjective reasons provide this kind of practical-epistemic grasp of or access to objective reasons in that they both constitute a kind of knowledge or awareness of facts (objective reasons) that consists in being in a position to intentionally respond to facts. But as we saw in the previous section, in the case of inter-subjective reasons, the way in which an agent is positioned to respond is less direct than the way in which an intra-subjective reason positions an agent to respond. Inter-subjective objective reasons are objectified and thus detached in that their vehicles are external to agents; they constitute perspectives that are themselves subject-matters of a more fundamental perspective.[123] I've suggested that intra-subjective reasons constitute this more fundamental perspective; a way of grasping or practically-epistemically accessing objective reasons that directly shapes how an agent is positioned to respond to them, i.e. how she will respond, if she does.

---

[123] Such vehicles might be "internal" or "internalized" --for instance, when we imagine places and things to describe and reason about them. These are still "external" in the sense that agents are related to them as objects-- as things to be described and reasoned about.

Plausibly, then, intra-subjective reasons are about objective reasons, i.e., facts. But in contrast to inter-subjective reasons, intra-subjective reasons don't stand-in-for objective reasons. Both inter-subjective and intersubjective reasons are about objective reasons by virtue of enabling, preparing, or positioning agents to respond intentionally to them; they're both forms of PEA to objective reasons (i.e., facts). But intra-subjective reasons constitute a more immediate or direct kind of access that doesn't involve standing-in-for objective reasons. They connect an agent to objective reasons simply by preparing actions that are directed at them.

Accordingly, I think we should consider whether intra-subjective reasons are also about these actions. That is, are they about both objective reasons and intentional actions?

If we were to ask this question about inter-subjective reasons, we'd probably say no. Even though, like intra-subjective reasons, inter-subjective reasons essentially position agents to produce intentional actions in response to objective reasons, i.e., even though they're both forms of PEA, inter-subjective reasons effect this positioning by standing-in-for the objective reasons (in the right way). And they don't *always* stand-in-for actions, though obviously some do. For instance, the proposition "it is raining" positions agents to respond to the state of affairs it stands-in-for (if it obtains), but it doesn't stand-in-for (name or describe) any particular actions.

Intra-subjective reasons certainly don't stand-in-for intentional actions, but they don't stand-in-for the objective reasons they're about either. One might wonder how intra-subjective reasons can be about or stand for actions or objective reasons if they don't stand-in-for them. The answer, I propose, is precisely by means of positioning or preparing agents to produce intentional actions: the intentional directedness of intentional actions is one and the same as the intentional directedness of the intra-subjective reasons that prepare them (or that are their preparedness) in the manner of an intentional action acorn (see Chapter 2). The suggestion, then, is that an intra-

subjective reason is about objective reasons insofar as those objective reasons are involved in the intentional action an intra-subjective reason prepares. So, when an agent is in a position to produce an intentional action in response to a target, that means she's in a position to treat that target as part of a means (or part of a part of a part of… a means) to some end, i.e., the satisfaction of some need. And when an agent is in a position to produce an intentional action in response to a need, that means she's in a position to respond to targets as part of a means to satisfying that need.

Even if intra-subjective reasons are about objective reasons through or by virtue of preparing intentional actions that are directed at them, it still might not be obvious that it makes sense to include intentional actions in their contents, i.e., the things they're about. Here, then, is a further reason: An agent may be related to multiple targets that could figure, in slightly different ways, as part of a type of means that aims at the satisfaction of one particular need. Plausibly, when this happens, an agent has slightly different subjective reasons involving those various targets corresponding to the slightly different actions in which she's positioned to involve them all of which would be means to the same end. In other words, the particular action an intra-subjective reason positions an agent to produce seems to be part of what individuates intra-subjective reasons. And this seems to be reflected in what an agent appreciates-- in the way in which she grasps objective reasons. The following example may make this latter point more salient:

> **Cat Treat:** When Ruch the cat is hungry, the fact of there being a treat on the floor in front of her is an objective reason for her to eat it. When Ruch's hunger has already been satiated and she needs to work off some energy, the fact of there being a treat on the floor in front of her is an objective reason for her to bat it around and chase it.

167

I think it's natural to say that Ruch appreciates the treat in different ways: in being prepared to eat the treat, she has one kind of intra-subjective reason that's about it and in being prepared to play with the treat, she has another kind of intra-subjective reason that's about it. Each of these intra-subjective reasons is its own sort of practical understanding. But since the target is the same in both cases, the target can't account for this difference. Nor need we appeal to actual couplings of target and need instances. Ruch may be neither hungry nor playful, but she is in a position to eat the treat should she become hungry or to play with it should she become restless. So, neither need obtains, yet Ruch has two distinct intra-subjective reasons involving one and the same target; two ways in which she's practically appreciating the self-same treat. In order to account for the different ways in which Ruch is practically appreciating the treat, and in order to account for why there are two intra-subjective reasons here rather than one, I think we should say that intra-subjective reasons are about both objective reasons (targets like the treat) and the actions one's in a position to produce in response to them. An intra-subjective reason is, in a sense, a mode of presentation of, i.e., a mode of PEA to, objective reasons by-means-of intentional actions, or of intentional actions and therefore of the objective reasons at which those actions are directed.

You might be thinking "what she's describing is unlike any kind of aboutness I've ever heard of! I hardly recognize it as aboutness!" If you're thinking thoughts like this, that's good. Because it means you recognize that we're departing from descriptive aboutness. If you're feeling really frustrated, you can skip ahead to Chapter 4. There I'll posit that intra-subjective reasons (original representations) are *generatively* about actions and the facts at which they're directed. Reading my description of generative aboutness might help you understand the kind of aboutness I'm attributing to intra-subjective reasons. But you're still not likely to discover much in terms of that comforting sense of familiarity. The point of this dissertation is, after all, to introduce

168

something genuinely, and perhaps even somewhat radically, new. Like Chapter 1, this chapter seeks to develop an *initial* case for thinking that the kind of aboutness I'm attributing to intra-subjective reasons (which I'll later name "generative aboutness") is genuine representational aboutness. In Chapter 1, this took the form of arguing that there's a promising representationalist construal of the intentional relation on which *de actu* contents might figure in a fundamental form of practical judgement that issues directly in actions.[124] In the present chapter, this case takes the form of a demonstration: generative aboutness, i.e. the kind of aboutness had by intra-subjective reasons or original representations (*de actu* contents (Chapter 1) or intentional action acorns(Chapter 2)), shares a genus with the familiar propositional (or what I'll later call "surrogative") kind of aboutness, as both are forms of PEA. Due to neglect, the intra-subjective (generative) species of aboutness is bound to look rather exotic. But it's important that we not let this distract us from what it has in common with propositional aboutness, namely: aboutness-- which I'm suggesting may be understood or analyzed as PEA. If there are two species of subjective reasons (intra- and inter-) and two species of representation (original and derived) there may well be two species of aboutness. (Indeed-- I suggested in Chapter 1 that there are two species of PEA.)

Fitness

Philosophers tend to think that representational success consists in a kind of "fit" between a representation (or representational content) and the facts it represents. As per the Descriptive Thesis, the relevant kind of fit is standardly cashed-out in terms of a kind of descriptive match or satisfaction or accurate labelling: representations that fit their objects are true, veridical, accurate,

---

[124] And in Chapter 4 I'll argue that *de actu* contents which generatively represent actions satisfy GENERALITY, ACTIVITY, and NORMATIVITY.

or satisfied. And a representation's fitness (or success) conditions are standardly considered its essential trait. I have suggested that the standard conception of representational fitness may be an artefact of the standard derived-representation (proposition)-based model of representation, and I've proposed to identify original representations with intra-subjective reasons and to develop a theory of representation on their basis. So, what would it mean for an intra-subjective reason to succeed? What kind of "fitness" befits an intra-subjective reason? Here is what I propose:

In the good case, there's a "practical-fit" between an agent's intra-subjective reasons and the objective reasons at which they're directed or which they represent: an agent's intra-subjective reasons constitute successful practical understanding or knowledge of the relevant objective reasons. An agent practically *mis*understands her objective reasons when her subjective reasons fail to practically-fit the objective reasons they represent, or with respect to which they direct an agent. Failure of practical-fit is a normative notion: intra-subjective reasons are supposed to, or function to, practically-fit the objective reasons at which they direct an agent. Let us consider some examples:

<div align="center">

Case A: Practical-Fit

</div>

Objective Reasons: Fact of need for food and fact of available food.

<Elicits:> ↓   ↓

Subjective Reasons: Motivation to eat and targeting food for eating (activity aimed at satisfying need for food).

<div align="center">

Case B: Need-Failure of Practical-Fit

</div>

Objective Reasons: Fact of *need for water* and fact of available food.

<Elicits:> ↓   ↓

Subjective Reasons: *Motivation to eat* and targeting food for eating (activity aimed at satisfying need for food).

Case C: Target-Failure of Practical Fit

Objective Reasons: Fact of need for food and *fact of available non-food object.*

<Elicits:>↓   ↓

Subjective Reasons: Motivation to eat and *targeting of non-food object for eating (activity aimed at satisfying need for food).*

Case A is an example of a good case-- there's a practical- fit between the agent's intra-subjective reasons and the objective reasons at which they're directed in that the agent's intra-subjective reasons prepare her to produce an intentional action that, if executed, would be appropriate: eating food is an action that exploits the properties of food to satisfy the agent's need for food.

In case B, there's a failure of practical-fit between the type of need the agent has (a need for water) and the action she's motivated (and in this sense prepared) to do. Had the objective reason been the need for food, there would not be a failure of practical-fit. Likewise, had the intra-subjective reason been the motivation to drink, there would not be a failure of practical-fit.

In case C there's likewise a failure of practical-fit, this time between the objective reason that is the fact of the available non-food object and the intra-subjective reason that targets that non-food object for the action or activity of eating. The failure is explained by the fact that the non-food item cannot be successfully targeted for eating-- ingesting it will not satiate the agent's need for food.

We may say, then, that intra-subjective reasons misrepresent when they fail to practically-fit the objective reasons at which they direct an agent. And there's a failure of practical-fit when an agent's intra-subjective reasons dispose (position) her to respond to objective reasons inappropriately; if an agent responds appropriately, she does so accidentally. Where an agent has successful intra-subjective access to her objective reasons, i.e., where she "understands" or has a successful "hold" or "grasp" of them, there's a practical-fit between the actions to which she's disposed on account of her intra-subjective reasons and the objective reasons at which those actions are directed. There's a practical-fit when an agent's intra-subjective reasons prepare her to produce *appropriate* intentional actions in response to the objective reasons at which they direct her; if she executes the action, she'll have acted appropriately, and it won't be by accident.[125]

The way in which intra- subjective reasons prepare an agent to produce an intentional response to objective reasons may be compared to the way knowing how to execute an instruction prepares an agent to do something. Consider, for instance, instructions for making a cake. If they are good instructions, they are appropriate instructions: they'll tell you to do what you need to do to make a good cake. Imagine that you have internalized these instructions-- you know how to make a cake by means of this recipe. Having these instructions doesn't mean that you will in fact make the cake-- knowing how to make it doesn't mean you have all the necessary ingredients, for instance. Nor does it guarantee that your oven will work. Knowing-how (or knowing what to do) just means that if the materials are present, you're in a position to interact with them so as to non-

---

[125] Importantly, misunderstanding or misapprehension, i.e. misrepresentation via an intra-subjective reason, is not the same as no understanding, apprehension, or representation. For an agent still apprehends a target or need insofar as her intra-subjective reason prepares her to act intentionally in a way that's directed at the satisfaction of need by means of a type of interaction with a target. To this extent, an agent is still intentionally directed at needs and targets. In contrast, an objective reason that figures neither as need nor target of an agent's intra-subjective reasons (not even indirectly, via a stand-in), is an objective reason of which an agent has no grasp, understanding, or representation. When this happens, an agent doesn't misrepresent objective reasons, rather she doesn't represent them at all.

accidentally make a good cake. Intra- subjective reasons, like instructions, practically-fit objective reasons insofar as they are potentials for appropriate responses to them; an instruction is *actually* appropriate if it is a *potential* for an appropriate action. Accordingly, we may think of practical-fit for intra-subjective reasons as consisting in *instructive appropriateness*. The conditions under which an intra- subjective reason succeeds (the good case) are thus the conditions in which it is instructively appropriate (henceforth "appropriateness conditions" for short).

Of course, the analogy with instructions isn't perfect: the vehicles of intra-subjective reasons are internal to agents, and agents act-out-of such reasons, whereas the vehicles of what we call recipes and instructions are external to agents-- they're presented to agents as subject-matters and thus seem to be a variety of inter-subjective reasons. If we're to understand intra-subjective reasons as instructions, we must understand them as a kind of fundamental interpretation-- they must be instructions *through which* an agent represents objective reasons and *out-of-which* an agent acts rather than instructions that are presented to her as a kind of subject-matter *on the basis of which* she might act.

Function

Many philosophers hold that representation is a functional kind and that this function is what grounds representational success or satisfaction conditions.[126] I've just suggested that the success of intra-subjective reasons has essentially to do with practical-fit and that they're successful in something like the way an instruction is successful, i.e., by preparing appropriate actions rather than by necessarily having those actions be actualized. Accordingly, I suggested that

---

[126] Millikan (1984), Ramsey (2007), Pierce (1883, 1977, 1998), Haugeland (1991), Cummins (1996).

intra-subjective reasons have appropriateness conditions rather than truth, veridicality, or satisfaction conditions. So, the question I now want to take up is whether we can articulate a function that's plausibly the function of a representation that would ground appropriateness conditions. I believe we can.

In the discussion on Aboutness above, I proposed that both inter- and intra-subjective reasons are about objective reasons (facts) by virtue of providing agents with PEA to them. To have PEA to an objective reason is to be in a position to produce an intentional action in response to it, where that intentional response is supposed to be appropriate. I suggested that intra-subjective reasons do this in a direct way: they directly prepare (determine and/or guide) intentional actions, while inter-subjective reasons do this in an indirect way by standing-in-for objective reasons in ways that enable agents to intentionally act on them in certain ways. So, I propose that subjective reasons have constitutive PEA functions where this means that they function to enable agents to produce appropriate intentional actions in response to objective reasons. And I propose that there are two varieties or species of PEA function, one of which is direct and grounds intra-subjective reasons, the other of which is indirect and grounds inter-subjective reasons.

Generally, then, to have practical-epistemic access (PEA) to an objective reason O is to have the ability (and in this sense be in a position or primed) to produce intentional action that aims to be practically-fit in response to O. Accordingly, something that has a PEA function functions to be an ability (to prepare, prime, guide, or position an agent) to produce an appropriate action in response to the relevant objective reasons (facts), or to be a projection or extension of such an ability.

PEA functions are intentional action *enabling or potentiating* functions; something with a PEA function functions to itself position an agent to produce a practically-fitting type of intentional

174

response to a type of objective reason. For I am assuming that intentional actions are themselves the sorts of things that are essentially supposed to practically-fit the objective reasons to which they respond (see Chapter 2). So, something that functions to enable or prepare an intentional action functions to enable or prepare a practically-fitting (i.e., appropriate) response. In other words, I propose that subjective reasons are constituted by PEA functions. And this is to propose that subjective reasons essentially function to provide agents with PEA to their objective reasons (facts). And this, in turn, is just to propose that subjective reasons essentially function to position agents to intentionally respond to objective reasons where being in a position to intentionally respond to objective reasons entails being in a position to produce actions that are *supposed* (or that themselves in some sense function) to practically-fit the objective reasons at which they're directed.

It follows that, on the view I propose, a subjective reason is essentially something that's supposed to (in that it functions to) practically-fit objective reasons, where a subjective reason R practically fits an objective reason O iff R positions an agent to produce a (non-accidentally) practically-fitting response to O. Accordingly, PEA-functions accommodate the earlier observation that an agent misunderstands the facts that are her objective reasons when there's a failure of practical-fit between her objective and subjective reasons (as in examples B and C above). When there's a failure of practical-fit, an agent is directed at her objective reasons insofar as she is positioned (prepared, primed) to produce an *inappropriate* action in response to them. When subjective reasons fail to practically-fit the objective reasons at which they direct an agent, this constitutes a practical-epistemic (normative) failure: a subjective reason fails to achieve its PEA function. This doesn't mean it fails to provide PEA, only that it fails to provide *successful* PEA, for successful PEA is being prepared to respond appropriately, or in a practically-fitting way.

175

An agent has PEA any time she's prepared to respond in a way that's *supposed* to be appropriate and therefore any time something that functions to provide PEA positions her to respond to objective reasons even if she's positioned to respond badly. So, PEA accounts for the way in which subjective reasons constitute a kind of understanding of objective reasons that's essentially assessable in terms of practical-fit.

4.4 The PEA Analysis of Subjective Reasons & Representations

Step 1 of the neglected strategy was to identify the genus *Subjective Reason* with the genus *Representation*. Step 2 was to develop a theory of subjective reasons-- in particular, intra-subjective reasons-- that could double as a theory of representations-- in particular, original representations. The last section worked towards such a theory, the results of which may be summarized as follows:

> **PEA Analysis of Subjective Reasons/Representations:** X belongs to the genus subjective reason (representation) iff X has a PEA-function, where X has a PEA- function iff X functions to position (to itself enable) an agent to produce intentional actions that are supposed to practically-fit the objective reasons to which they are responses.

The previous section proposed that there are two species of PEA-Functions corresponding to the two species of subjective reasons-- intra-subjective and inter-subjective reasons-- which I earlier mapped to original and derived representations respectively. To make this proposal more precise, and to demonstrate that the PEA-analysis of the genus subjective reason/ representation can accommodate these species, I propose the following distinction between "direct" and "indirect" PEA-functions:

Direct PEA-Functions: X has a direct PEA-function iff X functions to directly position an agent to produce an intentional action in response to objective reasons, where an agent is directly positioned to produce an intentional action in response to objective reasons iff she has an ability to produce (manifest) intentional actions in *direct* response (or that are direct responses) to objective reasons.

Indirect PEA-Function: X has an indirect PEA-function iff X functions to indirectly position an agent to produce an intentional action in response to objective reasons, where an agent is indirectly positioned to produce intentional actions in response to objective reasons iff she has an ability to intentionally act on and thus respond to objective reasons indirectly, i.e., by means of directly responding to something else.

Direct PEA-Functions constitute intra-subjective reasons/original representations. Something with a direct PEA-function (i.e., an intra-subjective reason) functions to enable or position an agent to produce a type of solution in *direct* response to an objective reason. A response is direct when its immediate target is the objective reason it's supposed to practically-fit. In the previous section I likened the way in which an intra-subjective reason is supposed to practically-fit objective reasons to the success of an instruction or recipe, only one that is in some sense "internal" to one's fundamental perspective.

I now propose that direct PEA-functions ground what Chapter 1 called *de actu* contents, or what Chapter 2 called *intentional action acorns*, or what I'll now call *Embodied Instructive Representations*. Embodied instructive representations are realized by *Instructive Dispositions*--composite dispositions or capacities whose components are direct-PEA functionally unified. Instructive dispositions consist in an agent's capacity to be sensitive to or affected by needs and targets (objective reasons) in such a way that an agent is primed or prepared to generate a particular intentional action teleologically/causally directed at generating a type of interaction with respect to the target that is teleologically/causally directed at satisfying an associated need. The vehicles of embodied instructive dispositions are internal to agents (though what activates them is often

external); they correspond to an agent's sensory-motor "wiring." I'll provide a detailed account of such representations and dispositions in the next two chapters (Chapters 4 & 5).

Indirect-PEA Functions ground inter-subjective reasons/ derived representations, or what I'll now call *surrogative representations* since they stand-in-for their objects (i.e., the things they represent). Something with an indirect-PEA function functions to enable or position an agent to indirectly intentionally respond to an objective reason. Indirect PEA- functions are mediator-functions; something with an indirect-PEA function functions to enable an agent to respond to objective reasons indirectly. A response is *indirect* when its immediate target is not the objective reason itself but an agent-external vehicle (concrete, sensible entity, or internalization thereof) that takes the place of (stands-in-for) the objective reason.[127] Paradigmatically, surrogative representations are realized by agent-external vehicles that come by their indirect-PEA functions via the intentional actions of human agents.

While I have identified intra-subjective reasons/original representations with embodied instructive representations, where these are paradigmatically mental representations, it should be noted that some mental representations might have indirect-PEA functions. In particular, mental representations that we relate to third-personally (as objects, or reflectively), such as, perhaps, mental imagery and mental maps, may be constituted by private PEA functional vehicles. However, such representations cannot be original (fundamental) in that not only are they not necessary-- it is possible to interact with the world directly just by means of vehicles with direct-PEA functions-- but they would also be useless to one who lacked the abilities, i.e., agent-internal vehicles with direct-PEA functions, i.e., embodied instructive representations, required to

---

[127] Where an objective reason may be the fact of an agent having certain intra- or inter- subjective reasons.

intentionally interact with the private indirect-PEA functional vehicles. This is because mental representations that are private indirect-PEA functional vehicles are, as such, third- (or second-) personal representations; they are representations on which one has a first-personal perspective or point of view, even if one is the only agent who can directly view them.

So far, I've proposed that direct-PEA- functions ground intra-subjective reasons/original representations while indirect-PEA functions ground inter-subjective reasons/ derived representations. I suggested above that the two species of subjective reasons are marked by different forms of practical-fitness or appropriateness. If so, then their corresponding species of representation should be expected to have different forms of representational success conditions. Embodied instructive representations, i.e., original representations, share their appropriateness conditions with the actions they instruct. But what about surrogative representations?

Descriptions, models, maps, and depictions or illustrations are paradigmatic representations. And it does seem as if it is essential to the success of these sorts of representations that they be "true of" or sufficiently similar to, or that they must resemble or correspond or map-to, the things that they stand-in-for, though what being sufficiently "true of" etc. comes to plausibly depends on the enterprises in which instances of such representations have their home.[128] And names and labels plausibly enable sentences to stand-in-for facts involving the things those names and labels name and label but can only do so if the names and labels trigger the right associations where this requires that they be used "correctly" or "truly" in the sense of "in accordance with a rule" that's been established by some social convention(s).

---

[128] This is not an entirely uncontentious claim. Some philosophers of science argue that scientific models are not properly truth-evaluable. See e.g., Cartwright (1980). In other work, I'm developing an account of scientific representation that I think can salvage a significant sense of truth evaluability for scientific models while accommodating their more pragmatic features on which Cartwright is right to place emphasis.

So alethic conditions seem to figure into the success of surrogative representations, i.e., representations that represent by standing-in-for facts. A description doesn't succeed, *qua* the type of representation it is, unless it is "true" (or true enough in the context) to what it describes. Whether an indirect response practically-fits an objective reason depends on whether the indirect-PEA functional vehicle is a successful surrogate (stand-in or mediator). In other words, it depends on whether the vehicle enables the relevant practically-fit interaction with the relevant objective reason. If this is right, then plausibly, what it is for a surrogative representation to succeed, *qua* representation, may be for it to be "true to" or "accurate with respect of" the target T for which it stands-in, meaning that it must functionally correspond to, map, or resemble certain of T's properties. To wit, in order to achieve an indirect-PEA function, a surrogative representation/inter-subjective reason (proposition, stand-in) must take the place of an objective reason (which might itself be the fact that someone has some intra-subjective or inter-subjective reason-- after all, these are facts too) in such a way as to enable an agent to produce appropriate intentional actions involving the thing it stands-in-for. To fulfill this function, a stand-in must itself have certain properties. At least in some cases, it may be necessary that there's a particular mapping or morphism between the stand-in and the objective reason it stands-in for, or a particular resemblance, correspondence, or similarity.[129, 130] Of course, being true to what it describes isn't what makes it about that thing-- otherwise it wouldn't be possible to misdescribe or mis-

---

[129] In some cases, practical correspondence supervenes largely on contingent (intersubjective) conventional facts, in which case the sense in which they are practically true or not is likewise contingent. In such cases, practical truth has a more deflationary-pragmatic flavor. In other cases, practical correspondence supervenes on non-conventional ("natural") facts. In such cases, practical truth has a more robust, less deflationary flavor.

[130] It should be noted that the fact of the production of a particular surrogative representation (inter-subjective reason) can itself be an objective reason, and instructive representations may instruct agents to produce particular surrogative representations.

descriptively-represent something.[131] I've proposed that what makes a surrogate representation a representation of something is its indirect-PEA function. But I believe that it may be possible to ground alethic representational success conditions in the special type of appropriateness condition that's associated with an indirect-PEA function.

Accordingly, I propose that alethic standards like truth, veridicality, accuracy, and satisfaction might be defined in terms of the properties a stand-in must have to fulfill its indirect-PEA function. That is, I think it may be possible to understand alethic success conditions as an aspect of a species of practical-fitness (or appropriateness). So, to say that a proposition (or some other descriptive content) is true or veridical may be to say that it fulfills its indirect-PEA function, and to say that it's false or falsidical is to say that it doesn't. Likewise for satisfaction. Accuracy may provide more information: to say that something is accurate is to say that it's got all the properties needed to fulfill its indirect-PEA function or functions (one stand-in vehicle might have multiple indirect-PEA functions, i.e. it might constitute multiple inter-subjective reasons). To say something is inaccurate to some degree is to say that it fulfills its indirect-PEA function but only to some degree-- i.e., not perfectly-- or it may be to say that it fulfills only a subset of the indirect-PEA functions it was assigned. On this proposal, then, the meaning of "true" "accurate" etc. cannot be given independently of specifying the operative indirect-PEA functions.

Consider an example: a child may draw a picture of her family that's just a bunch of scribbles that bear no visible resemblance to her family. This example shows two things. First, as we already observed, accurate description isn't what grounds representation. In this case the

---

[131] I take it that, at root, descriptions are descriptions OF some independent facts, and they are more or less successful insofar as they successfully render those facts (for indirect-PEA purposes). So if there's nothing a description is aimed at, it doesn't have a truth condition. Aiming a description (or other kind of stand-in) might itself be a relatively indirect affair. For instance, theorists may direct descriptions as whatever it is that ultimately explains some observed phenomena.

description (depiction) is highly inaccurate, but this doesn't stop the picture from being about, i.e., representing, the child's family. As I noted above, if accuracy were required to establish reference (indication, denotation) for surrogative representations, it wouldn't be possible to misrepresent. But this example of the child's picture reveals something about the way in which alethic conditions figure into the success of representations that stand-in-for what they represent. For despite failing to resemble her family, the picture is not obviously even a misrepresentation--- that is, it's not obviously a failure.[132] That's because there's much more to visual artistic representation than realistic rendering. On some views, the purpose of art is not to reproduce the visual world, but to create or express a different kind of perspective on or experience of it. So, a representation that stands-in-for X need not function to copy X. It may, rather, function to present X in a wholly new or different way.[133] Such a representation does not succeed by virtue of being true to how X is already or how X already appears---by providing a map to the mundane reality of X, as it were-- but rather by providing a map to how X *could be*, or by providing a *novel form of access* to and thus a different perspective on X or *enabling* agents to appreciate X under a different aspect.

More generally, it seems that a stand-in/surrogative representation's primary function is precisely to give agents a kind of PEA to what it stands-in-for, and alethic representational success conditions attach to the properties of the representational vehicle that enable whatever constitutes

---

[132] In a somewhat similar vein, Gabriel Greenburg (2013) argues that resemblance isn't necessary for accurate pictorial representation. (Thanks to James Shaw for the tip!)

[133] The child's drawing might also be produced as a part of her practice of learning how to produce visual stand-ins-- e.g. developing the necessary hand-eye coordination. We don't expect a child to produce realistic renderings when she's just learning how to draw, and it seems ridiculous to deem her drawings failures precisely because the point of producing them is not to already produce a realistic stand-in but to learn how to do so-- in the context of such practice, the stand-ins are instrumental. Of course, if one's taking a realist drawing class and one's drawings are not improving, the fact that one's drawings are not improving shows that one is not succeeding at learning.

successful access. But just what those properties are depends not only on the nature of the X the stand-in represents, but also our intentions in accessing X, as the following examples show:

Artistic Flower Picture: One may derive pleasure from looking at flowers in spring, and so produce or purchase a painting that provides one with access to their beauty even in the winter. Here, the relevant kind or purpose of access is aesthetic. Perhaps this purpose is only achieved if the painting is realistic-- perhaps one would prefer a photograph. Or perhaps one's aesthetic ends would be better met but a less realistic depiction-- perhaps one with softer edges, and slightly louder colors.

City Map: We create a map of a city which enables us to have a sort of birds-eye view access to it, which, in turn, makes it easier for us to find our way about it-- to get where we want to go. The map needs to be "true" to the city-- needs to picture or have corresponding parts with corresponding relative locations and dimensions (consistently scaled-down) if it's going to enable us to navigate the city.

Records: We create pictures and records to provide us with access to past events and facts. Records may be used to prove or provide evidence that something happened in the past, or to remind us of something we don't want to forget, or to provide future agents with access to something that's important to us, and for lots of other purposes. A photo of a crime scene that's to serve as evidence in a case needs to give us access to the crime scene that approximates direct (and clear) visual access to it as closely as possible. A biography

gives agents access to a person's life-- the story they might tell about them-self could they tell it directly-- even when the person is no longer living.

Scientific Model: A scientific model needs to be "true to" some causal structure that's instanced by some class of phenomena. Plausibly, being true-- corresponding in such a way as to provide access to and what we might think of as objective knowledge of the causal structure of this part of the world--is the objective of the model. But the evidence of its truth, and perhaps the substance of the sense in which it constitutes knowledge for us, is in its capacity to enable agents to predict and control (intentionally intervene on) instances of the relevant class of phenomena. In addition, such a model may be unlike any actual instance of the class of phenomena whose shared structure it functions to stand-in-for because it can provide access to such general structure only by abstracting away from particulars, i.e., by idealizing.

While all of these examples describe representations that stand-in-for things and may be appropriately assessed in terms of alethic standards (truth, matching, similarity, resemblance, pictorial correspondence, descriptive accuracy, etc.), the content or matter of the alethic standard varies according to the purpose of the stand-in-- the point of the indirect access it's supposed to afford, and the kind of thing to which it's supposed to afford such access. The photograph of the crime needs to stand-in-for a *particular* event, the objects it involved, their arrangements, etc. Such a representation succeeds by virtue of capturing (precisely corresponding to) particular details of an actual event. In contrast, the scientific model needs to stand-in for a *general* structural feature of some class of phenomenon that's realized somewhat differently by every actual member of that

184

class. Such a representation succeeds by abstracting away from the particular details of actual events. It needs to correspond to particular instances in the sense that it needs to enable us to predict certain things about them and successfully intervene on them in controlled ways. The picture of flowers may succeed by exaggerating certain details while minimizing or completely omitting others. It needs to take the place of the flowers in the sense that it needs to provide an agent with a kind of experience of or affective response to them. A painting might have the function of making salient-- providing an unusual kind of visual access to-- the particularities of a certain actual flower precisely by abstracting away from (omitting or minimizing) certain other of its actual features. A botanist's picture of the very same flower would have rather different pictorial properties and would in an important sense be subject to a different sort of standard of success. But it's hardly clear that only the botanist's picture is true or assessable in terms of truth, nor that it's "more true" of the flower than the artist's. They seem, rather, to be true of it in different ways corresponding to the different kinds of access they afford where the nature of the access they afford depends, in this case, on whether one's ends are empirical or aesthetic.

I have proposed that what representations that stand-in-for what they represent have in common is that they function to provide what we might call a *proximity-independent form of intentional access*-- i.e. indirect-PEA; PEA that doesn't depend on the agent's actual proximity (in time, space, etc.) to the things they stand-in-for.[134] For instance, a map affords visual access to the layout of a city that isn't otherwise available to one from the ground, and a record provides intellectual access to events or entities in the past, etc. Stand-ins (or surrogative representations)

---

[134] In Chapter 4 I'll suggest that they function to solve "Proximity Problems"-- a special, typically derivative kind of "practical problem" (need for flourishing-- see Chapter 2). In other words, indirect-PEA consists in the potentiations of application-accountable solutions to proximity problems.

take the place of the things they stand-in-for, for certain purposes (though those purposes need not all be decided in advance). To the extent that stand-ins succeed in taking-the-place-of the things they stand-in-for, they're true of or to them, and their vehicles have the relevant properties that are required for them to serve as surrogates for the things they stand-in-for, for the relevant purposes. For instance, the sound of a name must have the right associations for agents brought up in a particular culture, the scientific model must have the right number of parts connected at the right places, the painting must have the right balance of colors, etc. These observations suggest that the determinate content of "true" "accurate" etc. as applied to surrogative representations cannot be given independently of specifying their constitutive indirect-PEA functions.

4.5 Conclusion

Philosophers who adopt the descriptive thesis treat alethic success conditions as the essential feature of representations. The standard strategy for illuminating original intentionality (psychological subjective reasons) adopts the descriptive thesis. Accordingly, it treats all species of intentionality or representation on the model of propositions. The purpose of this chapter was to motivate and begin to develop a new paradigm that inverts the standard order of explanation by trying to understand representations in terms of an analysis of subjective reasons.

In order to show that the proposed paradigm shift isn't a non-starter, I addressed *prima facie* objections one might have to understanding representations in terms of subjective reasons. In doing so, I argued for the existence of both "private" and "shareable" subjective reasons, and I suggested that their mutual existence presents a sort of puzzle about how to relate them. I explained

that the propositional attitude model of subjective *qua* private or mental reasons may be understood as a strategy for solving this puzzle. On that strategy, the "private" nature of subjective reasons falls to the attitudes agents take towards what are essentially public reasons or interpretations of facts or else to "private" or "internal" propositions or proposition-like representations. I argued that this strategy faces serious challenges when it comes to illuminating the nature of original representations/ intra-subjective reasons (the "private" or "unshareable" species).[135] I therefore proposed to explore a neglected strategy for dealing with this puzzle which posits that there are two species of subjective reasons that correspond to two species of representation. I pursued this strategy by developing the PEA analysis of the genus subjective reason/representation and I analyzed the two species of subjective reasons/ representations in terms of direct- and indirect-PEA functions.

On the PEA analysis, the mental, non-sharable subjective reasons out of which we act are intra-subjective reasons.[136] Intra- subjective reasons are original representations, which have the form of embodied instructive representations as opposed to propositional attitudes. Embodied instructive representations are grounded in direct-PEA functions. Public or shareable representations are inter-subjective reasons on the basis of which we act. Inter-subjective reasons are surrogative representations which are grounded in indirect-PEA functions and realized by agent-external (but perhaps internalized) vehicles. Such vehicles, such as sentences, are created and used (uttered, written, thought) to indirectly position agents to produce intentional actions in response to objective reasons. Surrogative representations can only be representations for agents

---

[135] I also addressed similar problems in Chapter 1 and I'll address some more in Chapter 5.

[136] These are not sharable in the sense that they are embodied in particular agents. Of course, similar types of agents can have similar types of intra-subjective reasons. Indeed, creating overlapping intra-subjective reasons is part of the point of objectifying and sharing reasons, i.e., one of the functions of inter-subjective reasons.

who possess (and can deploy) the knowledge-how that's involved in employing them as mediators (surrogates), and they can be shared by multiple agents. On the other hand, because surrogative representations can stand-in for and objectify intra- subjective reasons they enable agents to intentionally modify their fundamental practical understanding of objective reasons.

Consequently, the PEA analysis accommodates the observation that two agents can represent the fact (objective reason) that it is raining via a descriptive surrogative representation of a state of affairs, and both act on the basis of that shared inter-subjective reason. Inter- subjective reasons-- shared, surrogative reasons-- are third-personal (objectified) purpose-defined perspectives of what they stand-in-for. Their vehicles are external to agents. Rather than being attached to individual agents, they have detached public meanings or significances-- they can represent objective reasons, i.e., facts, including the fact of having certain intra-subjective reasons or certain other inter- subjective reasons as such. But surrogative representations can only represent facts as such for agents who can take them to represent facts as such-- i.e., for agents who can have inter- subjective reasons.[137]

In addition to resolving the "puzzle," the PEA analysis is a unifying analysis. First, it unifies Subjective Reasons and Representations by defining a common genus in terms of the notion of a PEA function. Second, it unifies the species Intra-Subjective Reason with the species Original Representation, and it also unifies the species Inter-Subjective Reason with the species Derived

---

[137] The sentence "it is raining" is a surrogative representation for one who can respond to the sentence as if they had themselves seen that it is raining, or heard the rain falling-- as if they had had a direct encounter with the rain. But this is not, or at least not obviously, enough for the sentence to represent facts *as such*. Indeed, it seems possible to train a non-human animal to respond to the sound of the sentence "it is raining" by taking cover. I'm inclined to think that representing facts or objective reasons *as such* requires abilities (and the relevant needs to ground them) to give and ask for reasons as well as the ability to surrogatively represent oneself as an agent with a point of view (subjective reasons) in relation to objective reasons (facts) on which other agents may have other points of view. Agents that have these kinds of abilities count as having theoretical (intersubjective and objective) knowledge.

Representation. Finally, the PEA analysis also promises to be an illuminating analysis. Where propositional attitudes obscure the nature of original representations/ intra-subjective reasons by construing them as "private" or "internal" versions of derived representations/ inter-subjective reasons, I've proposed to analyze original representations in terms of direct-PEA functions. If I'm right that direct-PEA functions ground embodied instructive representations with constitutive appropriateness rather than alethic conditions, then a "surprising" result of my analysis-- though not so surprising if one has already shaken oneself free of the grip of the Descriptive Thesis--- is that alethic conditions are not an essential feature, let alone the essential feature, of representation. On the other hand, the PEA analysis is compatible with attributing alethic conditions to surrogative representations and may even illuminate notions like "truth," "veridicality," "accuracy," and "satisfaction" conditions by grounding them in the vehicular properties in virtue of which surrogative representations achieve their indirect PEA-functions.[138]

---

[138] One might object that the PEA (functional) analysis isn't much of an analysis precisely because it relies on the notion of a practical-epistemic access function. An unanalyzed notion of "practical-epistemic access" would indeed be problematic, since one might think that epistemic access just is representational access, and that anyway, it's just as mysterious as representation. However, I have proposed an analysis of PEA in terms of abilities to produce intentional actions in response to objective reasons or facts. In light of the problem-solving account in Chapter 2, we needn't worry about circularity.

5.0 Chapter 4: Embodied Instructive Representations & Problem-Solving Dispositions

This chapter has four objectives. The objective of §1 is to put some of the pieces together from the foregoing chapters. Specifically, I'll identify the notion of *de actu* contents from Chapter 1 with the notion of an "intentional action acorn" from Chapter 2 and with the notion of an embodied instructive representation (EIR) and hence intra-subjective reason (reason *out of which*) from Chapter 3. I'll also show that the definition of representation as PEA from Chapter 3 can be chained with the definition of intentional action as an application-accountable practical problem from Chapter 2. This will yield a unified account of intentionality which I'll have motivated from multiple angles in the preceding chapters.

The objective of §2 is to show that the notion of an intentional action acorn from Chapter 2 can be cashed-out non-metaphorically in terms of an account of the "problem-solving" dispositions that realize EIRs. Problem-solving dispositions will also start to put some skin on the bones of the conception of practical judgement that acts on *de actu* contents at which I gestured in Chapter 1. The objective of §3 is to illuminate the nature of EIR content-- what it means to be a *de actu* content or generative mode of presentation. Finally, the objective of §4 is to address the concern that EIRs are not really representations, and to clarify the kind of representations they are.

5.1 Putting the Pieces Together

Chapter 1 argued that, fundamentally, the intentional relation might consist in a representational relation involving *de actu* contents. Chapter 2 introduced the practical problem-solving analysis of intentional action. According to that analysis, intentional actions are application-accountable solutions to practical problems. That chapter also suggested that psychological representations may be understood as "intentional action acorns." Chapter 3 argued for identifying the genus "Subjective Reasons" with the genus "Representation," for distinguishing between intra- and inter-subjective reasons and identifying these with original and derived representations respectively, and for analyzing subjective reasons/ representations in terms of the notion of practical epistemic access (PEA) where embodied instructive representations (EIRs) provide direct-PEA while surrogative representations (SRs) provide indirect-PEA. In this section I'll show how these pieces may be put together to yield a unified and informative theory of representation.

I'll begin by connecting the problem-solving analysis of intentional action from chapter 2 with the PEA analysis of representation from Chapter 3 by chaining the proposed definitions (§1.1). I'll then consider what light the chained definition, together with the metaphor of an intentional action acorn from Chapter 2, can throw on the two species of PEA and corresponding species of representations identified in Chapter 3 (§1.2 & §1.3).

5.1.1 PEA & Problem-Solving

In Chapter 3 I defined PEA in terms of abilities for intentional actions in response to objective reasons (facts). I'll now show that this definition can be chained with the definition of intentional action from Chapter 2. Let's begin with a review of that definition:

> **Def. Intentional Action:** A is an intentional action iff A is a token application-accountable solution to a practical problem.

Unpacking this definition: A **practical problem** is a type of "need for flourishing," where such needs are objective facts about life forms. Examples include the need for nourishment, hydration, protection, shelter, friendship, etc. A **solution** is a type of process of interaction between a life-form and a type of target (or type of combination of targets) that is generated by the life-form and that is a (non-accidental) mechanism for satisfying a type of need for flourishing (i.e., solving a type of practical problem). Examples include anything from eating, drinking and evading predators to making friends to producing objects such as paintings and shoes.

There are different kinds of practical problems and corresponding solutions: A **local solution** solves a type of local problem; it answers: *What do I do with respect to some local problem P?* Every token local solution is a candidate global solution. A **global solution** solves a type of global problem by selecting the right local solutions at the right times; it answers: *Which of some set of the local problems I'm facing should I try to solve now to promote my overall flourishing?* or: *Which of these available local solutions should I prioritize, i.e., perform (now)?* Finally, there are also sub-solutions (and sub-sub-...+ solutions), which solve sub-+problems—instrumental problems that arise in the course of solving a larger practical problem. Roughly: a (super-)solution tells you *what to do;* sub⁺-solutions tell you *how to do it.*

Solutions (and therefore intentional actions) are fundamentally evaluable in terms of a norm of appropriateness. A local solution is appropriate iff, were it performed to completion, it would non-accidentally solve an instance of the type of local problem for which it is a solution. A local solution is inappropriate iff it fails to "practically fit" the context in which it is applied (potentiated), i.e., if it is applied outside its type's *domain of application* which consists of instances of the types of needs and targets which define the problem-type the solution-type it instantiates solves. A global solution is inappropriate iff it consists in the selection of a local solution that would not, were it performed to completion in the context, non-accidentally promote the subject's *overall* flourishing (and thus constitute an appropriate global solution in the context).

Importantly, the (in)appropriateness of a solution (intentional action) is not to be confused with the success or failure of its performance or execution (though the latter may redound on sub-solutions). After all, an appropriate action might be formed but never be performed, either at all or to completion (it may be interrupted, abandoned, conditions may not be favorable). And, of course, inappropriate actions can be performed perfectly to completion, but that doesn't make them appropriate.[139]

Finally, an intentional action is not merely a solution, but an application-accountable solution. Application-accountable solutions are sensitive to the appropriateness of their instances; their application algorithms (which will be described in more detail shortly) are disposed to be modified in response to whether or not the solution's tokens practically fit the situations to which they are applied.

---

[139] Appropriateness should also be distinguished from the standard of optimality; an appropriate solution need not be an optimal solution.

Since the practical problem-solving analysis defines intentional actions as solutions to practical problems, we may now define representation as that which is the potentiation of an application accountable solution (henceforth simply "solution") to a practical problem. First, recall the PEA analysis of representation:

Def. Representation: R is a representation of (about, directed at) a fact X iff R constitutes a form of PEA to X, where R constitutes a form of PEA to X iff, in having R, an agent is positioned to produce an intentional action in response to X.

Chaining this definition with the problem-solving definition of intentional action above, we get:

Chained Definition: R constitutes a form of PEA to X (and thus a representation of X) iff R positions an agent to produce an application-accountable solution in response to X.

Note that, since representations occur only on the left of the iff, the PEA analysis of representation constitutes a reductive analysis of representation in the sense that it reduces representations to abilities for application-accountable solutions to practical problems. This sort of analysis should satisfy fans of the rich correspondence view (see Intro and Chapter 2). Of course, such solutions were analyzed in terms of needs for flourishing, which are themselves intentionally directed. Hence, the PEA analysis is not the kind of reductive analysis of representation that reduces representation to something non-intentional. So, this analysis probably won't satisfy fans of the reductive view (see Intro and Chapter 2), but satisfying those folks isn't my objective.

5.1.2 Direct-PEA & Intentional Action Acorns

In chapter 3 I defined two species of subjective reasons, which I identified with two species of representations corresponding to two species of PEA. One of these species is direct-PEA:

> **Direct-PEA:** R provides an agent with direct-PEA to X iff in having R, an agent is positioned to produce an intentional action that is a direct response to X, where a response is direct when its immediate target is X.

Corresponding to direct-PEA we have *Embodied Instructive Representation*/ the intra-subjective reasons *out of which* we act. An embodied instructive representation (EIR) is a *generative* mode of presentation of one's own actions—the actions one is in a position to produce or in the process of producing-- and the facts at which those actions are directed. An EIR is a representation that stands-for an intentional action and the facts at which the action is directed by virtue of being an intentional action acorn for that very action. EIRs share appropriateness conditions with actions they generatively represent, and their *de actu* contents may be represented by the schema: "ϕ-THAT-THUS!"

Now, in light of the chained definition above, we may define original psychological representations, or intra-subjective reasons, as potentiations of direct solutions, i.e., solutions that are directly applied to a fact X. A potentiated direct solution means an agent is in a position to produce an intentional action that is a *direct* response to a fact X. We may also identify embodied instructive representations (EIRs), that is, representations which constitutively provide direct-PEA, i.e., intra-subjective reasons, with intentional action acorns.

In Chapter 2 I suggested that we might understand psychological representations in these sorts of terms. There I used the metaphor of intentional action acorns. As a reminder, here's the

basic idea: An intentional action, like the life of an individual oak tree, is a developing process. The oak tree begins its life as an acorn— the acorn is the embryonic phase of a particular oak tree; it contains the potentiality (the genetic instructions for) the essential features of the entire oak tree. The oak tree is there, as it were, from its conception. Likewise, an intentional action begins its life as an intentional action acorn—the embryonic form of an action that contains the potentiality of (instructions for) the essential features of the entire action. The intentional action is there, as it were, from its conception. And just as many acorns will never move beyond the acorn phase— they'll be eaten, rot, dry up, or begin to grow only to then be eaten, rot, or dry up— so many intentional action acorns do not develop beyond that initial phase; they are not selected to be performed or they might begin to be performed only to be abandoned. As noted in Chapter 2, by identifying psychological representations with intentional action acorns, we can accommodate the observation that mental representations like beliefs and desires can exist without the *performance* of the relevant actions. At the same time, we can hold onto the idea that the intentionality of action and the intentionality of psychological representation are essentially connected. Moreover, we can hold onto a genuine sense in which a mental representation causes— can figure into a causal explanation of— an intentional action. This is the same sense in which an acorn can figure into the causal explanation of the oak tree it grows into.

Going forward-- for the rest of this Chapter and the next-- I'll employ the metaphor of the intentional action acorn to illuminate EIRs. I'll also provide an account of EIRs which will, in turn, cash-out the metaphor of intentional action acorns. First, however, let us see what light the chained definitions can throw on the other species of PEA identified in Chapter 3, namely, indirect-PEA.

5.1.3 Indirect-PEA & Proximity Problems

This second major species of PEA identified in Chapter 3 is indirect-PEA:

> **Indirect-PEA:** R provides an agent with indirect PEA to X iff in having R, an agent is positioned to produce an intentional action that is an indirect response to X, where a response is indirect when its immediate target is something that takes the place of X-- something one acts on directly so as to act on X indirectly. [140]

Corresponding to indirect PEA, we have Surrogative Representations (SRs) or inter-subjective reasons *on the basis of which* agents act. SRs are *surrogative-generative* modes of presentations of facts. They stand-*in*-for facts thereby potentiating some indirect action involving those facts. SRs may also be understood as intentional action acorns: they too represent by virtue of potentiating an intentional action—this time an indirect intentional action. SRs don't stand-in-for actions. Rather, they stand -in-for facts and thereby potentiate indirect actions with respect to those facts.[141] SRs ride piggyback on EIRs because SRs can only constitute acorns for indirect intentional actions for agents who are in a position to use them to do so, i.e., for agents who have EIRs to produce and respond to SRs. So, where EIRs are fundamental or original, SRs are in this sense "derived." SRs also differ from EIRs with respect to their success conditions. Because SRs are genuine stand-ins, they can potentiate non-accidentally successful indirect actions with respect to their targets only if they bear certain sorts of structural (though they may be heavily conventional) correspondence relations to the facts for which it stands-in. Accordingly, I suggested in Chapter 3 that descriptive standards of success such as truth, veridicality, accuracy,

---

[140] Indirect PEA presupposes direct PEA: one must have direct PEA to the R that provides indirect PEA.

[141] Note that a single vehicle (e.g., a single diagram or sentence form) will often potentiate a variety of different indirect actions and therefore realize multiple SRs.

and satisfaction may be understood in terms of the properties an SR must have to constitute the potentiation of an appropriate indirect actions.

Now, in light of the chained definition of representation, we can conceptualize SRs, and intra-subjective reasons, in terms of the notion of an application-accountable solution to a practical problem. Roughly, we may think of EIRs as an agent's embodied problem-solving abilities and we can think of SRs as extensions of these abilities.[142] More precisely, indirect-PEA may be defined as a potentiated indirect solution that consists in an interaction with a vehicle Y that stands-in-for a target X and that constitutes an agent's being in a position to produce an intentional action that is an indirect response to a situation X; the solution is itself directly applied to Y, and the interaction with Y constitutes an indirect interaction with X.

We may also understand indirect-PEA, or SRs, more straightforwardly as a type of intentional action acorn by understanding SRs as potentiations of solutions to a special type of practical problem I call a "proximity problem."[143] Under some conditions, targets (facts) are "unavailable" for certain kinds of direct interactions. A particular kind of direct interaction may, for instance, be too dangerous or otherwise costly or difficult, or it may not be possible because the relevant targets are facts that obtained in the past but no longer, or facts that may come to obtain in the future but not yet, or facts that are mere logical possibilities, or facts that consist not

---

[142] The vehicles that realize EIRs are *problem-solving dispositions*—neurologically realized (though arguably more deeply embodied and embedded) functional capacities that I will describe in detail below and that constitute an agent's standing practical knowledge understood as their abilities to solve practical problems. Paradigmatically, the vehicles that realize surrogative representations are sensible media that are embedded in communities of users.

[143] EIR and SR acorns persist in a way that oak tree acorns do not, and they persist in different ways. EIR intentional action acorns persist in the form of DNA instruction which guides the development of the action as it is performed (if it is performed). The SR potentiates an indirect action in the sense that it is a target for an EIR that is the potentiation of that indirect action (solution to proximity problem) without which there could be no such EIR. It persists as an object separate from that action precisely because it contributes to solving a proximity problem by standing-in for a target (fact).

of particulars but of generalities, patterns, universals, etc. if such things can exist. When an agent has a practical problem whose solution requires a type of interaction with an unavailable target, she faces a proximity problem.[144] When an agent faces a proximity problem, she needs something to go proxy for, i.e., stand-in-for, an unavailable target, and interacting with the stand-in non-accidentally enables her to satisfy her overarching need. The solution to a proximity problem essentially involves the production or use of some item as a vehicle of indirect-PEA: the vehicle of indirect-PEA (the proxy in the context of the interaction) itself potentiates the indirect interaction that constitutes the larger solution. Since indirect-PEA solves a proximity problem, it constitutes a phase in a larger solution, as a proximity problem arises in the context of a practical problem that gives rise to the proximity problem, i.e., the need for a type of indirect interaction. Indirect-PEA is what solves the proximity problem: if it is appropriate, it is the potentiation of (part of) a solution to the larger practical problem in which the proximity problem it solves arose.[145] [146]

---

[144] Different varieties of proximity problems will produce different varieties of surrogative representations. I do not attempt a taxonomy of these varieties in this dissertation.

[145] The limiting case of this might be cases of empty reference, where the target is merely posited -- a merely possible target that turns out not to be real.

[146] A solution to a proximity problem is the potentiation of an action that is indirectly directed at a state of affairs X and consists in the production of a Surrogative Representation (SR), i.e., an indirect-PEA vehicle, that goes proxy for X, and with which one is in a position to interact directly and to thereby interact with X indirectly. The realizer of the SR that goes proxy for X is the *vehicle* of indirect-PEA: it is the thing with which one interacts directly and that potentiates an indirect interaction with a target X. Philosophers and cognitive scientists often refer to a representation's physical realizer as its "vehicle." According to the term's everyday meaning, a "vehicle" is a tool -- something that mediates some activity, in particular activities of getting from here to there. Thus, the notion of a vehicle makes sense in the context of indirect-PEA: indirect-PEA is a tool that extends the reach of our PEA -- we have direct-PEA to the surrogative vehicle (tool) and that tool is what *takes us to* the target to which it provides us, qua proxy, indirect-PEA. In other words, there's a real cognitive sense in which SRs get us from here to there. Most of the things people who aren't doing philosophy or cognitive science call vehicles require a user or driver. Likewise, indirect-PEA requires a user: an agent must have direct-PEA to the vehicle of indirect-PEA: she must produce a proxy/stand-in and then use it or use a proxy/stand-in that someone else produced to solve her proximity problems. In other words, an agent interacts with a vehicle of indirect-PEA as an object. In the context of direct-PEA, "vehicle" merely means realizer; this is an *extension* of what I take to be the word's everyday meaning. So extended, it does not carry its everyday sense.

An SR is appropriate iff the indirect solution it potentiates (the indirect action direct interaction with it makes possible) is appropriate.[147] By identifying derived representations with SRs, we understand derived representations as extended practical-problem-solving abilities: they serve as surrogates or stand-ins for facts in order to enable agents to *indirectly* solve the practical problems those facts (partially) define.

Whether an SR will potentiate an appropriate indirect action depends on whether it is an appropriate SR. The appropriateness of an SR is a matter of having the properties (i.e., functional similarity) required to solve the relevant proximity problem and thereby enable agents to produce the relevant indirect solutions involving the objective reasons for which they serve as surrogates. I suggested in Chapter 3 that standards of truth, accuracy, etc. may measure the presence/ absence of these properties. We may now put the idea this way: whether an SR is (in)appropriate with respect to its proximity-problem-solving function (indirect PEA-function) depends on whether it is true or accurate with respect to the objective reason for which it stands-in. That is, we may define an SR's alethic representational success conditions as the conditions under which a surrogative representation (i.e., a representation that provide indirect-PEA) solves the proximity problem it functions to solve at all (true/false standard) or to some degree (accuracy standard). [148]

Finally, defining SRs, i.e., derived representations, in terms of proximity problems helps bring out a reason they are less fundamental than EIRs, i.e. original representations. After all, proximity problems arise in the context of more basic (metaphysically speaking) practical problems, and they are essentially tools we use to extend our direct PEA- abilities; they enable us

---

[147] Recall that the relevant notion of appropriateness is an anti-luck notion: solutions are not solutions by luck.
[148] Chapter 3 also discussed some examples of different types of SRs. I do not attempt a complete taxonomy in this dissertation.

to intentionally respond to objective reasons indirectly. To be sure, we would not be able to solve certain kinds of practical problems without surrogative representations. So, there's a real sense in which they bring new abilities into existence. But they do so by extending already existing abilities, including the ability to interact with surrogates so as to solve proximity problems.

Now, one might object that sentences, photographs, and scientific models don't appear to have any essential relation to any practical activities. Afterall, there's no one thing they seem to "say." But this is a misleading appearance which stems from the fact that a single SR vehicle may actually be a solution to a multiplicity of proximity problems. To wit, the multiplicity of practical functions (solutions to proximity problems) that can be identified with a single SR vehicle gives rise to the illusion that SRs are not essentially practical because there's no *single* practical problem they function to solve. Reinforcing this illusion is the excessively narrow conception of the practical with which philosophers are often operating. Indeed, there's a tendency to reduce the "practical" to some minimal set of activities that are part of the survival of any animal. But it's not obvious what the motivation for this is. On the proposed view, a representation is essentially practical because it essentially solves a type of practical problem. Some practical problems may have more urgency than others when it comes to overall flourishing-- you need to eat, drink, and avoid predators if you're going to be able to dance and do philosophy. And going along with this, it may be necessary to satisfy certain "basic" needs before we can attend to others. But it doesn't follow from this that the "less basic" needs are derived from the more basic needs-- that dancing and doing philosophy ultimately emerge as solutions to problems of eating, drinking, and avoiding predators. Rather, human practical problems may be of a *fundamentally* intellectual or ethical nature and may involve the human need for things like freedom, friendship, and justice, where solutions to such problems may mark a distinctive species of practical understanding.

As I mentioned in the introductory chapter, there is much more to say about SRs. But this dissertation will focus more or less exclusively on developing an account of EIRs.

5.2 EIRs & Problem-Solving Dispositions

In order to cash-out the metaphor of an intentional action acorn, and to better understand the sense in which *de actu* contents are judgeable, and hence to better understand the nature of EIRs, this section will provide an account of the problem-solving dispositions that realize EIRs.

This section is very detailed, but the big picture is this: Problem-solving dispositions are abilities for solving practical problems. They're composed of functionally unified sub-capacities, namely, need- and target- input capacities and ϕ-output capacities, where causal inputs to the input sub-capacities activate the dispositions they compose (**Figure 1**). The function that unifies the sub-capacities into a problem-solving disposition is a type of problem-solving or solution function. The previous section reviewed the distinctions between local, global, and sub-problems and solutions from Chapter 2. Corresponding to these are local, global, and sub-solution functions, and these, in turn, individuate different kinds of problem-solving dispositions. "Instructive Dispositions (IDs)" have constitutive local solution-functions. They realize what I call "embodied instructive concepts" applications of which constitute EIRs. "Organizing Dispositions (ODs)" have constitutive global-solution functions and act on activated IDs, that is, EIRs, and hence de actu contents. Organizing dispositions select potentiated actions (EIRs/ *de actu* contents) for performance. Their activities constitute a fundamental form of practical judgement-- the form,

introduced in Chapter 1, which issues in actions directly. Finally, "Sub-Instructive Dispositions" (Sub-IDs) have constitutive sub-local solution functions.

In order to map the properties of problem-solving dispositions to the properties of embodied instructive representations, I'll use the following schema for representing embodied instructive representational (*de actu*) content:

"φ-THAT-THUS!"

On this schema, φ is a type of solution (action), e.g., eating, THAT is a target (e.g. a particular apple), and THUS is a manner or realization of φ which often involves various sub-acts (e.g., climbing an apple tree). "!" indicates the "force" or urgency of the instruction (this is a pragmatic feature of EIRs that ODs exploit in determining which to select as global solutions).

In what follows, I'll discuss the relationships between the three types of dispositions to illuminate the ways in which problem-solving dispositions constitute abilities for the three types of solutions to practical problems (§2.1-§2.4). In the subsequent section (i.e., §3) I'll have more to say about instructive dispositions *qua* embodied instructive concepts.[149]

5.2.1 Types of Problem-Solving Dispositions

Problem-solving dispositions are abilities or capacities to solve, or to apply solutions to, practical problems. They come in three types: ***Instructive Dispositions,*** i.e., abilities to solve local practical problems, ***Sub-Instructive Dispositions***, i.e., sub-abilities involved in solving local practical problems, and ***Organizing Dispositions***, i.e., abilities to solve global problems that

---

[149] Concepts or standing practical knowledge/ understanding, standing beliefs, or perhaps in some cases, prior intentions.

piggyback on instructive dispositions *qua* abilities to solve types of local problems. EIRs only exist where all of these problem-solving dispositions are present, but they are most immediately connected with instructive dispositions (IDs). IDs are capacities for types of local solutions. Their exercises (activations) function to solve tokens of the types of problems the solutions for which they are abilities solve. They achieve their function only if they are activated within the relevant solution type's domain of application. In other words, they achieve their function only when the solution type is appropriate. Unsuccessful exercises of IDs potentiate tokens of a type of solution in a context in which the relevant solution cannot be actualized, for solutions can only be fully actualized in their domains of application. Accordingly, inappropriate solutions are defective solutions: solutions that cannot be fully actualized.

The causal properties (sensitivities) of the input capacities of all types of problem-solving dispositions that causally determine their application conditions are "application algorithms" (see Chapter 2). As noted in Chapter 3, an important component of these abilities consists in an agent's "sensory wiring," i.e., the causally sensitive sub-capacities of instructive dispositions that are responsible for their input-activations. Abilities for application-accountable solutions, and hence problem-solving dispositions, also include mechanisms that auto-adjust their application algorithms according to whether they are succeeding at solving an agent's practical problems (see Chapter 2). In other words, problem-solving dispositions must include mechanisms that realize application-accountability. I'll now describe the properties of each type of problem-solving disposition in more detail, starting with Instructive Dispositions.

5.2.2 Instructive Dispositions

Instructive dispositions (**Figure 2**) are, first and foremost, a kind of active causal capacity: they function to enable agents to produce or generate intentional actions. So, they essentially involve an active causal component or sub-capacity; what I'll call a "ϕ-output sub-capacity." But intentional actions constitutively function to solve practical problems. So, the ϕ-output sub-capacity of any given instructive disposition must be modulated by two kinds of passive causal sub-capacities: a "need-input sub-capacity" that responds to instances of a type of need (local problem), and a "target-input sub-capacity" that responds to instances of the type(s) of targets that, through a type of interaction, satisfy those needs (**Figure 1**). The input causal capacities of IDs encompass what I earlier referred to as a solution's application algorithms. These sub-capacities can have quite complex causal structures that allow them to respond to need and target types in different contexts.
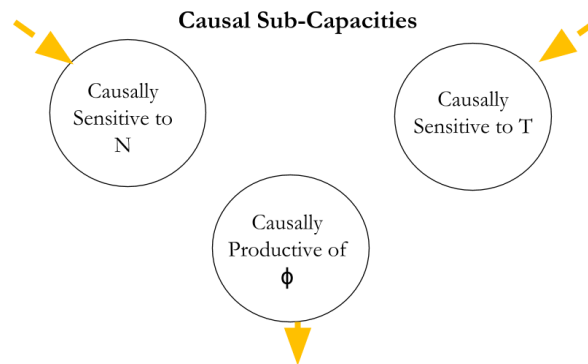
Causal Sub-Capacities

Figure 1 Causal Subcapacities of Problem-Solving Dispositions
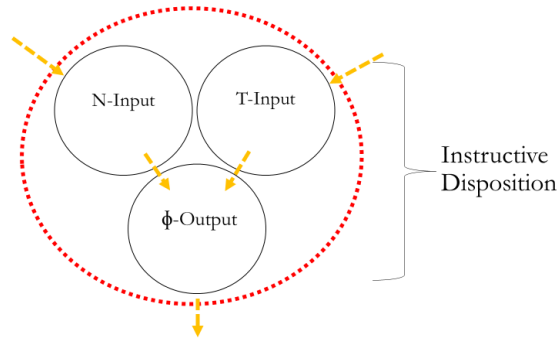
Figure 2 Instructive Disposition

Some detail should help to flesh-out the picture. First, a type of need (problem) is typically for a type of "terminal target" and for a type of interaction with that terminal target. For instance, if the need is nutrition, the terminal target is food (something edible), and the type of interaction is eating. So, an agent's "Eat" instructive disposition functions to respond to the agent's hunger and to prepare and thus enable the agent to eat something edible (where eating involves getting a terminal target into the mouth and chewing so as to begin the process of digestion).

The N-input sub-capacity functions to respond to an agent's hunger, i.e., to activate the "Eat" instructive disposition when and only when the agent needs to eat food. It may fail to fulfill this function by activating the "Eat" instructive disposition when the agent doesn't need to eat food. As we'll see below, when this happens, the agent has an inappropriate N-instruction. In order to fulfill the function of preparing and thus enabling the agent to eat something edible, the T-input sub-capacity functions to respond to terminal targets. It is important to note that a single particular P can be an N-input and a T-input at one and the same time. That is, a single particular P may activate an instructive disposition both via its N-input and via its T-input sub-capacities (or components). For example, a rabbit's protective "Avoid" instructive disposition may be N-activated by the presence of a predator such as a fox. In other words, the fox may figure as an N-

input. That very same fox may also figure as a T-input; it may T-activate the rabbit's "Avoid" instructive disposition. Depending on the situation, the rabbit will avoid the fox by e.g., getting a safe distance from it, escaping into a hole into which the fox can't fit, or hopping over to a camouflaging spot and remaining utterly still (i.e., hiding).

If an ID is activated outside of its solution-type's domain of application and the solution/action that activation potentiates (the EIR it constitutes/realizes) begins to be performed, the fact that a token of the relevant kind of need (practical problem) doesn't get satisfied (solved) will at least increase the probability that that ID's application algorithms are adjusted. It's important that the (sub-individual) mechanisms responsible for adjusting an ID's application algorithms are sensitive to whether an activated ID (EIR) is selected for performance as well as whether it continues to be selected for performance. For if an activated ID (EIR) is not selected for performance, then a token of the relevant kind of need (practical problem) won't be satisfied (solved). But since being a local solution is not the same as being a global solution, an activated ID (EIR) might be locally appropriate (applied within its solution's domain of application), and yet not selected for performance. In addition, an activated ID (EIR) might be selected for performance and begin to be performed, but then interrupted, and so never finished, because it ceased to be a global solution. This means application algorithms should not be adjusted just because an ID is activated without solving the relevant type of problem nor just because an ID is activated, selected, and partially performed.

To activate an ID is to potentiate an intentional action (local solution). To potentiate an intentional action is not yet to perform it. Accordingly, there can be activations of IDs without any outward manifestations of actions. IDs can be fully or partially activated. They are fully activated if they are both N- and T-input activated. They are partially activated if they are only T-input

207

activated (or if they are N-activated in the absence of any T-inputs for sub-IDs that activation recruits). Whether a fully activated ID is selected for performance depends on the activities of the organizing dispositions (ODs) with which it's connected.

5.2.3 Sub-Instructive Dispositions

The following situation can occur: an agent may be hungry when a terminal target is not immediately present. In such cases, an agent will need to take some action to bring it about that she's in a position to directly respond to a terminal target. Indeed, in many contexts, the interaction with a terminal target that's required to solve a problem is "extended": it involves sequences of sub-actions and sub-targets by means of which an agent "secures" a terminal target. In order to secure something edible, an agent may need to check in certain places where food might be: a lion will scan the savanna for the movement of potential prey, a squirrel will start digging for nuts, a person will look in the fridge. An agent may also need to do some travelling: a lion will have to chase its prey, a dog will have to follow a scent trail, a person will have to go shopping or to go put in an order at a restaurant. An agent may have to do other sorts of things as well, such as crying or asking for food. And sometimes an agent may need to do all of these things.

It follows, then, that an agent must also be sensitive to the absence of terminal targets. That is, if an agent is hungry and food isn't immediately available, an agent will need to deploy a sequence of sub-actions (or subroutines) and interact with non-terminal targets in order to secure some food. These are sub-solutions; they solve the problem of eating when food must be secured by solving the derivative problem of securing food. Such sub-solutions are embedded in token acts

of eating; they are parts of the intentional actions of eating. They are constituted by exercises or activations of *Sub-Instructive Dispositions* (**Figure 3**).
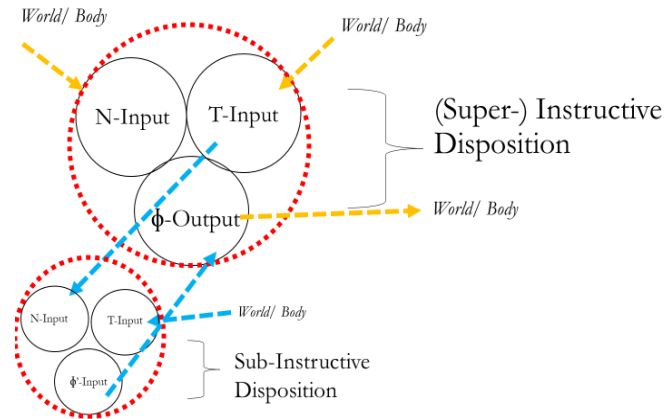
Figure 3 Sub-Instructive Disposition in Relation to (Super-) Instructive Disposition

Sub-instructive dispositions are abilities for solutions that indirectly serve an agent's flourishing by contributing to solutions that directly serve an agent's flourishing. [150] So, when an instructive disposition is N-activated but not T-activated because a terminal target is not present, the disposition's N-activation may N-activate the N-input sub-capacities of some sub-instructive dispositions. The T-input sub-capacities of these sub-instructive dispositions are T-activated by the targets that *are* present in the context. Their φ-outputs are tokens of types of sub-actions (e.g., searching for food by going to particular places). There may be sub-sub-instructive dispositions

---

[150]This does not mean that sub-instructive dispositions provide indirect-PEA. Rather, sub-instructive dispositions have derivative direct-PEA functions in that they function to enable an agent to produce a direct solution (a type of sub-action). They are indirect only in the sense that their PEA functions are derivative; they depend or piggy-back on instructive dispositions. In other words, the "indirectness "of sub-instructive dispositions consists in their essentially functioning to enable an agent to solve derivative or instrumental practical-problems, i.e. problems that arise in the context of solving more fundamental or basic problems that are directly concerned with a need for flourishing.

that are N-activated by N-activated sub-instructive dispositions, and so on. Sub-sub instructions are T-activated by fine-grained features of contexts and their ϕ-output is just the modulating effect they exert on the ϕ-outputs of the (super) instructive dispositions that ultimately N-activate them.[151]

5.2.4 Organizing Dispositions

Where instructive dispositions tell us what to do, organizing dispositions (**Figures 4, 5** and **6**) determine which of the whats we'll actually attempt to do-- which of the instructions will get a chance of becoming an oak tree, as it were, and when. Organizing dispositions (ODs) function to solve global problems (**Figure 4**). A global problem is defined by the logico-practical relations (compatibility, incompatibility, complementarity) that obtain between one's available local solutions in a context. An appropriate solution to a global problem consists in prioritizing and selecting for performance the action(s) that will most promote an agent's overall flourishing in a context. ODs solve global problems by selecting and prioritizing EIRs (de actu contents, potentiated local solutions, activated instructive dispositions (IDs)) for performance. In this way, ODs decide which of the available behavioral options to implement where such decisions are evaluable in terms of a rational standard of appropriateness. Solutions are application-accountable with respect to both local and global appropriateness. Accordingly, ODs must involve mechanisms

---

[151] Instructive dispositions will typically deploy sub-instructive dispositions and/or sub-sub-instructive dispositions even when a terminal target is present. This is because securing a terminal target pretty much always involves some sub-actions that need to be responsive to the particular properties of a terminal target that are relevant for the relevant type of interaction as well as to the particular properties of the context in which it is secured.

that are N-activated by N-activated sub-instructive dispositions, and so on. Sub-sub instructions are T-activated by fine-grained features of contexts and their ϕ-output is just the modulating effect they exert on the ϕ-outputs of the (super) instructive dispositions that ultimately N-activate them.[151]

5.2.4 Organizing Dispositions

Where instructive dispositions tell us what to do, organizing dispositions (**Figures 4, 5** and **6**) determine which of the whats we'll actually attempt to do-- which of the instructions will get a chance of becoming an oak tree, as it were, and when. Organizing dispositions (ODs) function to solve global problems (**Figure 4**). A global problem is defined by the logico-practical relations (compatibility, incompatibility, complementarity) that obtain between one's available local solutions in a context. An appropriate solution to a global problem consists in prioritizing and selecting for performance the action(s) that will most promote an agent's overall flourishing in a context. ODs solve global problems by selecting and prioritizing EIRs (de actu contents, potentiated local solutions, activated instructive dispositions (IDs)) for performance. In this way, ODs decide which of the available behavioral options to implement where such decisions are evaluable in terms of a rational standard of appropriateness. Solutions are application-accountable with respect to both local and global appropriateness. Accordingly, ODs must involve mechanisms

---

[151] Instructive dispositions will typically deploy sub-instructive dispositions and/or sub-sub-instructive dispositions even when a terminal target is present. This is because securing a terminal target pretty much always involves some sub-actions that need to be responsive to the particular properties of a terminal target that are relevant for the relevant type of interaction as well as to the particular properties of the context in which it is secured.

that are sensitive to whether the EIR selections and orderings they determine constitute global solutions. In these ways, I posit that the activities of ODs constitute a primitive form of practical judgement-- judgements that act on *de actu* contents.
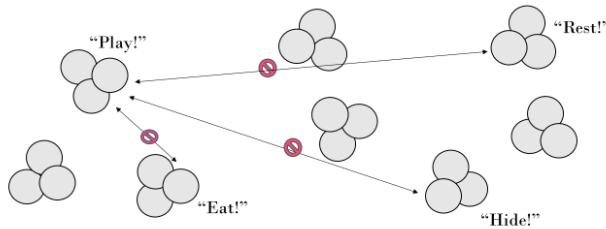
Example of how instructive dispositions may be related via organizing dispositions.

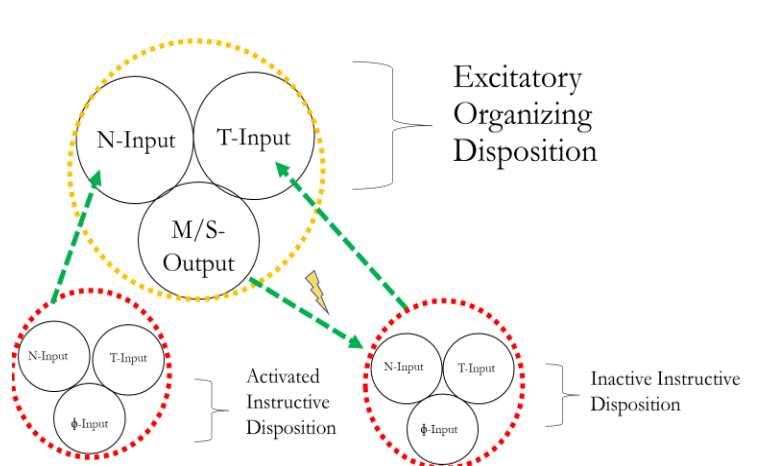**Figure 4 Globally Relevant Relations Between Instrucitve Dispositions**

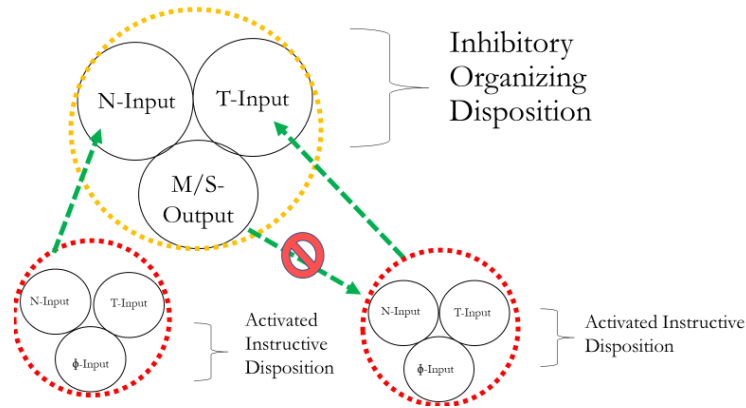**Figure 5 Excitatory Organiszing Disposition**

Figure 6 Inhibitory Organizing Disposition

Organizing dispositions function to modulate instructive dispositions in order to solve global practical problems (**Figure 4**). They enable short-term overall flourishing by, e.g., taking as input the strength of two activated instructive dispositions (two instructions) and producing as output suppression of the weaker input (**Figure 6**), or excitation of the stronger input (**Figure 5**), or both, thereby enabling an agent to execute the instruction that will, in the context, most promote her overall flourishing.

Where application-and-agent-accountable solutions, i.e., rational actions, are concerned, agents can intentionally intervene at the level of action selection (see Chapter 2 and conclusion for further discussion). When this happens, self-directed human actions indirectly do the work of ODs. Humans have abilities (IDs) to notice and inhibit their own EIRs and to create SRs and scaffold their environments in ways designed to elicit competing EIRs. Humans also have abilities (IDs or sub-IDs the operations of which may dynamically involve ODs) to intentionally shift their attention towards certain facts and thus away from others so as to control which of their IDs are activated and when. Attentional actions can indirectly play the role of ODs but increasing or decreasing the amount of activation an ID is receiving. These capacities, in conjunction with other

212

capacities including social capacities and capacities for producing SRs, underwrite the human capacity to act on the basis of intra-subjective reasons (see Chapter 3). Agent-accountability also takes the form of humans performing intentional actions aimed at modifying their existing IDs and ODs and acquiring new IDs. Such intentional actions include practicing habits and skills, but they also include group actions of establishing educational and other social institutions.[152]

5.3 EIRs, *de actu* Contents, & Generative Modes of Presentation

EIRs have *de actu* contents; they are generative modes of presentation of actions and the facts at which those actions are directed. *De actu* contents are a novel concept, but they have features in common with the more familiar concept of demonstrative contents. So, it may be helpful to begin with a brief compare and contrast.

First, I take it that demonstrative contents employ demonstrative concepts which are more similar to concepts in Frege's sense than are *de actu* contents.[153] *De actu* contents are constituted by applications of embodied instructive concepts (which are realized by instructive dispositions). These are importantly different from Fregean senses (perhaps closer to what Frege calls "ideas" than what he calls senses, but take the comparison with a grain of salt). For instance, where Fregean senses are "objective," *de actu* contents are subjective in the sense that they account for how an

---

[152] I haven't yet speculated about the nature of group actions, and I won't attempt to do so in this dissertation.

[153] See McDowell (1984, 1994). Non-conceptualist varieties of broadly demonstrative content were developed by Evans (1982), Peacocke (1998), Dretske (1969, 1981), Tye (2005), Bermudez (1998).

individual subject fundamentally understands or grasps things in the world. They are also subjective in the sense that they are embodied.

In connection with the objective as opposed to subjective nature of demonstrative and *de actu* contents respectively, demonstrative concepts are true or false of what they properly apply to, whereas *de actu* or embodied instructive concepts are appropriate or inappropriate to what they properly apply to. Put another way, demonstrative contents are objective contents in that they present objects as such, that is, as belonging to certain categories or having certain kinds of properties or attributes. And their being objective in this sense is essentially connected to the fact that applications of demonstrative concepts position a subject to have truth/false thoughts about the objects to which they're applied. *De actu* contents, in contrast, have contents that are more subjective-- not in the sense that they present sense data or some such, but rather in the sense that applications of embodied instructive concepts present objects in terms of intentional actions. They are egocentric-- they fundamentally relate objects to a subject's actions rather than relating objects to other objects or classes of objects. Just as the objectivity of demonstrative contents is essentially connected to their having truth conditions, so the egocentricity of *de actu* contents is essentially connected to their having appropriateness conditions. Demonstrative concepts and embodied instructive concepts fit their objects in different ways.[154]

Second, demonstrative contents are supposed to account for the way in which thought directly involves objects-- the way in which subjects can be acquainted with objects, or directly cognitively related to them, rather than being related to objects by means of descriptions those objects uniquely satisfy. And in connection with this, philosophers appeal to demonstrative

---

[154] Chapter 3 described the fitness of embodied instructive concepts (actions) in terms of the notion of "practical fitness."

contents primarily to account for the contents of perceptual experiences, but I take it, distinctively human perceptual experiences. *de actu* contents are also supposed to account for the way in which subjects directly grasp objects, or the way in which subjects are directly cognitively acquainted with objects, and in connection with this, they are essentially bound up with sensory or perceptual capacities. However, *de actu* contents are supposed to account for the fundamental way in which both humans and animals are fundamentally cognitively related to the world.

Third, both demonstrative concepts and contents and embodied instructive concepts and *de actu* contents refer to objects in a way that's naturally captured by "that…" And for both kinds of contents, what "that" picks out is a constituent of the content. But in the case of demonstrative contents, I take it that this way of referring is genuinely a kind of demonstration in the sense that it is accompanied by a kind of description or descriptive element such that the complete content has the form "That F" or "That S is F." In contrast, *de actu* contents target their objects; they relate a subject to an object by means of constituting a potentiated action that's directed at that object. A complete EIR content has the form "ɸ-THAT-THUS!"

Now, some "ɸ-THAT-THUS!" contents are "Judge-That-F" (or related) contents. Since acts of judgement produce or constitute SRs, and SRs can be (broadly) true or false, the appropriateness conditions of such *de actu* contents are one and the same as the truth conditions of the judgements they instruct. For this reason, I am inclined to think that demonstrative contents might be understood as judgement-types of *de actu* contents, that is, as *de actu* contents about acts of judging.[155] But this is just a hypothesis.

---

[155] In the next chapter, I'll argue that an EIR psychology can subsume propositional attitudes in a demanding sense. I'll argue that propositional attitudes properly correspond to clusters of IDs which include IDs for distinctively rational sorts of behavior. If we allow that SRs can be complex, i.e., constructed out of elementary SRs according to certain rules of composition, then we may be able to argue that if one has an

Hopefully these comparisons and contrasts are helpful. I'll now develop the ways in which I've just characterized *de actu* contents in more detail in part by trying to illuminate the nature of a generative (as opposed to surrogative) mode of aboutness.

Since EIR contents are essentially embodied, to better understand them we need to better understand their realizers. As noted, the instructive dispositions (IDs) an agent possesses are her embodied instructive concepts. (**Figure 7**) They play the role of general embodied practical knowledge the application of which is constitutive of fundamental cognition. Concepts are sometimes called representations, so IDs may also be thought of as representations, but representations in the sense of concepts rather than in the sense of articulated representations (or judgments) that are composed out of concepts. IDs, or clusters of them, are also what we refer to when we talk about an agent's dispositional beliefs or standing intentions. These will be discussed in more detail in Chapter 5. Activations of IDs constitute applications of the embodied instructive concepts they realize which, in turn, constitute embodied instructive representations (EIRs). Embodied instructive concepts are general in that they stand-for (constitute abilities or standing-knowledge for) a type of application accountable solution (intentional action). Input activations of IDs potentiate (causally determine) tokens of those types of action by supplying the token facts at which they're directed. This is what accounts for the sense in which EIRs have the general-particular form of judgeable contents (Chapter 1). Embodied instructive representations *(de actu* contents) just are potentiated intentional actions-- particularized tokens of general types of action-

---

ability to employ SRs as such (an issue I'll discuss in the next chapter), then one will meet Evans's (1982) "generality constraint." But I won't develop the account of SRs enough in this dissertation to make such an argument here.

- that are directed at activating-input facts; potentiated intentional actions (EIRs) are generative modes of presentation of such facts.

Let us take as an example Raquel the Raccoon's "Eat" instructive disposition (ID). This ID constitutes her "Eat" or, schematically, her "Eat-[THAT]-[THUS]" embodied instructive concept.[156]
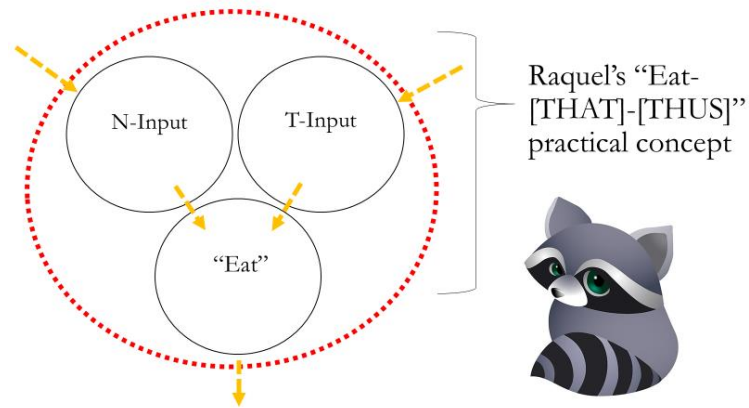

Figure 7 Example of an Embodied Instructive Concept

Raquel's "Eat" embodied instructive concept stands-for eating: it constitutes an ability that is for a type of solution to a practical problem, namely, eating (which solves the need for nourishment). The concept is applied any time the ID is activated (be it fully or partially). When the concept is applied, a more or less determinate act of eating is instructed (potentiated). So, when the ID is inactive, it stands for eating where that action is entirely determinable. The context of activation of the ID is the context of application of the embodied instructive concept and this context-- the activating inputs-- determines or potentiates a token act of eating. The N-input that

---

[156]Exploring the connections between embodied instructive concepts and the embodied concept literatures in philosophy and cognitive science, as well as objections to embodied concept theories, is important work but there isn't space to do it here.

N-activates the ID determines the need (or pseudo-need) at which the token is directed. The T-input that T-activates the ID determines the terminal target ("THAT") at which the token is directed. Other aspects of the context in which an ID is N- and/or T- activated determine the manner ("THUS") of the token action which involves a sequence of activations of sub-IDs for sub-actions, and the context will determine the targets at which those sub-acts are directed. The conditions of activation and thus application of the concept thus determine the appropriateness value for the instructed act of eating.[157]

Embodied instructive representations represent actions, targets, and needs. But they represent all of these things *generatively.* An agent is generatively related to the actions she's in a position to produce because she's in a position to generate them. Generation is a kind of causal relation.[158] Thus, the generative aboutness relation is a kind of causal relation. However, it is not primarily a *passive* causal relation.[159] Rather, the generative causal relation is essentially a kind of *active* causal relation albeit one that involves a crucial passive *dimension*. The passive dimension of generative directedness consists in the fact that needs and targets (states of affairs, facts) activate instructive dispositions for types of actions (local solutions) and in so doing, potentiate determinate tokens of those types of actions (local solutions). The facts that need- and target- activate an instructive disposition become the facts with respect to which the action that's thereby potentiated is directed. Need and target inputs constitute the causal matter (see Chapter 2 §3.2) that

---

[157] There are thus four modes of activating instructive dispositions: isolated N-input activation, isolated T-input activation, simultaneous N- and T-input, and the activation from organizing dispositions. I'll articulate an account of embodied instructive representational attitudes in terms of these modes of activating instructive dispositions in Chapter 5.

[158] This is not to say that it is a merely efficient causal relation; it isn't.

[159] This is a key point of divergence between my account of original representation and other reductive teleosemantics accounts, for the other accounts privilege passive causal relations (though some, i.e., producer-side teleosemantic theories, more than others, e.g., consumer-side teleosemantic theories).

individuates instances of the type of solution the dispositions function to prepare, and they passively causally determine the particular shape or manner of the actions they potentiate. Accordingly, in being generatively directed at an intentional action via an EIR, one is at once generatively directed at the needs and targets at which that action is directed. At the same time, one cannot be generatively directed at needs and targets except by being generatively directed at the intentional action that's directed at them. In this way, *de actu* contents "internalize" the facts at which potentiated intentional actions are directed by being about intentional actions that are directed at them.

To see how this is supposed to work, recall the "ϕ-THAT-THUS!" In representing via an embodied instructive representation, an agent is generatively directed at ϕ and THUS and generatively directed at THAT. The agent is generatively directed at ϕ and THUS because these symbols correspond to a manner that partly individuates a token of a type of action that the agent is prepared to produce by virtue of the activation of the relevant instructive disposition. The agent is generatively directed at THAT because (1) this symbol corresponds to a target that individuates the token of ϕ and contributes to determining the individuating THUS, and (2) the agent is generatively directed at ϕ and THUS. In other words, the action to which the agent is generatively directed is itself directed at THAT.[160] So, an EIR is generatively about an intentional action by

---

[160] Of course, THAT is not something the agent generates. There are two reasons I prefer not to distinguish between the way in which an agent is generatively related to THAT and the way she's related to ϕ and THUS. First, an agent is only generatively related to a particular ϕ insofar as she's related to a particular problem (or pseudo-problem), i.e., a need. And an agent is only generatively related to a particular need insofar as she's related to a ϕ that functions to solve the practical problem it defines, where a type of solution implicitly identifies a type of target, i.e., a THAT. ϕ is, after all, a type of interaction with a type of target. So, the way in which an agent is related to a ϕ is deeply bound-up with the way she's related to a THAT. That said, there is some independence-- the agent isn't related to a particular target until the THAT slot in the schema is saturated by the particular target that activates the instructive disposition. And the THAT slot may be saturated without fully activating the ϕ slot, for that slot is saturated by instances of (disposition activations by) particular needs. In addition, an agent's needs in a sense begin as internal to her while the target-interactions that satisfy them in a sense begin as external to her. This may seem to speak in favor of marking the distinction between how she's related to them. But the second reason I prefer not to mark this distinction is that a token ϕ is always a token ϕ -THUS, and

virtue of being the potentiation of that particular action, and because that particular action is causally directed at the needs and targets that determine it, an EIR is also generatively about needs and targets, where needs and targets (activating inputs) determine the particular shape or manner of the actions they potentiate.[161]

Surrogative directedness extends our generative directedness; we are generativity directed at surrogates which are surrogatively directed at facts. So, in a sense, we are generatively-surrogatively directed at facts when we employ SRs, though for simplicity I'll just say we're surrogatively directed at facts when we employ SRs. We are related to an SR as itself a fact -- as a target of an intentional action. In the language of Chapter 3, SRs are themselves interpretables. In contrast, while EIRs appear to be interpretables for self-conscious creatures like ourselves, we are not fundamentally related to them as interpretables. Rather, fundamentally, they constitute our interpretations. In other words, we are not fundamentally related to them as objects. Rather, they are that through which we are fundamentally intentionally related to objects.

THUS is determined by particular contextual targets (often by sub-instructive dispositions) which may already contain the terminal target THAT. Since a token ϕ is not independent of a token THUS and a token THUS is not independent of targets, the way in which an agent is directed at ϕ seems to swallow-up the way in which she's directed at targets.

[161] In so doing, they also determine whether the potentiated action's appropriateness conditions are met. In the good case, an instructive disposition is need-activated by a token of the type of practical problem it functions to enable an agent to solve. In other words, a token of a type of solution is determined for or in response to a token of the type of practical problem it solves. So, in the good case, a potentiated action is an instance of a solution that is directed at an actual instance of the practical problem (need) it functions to solve (satisfy). And in the good case, an instructive disposition is target-activated by a token of the type of target that is the proper complement of the type of solution the disposition functions to prepare to produce. For recall that solutions to practical problems consist in the type of interaction with the type of target that satisfies the type of need for flourishing that defines the type of practical problem for which it is a solution. In other words, a solution is always a type of interaction with respect to a type of target, and that type of target is the "right target" for a type of problem-solving interaction. So, in the good case, the target that activates an instructive disposition is an instance of the right type of target; a potentiated action is an instance of a solution that is directed at an instance of the right type of target.

What about the bad case? When the activating need-input is not within the domain of application of the local solution for which that instructive disposition stands/ is an ability, the potentiated action is inappropriate: it is directed at a need-input it fails to practically-fit. Likewise, if the activating target at which the potentiated action is intentionally directed is not within the domain of application of the local solution for which that instructive disposition stands/ is an ability, the potentiated action is inappropriate: it is directed at a target-input it fails to practically-fit.

When we are directed at an action via an SR (that is, when we are surrogatively-generatively directed at an action), we are related to that action as an object and in the same way as others may be related to it. In contrast, generative directedness is an essentially first-personal way of being directed at one's actions. Indeed, the notion of generative directedness provides a way of explaining what it is to have a first-personal perspective or a fundamental interpretation: To have a first personal perspective or fundamental interpretation of or on something is to represent it via an EIR; it is to have something presented to you as the object (need, target) of one (or more) of *your* potentiated actions where this consists in being in a position to respond intentionally to it; it is for that thing to be involved in one or more of your intentional action acorns. So, to represent something in the empirical world first-personally is to be generatively related to an intentional action that is token-individuated by being causally directed at that thing. This is what accounts for the fact that one's first-personal perspective is essentially one's own-- no one else can have it-- because one can only be generatively directed at the very actions one is in a position to produce. That is, this accounts for the way in which we first-personally represent parts of the world in terms of the fact that we represent them through being generatively directed at our actions.

Generative directedness constitutes our fundamental sensory mode of awareness of the world: need and target activations of instructive dispositions correspond to sensory, perceptual, emotional, and affective capacities.[162] But while sensory, generative directedness is *not* observational; generative directedness is first-personal in the sense that we are not related to our embodied instructively represented actions as objects-- as something *on which* we have

---

[162]Generative directedness is sensory insofar as T- and N-instructions correspond to perceptions and emotions. Since T- and N-input sub-capacities are activated via causal impacts on sensory organs, when one T- or N-instructs, one is responding to the causes of such impacts. Indeed, sub-capacity T- and N-inputs are perceptual experiences that instruct a response from the agent who has them.

perspective-- except insofar as it is precisely through the actions we represent by means of our embodied instructive representations that we have first-personal perspectives on the needs and targets at which those actions are directed. The actions we instructively represent are those we are in a position to generate; they constitute our inclinations or attitudes towards objective reasons (facts). They have a sensory component in that sensory activations set or determine the targets and needs at which our potentiated actions or intentional action seeds are directed and thereby determine or individuate those actions. In addition, sensory sub-actions will be involved in a great many instructed actions.[163]

I mentioned above that embodied instructive representations are *de actu* contents, which are in important ways subjective as opposed to objective in the way Fregean senses are supposed to be objective. Importantly, however, this does not mean that EIRs/ *de actu* contents are in no interesting sense also objective. The *de actu* perspective is objective in the sense that we are directed at facts as the objects or objectives with respect to which our potentiated actions are directed. So implicit in *de actu* contents is a primitive kind of subject/object distinction in the form of an agent/ target distinction. EIRs are also objective in the sense that the appropriateness of an EIR turns on whether the objects at which the represented action is directed are a practical fit for that action (See Chapter 3). An EIR's success is beholden to the facts in that an EIR fails to

---

[163] There is, then, a nuanced sense in which EIRs "present" us with our own potentiated intentional actions. I say "nuanced" because if one thinks of "being presented with X" as an exclusively passive relation to X, generative aboutness is not a presenting kind of aboutness when it comes to how it directs one at one's own potentiated actions. However, since passivity and activity are inextricably bound up in generative aboutness, a more liberal sense of "being presented with X" does appear to apply: EIRs "present" us with our potentiated actions insofar as they embody our readiness to do them, that is, insofar as we are in a position to intentionally respond to the needs and target at which they direct us. EIRs also "present" us with our potentiated actions is precisely by presenting us with the needs and targets at which they are directed. After all, needs and targets are given to us through those potentiated actions. But EIRs do not passively "present" us with our intentional actions as passive objects: they do not present us with our own actions the way they present us with the actions of others. One might also think of "being presented with" as the passive way of putting what happens when one "grasps" something. An EIR constitutes an embodied grasp of one's actions, both as merely potentiated and as they're performed/ executed, and a fundamentally practical grasp (an embodied instructive grasp) of the parts of the world at which the action is directed.

successfully perform its direct-PEA function when it consists in a potentiated action that fails to practically fit the facts at which it's directed (when it's directed at facts outside the action type's domain of application). EIRs are beholden to the facts also in the sense that embodied instructive concepts (IDs) are abilities for *application-accountable* solutions.

Of course, this is not the kind of "objective" perspective we have on things via surrogative representations. An EIR/*de actu* perspective does not *generally* involve any attribution or description; it is not generally a perspective on an object as a target of attribution or description. Rather, attributing properties to or describing a target is a type of intentional action which requires that one already have a kind of perspective on or psychological grasp of the relevant target. So, this grasp can't itself be the product of actualizing an act of attribution or description. Rather, EIRs constitute our direct-PEA and we have direct-PEA to many things that are not SRs. Direct-PEA consists in intentional action acorns (potentiated actions, embodied instructive representations, *de actu* contents), and the actions for which they are acorns often have nothing to do with producing or exploiting surrogative representations. That said, it is certainly possible to grasp an object as the target of a potentiated act of describing or attributing. In such a case, one's first-personal (direct) intentional access to or grasp of (or perspective on) the object is via the potentiated action of describing or attributing. But crucially, one's access, grasp, or perspective is not achieved through the actualization of the act of describing or attributing. Rather, the actualization of the action depends on one already having direct-PEA to the object.

5.4 The Representational Status of EIRs

Since EIRs are realized by problem-solving dispositions, and since they are constituted by potentiations of intentional actions, I expect many readers to wonder: "why isn't this just dispositionalism?"[164] That is, why should we count EIRs as representations?

Backing-up a bit, we might begin by asking: What is the relationship between psychological representations and behavioral dispositions? Some seem to think that representational capacities rest on behavioral dispositions (e.g. Evans 1982, Schellenberg 2018) Others seem to think that behavioral dispositions rest on representations (e.g. . Fodor 1968).[165] This much seems clear: where we find psychological representations, we find behavioral dispositions. This is sort of the flip side of the relationship between intentional actions and psychological representations discussed in Chapter 2. There it was noted that intentional actions seem to entail that their agents had/have certain psychological representations. Here the point is that the presence of certain behavioral dispositions seems to be a condition on the appropriateness of intentional state ascriptions. After all, what would it mean to say that someone, e.g., believes apples are edible, if they had absolutely no related behavioral dispositions? -- if they were neither disposed to eat apples, nor call them edible when they see them, nor suggest to others that they eat them when hungry, nor answer the question "are apples edible?" affirmatively, etc.?

In Chapter 2 I made a case for understanding the connection between psychological representations and intentional actions in terms of their being two sides of the same coin:

---

[164] One way to answer this question would be to consider the relationships between dispositional, functionalist, and representationalist views to see whether or the extent to which they are (necessarily) mutually exclusive. However, this is not the route I'll take.

[165] See also the concluding chapter of this dissertation-- specifically the discussion of Williamson's reply to Hyman's dispositional theory of knowledge.

psychological representations just are intentional action acorns. On this view, occurrent beliefs, desires, and intentions are related to intentional actions as their potentiations. In keeping with this, in this chapter I'm making a case for understanding behavioral dispositions and psychological representations as two sides of the same coin: standing psychological representations or concepts are realized by problem-solving dispositions (abilities), while occurrent psychological representations are realized by the activations of said dispositions.[166]

Here, in a nutshell (pardon the pun!), is the proposed view: Activations of an agent's instructive dispositions constitute applications of her embodied instructive concepts. And these constitute embodied instructive representations (EIRs). EIRs just are intentional action acorns -- the initial stage in a process of acting that contains the potential of (instructions, recipe for) the full action.[167] They are, in other words, potentiated (application accountable) solutions. The potentiated solution stands-for a whole solution because it is the potentiation of the whole; just as an acorn "contains" the form of the oak tree of which it is the potential, an intentional action acorn "contains" the form of an action of which it is the potentiation. Surrogative representations are objects-- vehicles with which we engage-- that stand-in-for and thus go proxy for what they're about; they are solutions to proximity problems because they potentiate indirect solutions (solutions that involve indirectly interacting with something X by way of directly interacting with some other thing Y). In contrast, embodied instructive concepts are embodied abilities (problem-solving dispositions) that stand-for a type of action (solution) by virtue of being abilities to perform

---

[166] It's worth noting that there are quite a few broadly action-based or dispositional theories of desire, e.g., Smith (1987, 1994), Millikan (1984), Papineau (1987), Scanlon (1998), Lewis (1972).

[167] Of course, just as a recipe does not tell you how to break the eggs, so the acorn does not contain all of the THUS material from the start-- much is supplied by the context of execution and draws on sub-solutions (sub-EIRs).

them and embodied instructive representations are applications of those abilities -- potentiations of intentional actions.

Now, the fact that EIRs are so different from representations as standardly construed (the model of which is the proposition) may cause some philosophers to doubt that they are really representations. Indeed, abilities and dispositions and their exercises are sometimes thought to be alternatives to representations. [168] Still, philosophers might agree that intentional action acorns are similar to representational *attitudes*, but not to representational contents. Alternatively, one may object that EIRs are either propositional representations in disguise -- e.g., imperatives -- or else they are not really representations at all.

Let's start with the charge that EIRs are not really representations. One might simply insist that representations must have alethic success conditions. However, unless such claims are accompanied by an account of what it is to be a representation and why, and unless such an account did a good job of explaining why all the things that intuitively are representations are, and all the things that intuitively are not are not, it will simply beg the questions. Chapters 1-3 provided different sorts of reasons for thinking that alethic conditions are not essential to representations. We may now appeal to some of those criteria to make a case for counting EIRs and thus intentional action acorns as representations.

Chapter 1 discussed two ultimately convergent accounts of what is essential to representation both of which were arrived at by considering motivations for the descriptive thesis. On the first account, what's essential to representations is that they are judgeable contents, where these have certain features. On the second account, the essential feature of representations is that

---

[168] See e.g., Schwitzgebel (2002) and Quilty-Dunn (2018).

they have informing functions. I already argued in that chapter that informative function might be analyzed as PEA. In addition, the informing function criterion is arguably similar in spirit to the subjective reasons criterion I argued for in Chapter 3, where I also argued that subjective reasons can be analyzed as PEA. So I'd like now to focus on the idea that representations must be or have the form of judgements.

The tendency to assume the descriptive thesis goes hand-in-hand with a tendency to treat all judgement as theoretical judgement. This latter tendency was noted by Dancy (2018) and discussed in Chapter 1. Following Dancy, I suggested that rather than understanding practical judgements as a species of theoretical judgement, theoretical judgement might be understood as a species of practical judgement. There I also suggested, again following Dancy, that the conclusion of a practical reasoning (of a fundamental kind) might itself be an action. On such a view, then, the OD-selection of EIRs (*de actu* contents) for performance may be understood as contents of practical judgments. Indeed, IDs and EIRs are related to reasoning in the following ways:

First, the activation of an ID may be modeled in terms of an Aristotelian practical syllogism: the premises are "apparent" needs (instructive desires) and targets (instructive means-end beliefs), where "apparent" just means that some facts have activated N- and T- inputs to and ID, and the conclusion is the potentiation of a token of the type of *local* solution (action, activation of Φ-output sub-capacity) the ID is an ability for. Such reasoning is successful iff the conclusion is *locally* appropriate, i.e., if it practically fits the actual (rather than apparent) facts, i.e., if the premises are sound (if the activating inputs were appropriate).

Second, the selection of potentiated local solutions (EIRs, activated IDs) for performance by ODs may likewise be modeled in terms of an Aristotelian practical syllogism, only this time the premises are EIRs (activated IDs that are inputs to ODs), and the output is the selection of the

winning EIRs, where such reasoning is successful iff the conclusion is *globally* appropriate. Moreover, ODs relate IDs and thus EIRs to one another by practical-logical relations of compatibility, inclusion, and exclusion (and perhaps others). In other words, ODs realize the holistic nature of embodied instructive concepts.

Third, sub-EIRs (sub-IDs) may be modeled in terms of an Aristotelian practical syllogism, only their need-inputs will be activated IDs, while their target inputs will be context-specific facts, and their outputs will be modifications of the Φ-output of the overarching activated ID (EIR) in response to a dynamic succession of targets. Sub-EIRs may also be modeled in terms of instrumental or means-end reasoning, as an activated overarching or super-ID will, together with the context in which it is activated, recruit sub-IDs (sub-EIRs) as the *means* of performing the overarching instructed action.

Now, in all of these cases, reasoning is understood causally-- it's not itself an intentional activity. But this isn't intended to be an account of reasoning as a rational activity. It is, rather, an account of what provides reasoning as a rational activity with material to reason about. In other words, EIRs are also involved in intentional activities of reflection and reasoning. EIRs themselves are the products of causal-teleological processes that can be modeled as processes of reasoning, but those processes are not themselves intentional reasoning activities. If they were, there would have to be some more fundamental kind of psychological representations in terms of which that reasoning is conducted, and EIRs would not be the fundamental psychological states I'm claiming they are.

We can also understand why EIRs (*de actu* contents) should count as representations in terms of the broadly Kantian representationalist criteria for representation set out in Chapter 1:

GENERALITY: Understanding is in some sense more general than the object it understands: it identifies the object as falling under some category/ies, instantiating some property/ies, partaking in some kind, or as being an individual (self-identical) that may be encountered in variant contexts.

ACTIVITY: Understanding is in some sense actively supplied by the subject and imposed on or applied to the object the subject grasps in its terms.

NORMATIVITY: Understanding succeeds where it "fits" the object to which it's applied or on which it's imposed; understanding fails where it fails to "fit" the object to which it's applied or on which it's imposed.

First, EIRs have the general structure of judgeable contents: they are constituted by applications of embodied instructive concepts and therefore consist in a general element, namely, the action/ solution type, as well as a singular element, namely, the token of the type of action that's particularized by the particular facts to which the embodied instructive concept and thus potentiated intentional action (EIR) is applied.[169] Embodied instructive concepts are general because they stand-for intentional action types. Intentional actions are typed according to the type of solution they constitute, and solutions are types according to the type of practical problem they function to solve. Token intentional actions (solutions) are particularizations or determinations of a determinable type of solution (the kind of solution an instructive disposition is an ability for/ an embodied instructive concept is a concept of). An EIR/ *de actu* content/ intentional action acorn potentiates a token of a type of action. So, an EIR is constituted by the application of something general-- the type of action an embodied instructive concept stands for. In this way, EIRs satisfy GENERALITY.

---

[169] Borrowing from Millikan (2017), we might say that embodied instructive representations are "articulated into components or aspects" (125-132) some of which are variants (token need and target facts, the time and context of activations) others of which are invariant (the solution/ action type).

Second, an intentional action acorn/EIR constitutes a subject's fundamental practical understanding or grasp of a situation: a subject grasps a situation by virtue of applying an ability to it, where the application of this ability just is the potentiation of a token of the type of action the ability is for, and this potentiation just is an intentional action acorn/EIR. So, the way in which a subject fundamentally grasps or understands a situation (fact) is in terms of her own potentiated activity. In this way, intentional action acorns satisfy ACTIVITY.

Finally, an EIR is appropriate iff the action it generatively represents (first-personally instructs) is appropriate; otherwise it is inappropriate. Moreover, EIRs are potentiations of token application accountable solutions which are, as such, sensitive to whether EIRs are appropriate. In this way, EIRs satisfy NORMATIVITY.

I offered a different sort of criterion for representation in Chapter 3: representations are subjective reasons. EIRs are intra-subjective reasons while SRs are inter-subjective reasons, and the PEA analysis was offered as an account of representations and subjective reasons generally: both propositional (surrogative) and embodied instructive representations essentially function to provide PEA. Chapter 2 suggested that the psychological representations that explain actions may be understood as intentional action acorns, and the present chapter identifies intentional action acorns with EIRs and thus intra-subjective reasons and direct-PEA. No one doubts that SRs are robust representations, and on the PEA analysis, SRs and EIRs just provide different kinds of PEA. And since PEA provides a general analysis of representation, EIRs fit into an encompassing account of representations that can do interesting explanatory (illuminating and unifying) work. If this isn't a good reason to count them as representations, I don't know what would be.

Now let us finally turn to the reasons why EIRs are not just directives or imperatives. Directives and imperatives are SRs: they provide indirect-PEA. They *stand-in-for* actions and can

230

thereby *indirectly* potentiate corresponding actions. They constitute surrogative modes of presentation or senses. In contrast, EIRs provide direct-PEA. They do *not* stand-*in*-for those actions. Rather, they *stand-for* the actions they instruct by virtue of being the potentiations of those very actions; they're generative modes of presentation. A directive or imperative is a representation that can be an input to our output of an EIR: one may be the recipient of a directive or imperative or one may issue a directive or imperative. If one is inclined to think that EIRs are directive or imperatives, one is presumably thinking in terms of the recipient-side. But if one is the recipient of a directive or imperative, it will only potentiate a corresponding action if it activates an instructive disposition (ID) to perform the corresponding action. That is, the directive or imperative only potentiates the corresponding action if the recipient produces an EIR for the corresponding action, whereas the EIR is identical to the potentiation of the corresponding action. So, an EIR for an action ϕ entails that an instance of ϕ has been potentiated, whereas an SR that stands-in-for ϕ functions to potentiate an instance of ϕ but does not entail that an instance of ϕ has in fact been potentiated. In addition, EIRs have constitutive appropriateness conditions but not alethic success conditions, while imperatives or directives, qua SRs, will at least involve parts that have alethic success conditions, i.e., satisfaction conditions, even if their overall function is not to describe. Finally, EIRs constitute intra-subjective reasons while SRs, and thus imperatives and directives, constitute inter-subjective reasons.

5.5 Conclusion

In this chapter, I connected the key concepts introduced in the preceding chapters. previous chapters. I identified the notion of *de actu* contents from Chapter 1 with the notion of an "intentional action acorn" from Chapter 2 and with the notion of an embodied instructive representation (EIR) and hence intra-subjective reason (reason *out of which*) from Chapter 3. I also chained the definition of representation as PEA from Chapter 3 with the definition of intentional action as an application-accountable practical problem from Chapter 2 yielding a unified and informative account of intentionality. I also introduced an account of the problem-solving dispositions that realize EIRs which I used to begin to cash-out the metaphor of an intentional action acorn and to put some skin on the bones of the conception of practical judgement that acts on *de actu* contents at which I gestured in Chapter 1. Finally, I attempted to illuminate the nature of EIR content, i.e., what it means to be a *de actu* content or generative mode of presentation, and to address potential objections and confusions about the representational status of EIRs.

6.0 Chapter 5: An Embodied Instructive Representational Psychology & Its Advantages

The last chapter began to flesh-out an account of embodied instructive representations (EIRs). In this chapter I'll develop that picture into an alternative to the propositional attitude psychology and argue that, as a substantive conception of psychological representations, the EIR psychology is preferable. By substantive conception, recall, I mean a conception that is committed to treating psychological representations as causally efficacious entities which true ascriptions of beliefs, desires, and other attitudes or other mentalistic talk refer to but do not necessarily describe accurately. On such a conception, the right way to characterize the nature of these entities is a live philosophical question.

I'll begin in (§1) by developing an intentions-first EIR psychology spelled out in "embodied instructive attitudes": N-instructions, T-instructions, and i- and I-instructions. I'll explain how these are realized by different ways of activating instructive dispositions and I'll argue that they plausibly constitute the basic form of occurrent desires, beliefs, and intentions.

In (§2) I'll argue that EIRs can subsume propositional attitudes. I'll distinguish between "full-blooded" propositional attitudes and IDs/ EIRs that theorists might call "propositional attitudes" by extension. I'll propose that full-blooded propositional attitudes (or "propositional attitudes proper") are constituted by clusters of IDs (and EIRs resulting from activations from IDs in these clusters) for broadly critical or rational activities. Such activities function to solve "rational" practical problems, which are often proximity problems. Hence, they typically involve responding to and producing SRs. Rational practical problems may be grounded in something like a need for self-consciously acting (including epistemically acting) as one should-- a need for

justification, understanding, or the ability to give an account to answer all varieties of "Why?" -questions. Examples of such activities include seeking and constructing explanations and testing and evaluating theories and asking after the justification for explanations.

The subsequent sections will argue that an EIR psychology has advantages over the standard propositional attitude psychology. Specifically, in (§3) I'll argue that the standard propositional attitude model faces two clusters of challenges. The first cluster concerns obstacles to illuminating psychological representations when the descriptive thesis, and thus the propositional attitude model, is assumed. The second concerns limits on the explanatory power of propositional attitudes. Finally, in (§4) I'll argue that the EIR model addresses these challenges.

6.1 An EIR Psychology

In what follows, I will sketch an account of embodied instructive representations (EIRs) in order to argue that they can subsume propositional attitudes as a sub-species. As I did in the previous chapter, I'll use the following schema to model the form of *de actu/* embodied instructive representational content:

"ɸ-THAT-THUS!"[170]

---

[170] On this schema, ɸ is a type of solution (action), e.g., eating, THAT is target (e.g., a particular apple), and THUS is a manner or realization of ɸ which often involves various sub-acts (e.g., climbing an apple tree). "!" indicates the "force" or urgency of the embodied instruction.

As described in the last chapter, EIRs are realized by activations of instructive dispositions. I'll review some basic features of instructive dispositions before I turn to the explanation of how propositional attitudes can be subsumed under EIRs as a special type.

6.1.1 Review of Problem-Solving Dispositions

Problem-solving dispositions are abilities or capacities to solve, or to apply solutions to, practical problems. They come in three types: *Instructive Dispositions,* i.e., abilities to solve local practical problems, *Sub-Instructive Dispositions*, i.e., sub-abilities involved in solving local practical problems, and *Organizing Dispositions*, i.e., abilities to solve global problems that piggyback on instructive dispositions *qua* abilities to solve types of local problems.

The instructive dispositions an agent possesses are her embodied instructive concepts. Instructive dispositions (IDs) are composed of a "ɸ-output sub-capacity" that produces a type of activity, a "need-input sub-capacity" that responds to instances of a type of need (local problem), and a "target-input sub-capacity" that responds to instances of the type(s) of targets that, through a type of interaction, satisfy those needs. Activations of IDs constitute applications of these embodied instructive concepts which, in turn, constitute EIRs. I'll continue using the example from the previous chapter: Raquel the Raccoon's "Eat" ID (**Figure 7**). This ID constitutes her "Eat" or, schematically, her "Eat-[THAT]-[THUS]" embodied instructive concept.[171]

---

[171]Exploring the connections between embodied instructive concepts and the embodied concept literature as well as the more general embodied, embedded, enactive, and extended mind literatures in philosophy and cognitive science is important work but there isn't space to do it here.
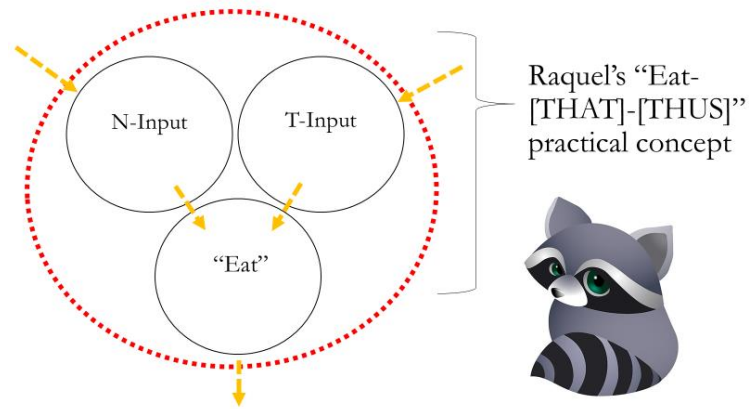
Figure 7 Example of an Embodied Instructive Concept

Raquel's "Eat" embodied instructive concept stands-for a determinable type of action, namely, eating. The concept is applied any time the instructive disposition that realizes it is activated. When the concept is applied, a more or less determinate act of eating is instructed (potentiated). In other words, when the instructive disposition is inactive, it stands for eating where that action is entirely determinable. The context of application of the concept determines or potentiates a potential act of eating in that 1) the N-input that N-activates the disposition determines a particular instance of the type of practical problem (nutrition), 2) the T-input and its context determine a particular terminal target ("THAT") and manner of securing a terminal target ("THUS") which involves a sequence of sub-acts (sub-instructive dispositions) and their targets. The conditions of activation and thus application of the concept thus determine the appropriateness conditions for the instructed act of eating.

There are thus four modes of activating instructive dispositions: isolated N-input activation, isolated T-input activation, simultaneous[172] N- and T-input activation, and the activation from organizing dispositions. The next section will articulate an account of EIRs in terms of these modes of activating instructive dispositions. In so doing I'll describe features they have in virtue of which they can play the role of beliefs, desires, and intentions.

Table 2 Problem-Solving Dispositions & EIRs

Types of local solutions	IDs (take facts via T- and N-inputs, potentiate/ produce behaviors as outputs)	Embodied Practical (Proto) Concepts	Activations constitute EIRs (i, N-, and T- instructions)
Global solutions	ODs (take activated IDs as inputs and order them as outputs)	Decision making procedures (holism)	Select/prioritize EIRs for performance (l-instructions)
Token local solutions (complete)	Fully activated instructive disposition to ϕ	Complete EIR to ϕ-THAT-THUS!	i-instructions (candidate l-Instructions)
Token local solutions (partial/ aspects)	Activation of sub-instructive dispositions (N- and/or T-input activation)	Partial EIR to ϕ-THAT-THUS!	N- & T-Instructions i
Sub-solutions (means)	Activation of sub-instructive dispositions	Determine the manner of token instructed action (the THUS)	Sub-+ instructions

---

[172] The precise nature of the activations is an empirical question. By "simultaneous" activations I mean that an ID is both T- and N-activated at the same time. This might happen if it is first T-activated and then a little later N-activated so long as it continues to be T- (or sub-T) activated. Likewise, and ID may be simultaneously T- and N- activated if it is first N-activated and then, at some later time, T-activated, so long as the N-activation persists to the time it's T-activated. It's worth noting that one of the functions of SRs, including potentially psychological SRs, is that they can stand-in-for need and target facts so as to T- (and also N-) activate IDs in their absence (when those facts cannot directly activate the relevant IDs).

6.1.2 Intentions-First

The propositional attitude model of psychological representation is a belief-first model of psychological representation in the sense that it begins with a theory of the form and nature of belief and uses that as a blueprint for the other attitudes. Its conception of beliefs is broadly "theoretical": beliefs are understood as states that are not in general essentially related to any actions and that aim at the truth and thus have constitutive truth, veridicality, accuracy, or satisfaction conditions. Beliefs serve as a blueprint for desires in that desires are understood as states that are not in general essentially related to any actions-- though since they're motivating they're more closely related to actions than beliefs-- and that aim to have their contents made true and thus have satisfaction conditions. In other words, both beliefs and desires have satisfaction conditions-- their contents are supposed to "fit" in the sense of match or be descriptively satisfied by the facts. The difference between them is simply their direction of fit: beliefs aim to have their contents descriptively fit the world; desires aim to make it the case that the world descriptively fit their contents. Finally, beliefs serve as a blueprint for intentions in that intentions are typically thought to be constituted by, composed out of, or derived from belief-desire pairs and to constitute a special kind of belief or desire (e.g., a desire with a particular kind of content and thus satisfaction condition). Intentions come the closest to being essentially related to actions on this model in that they're the most proximal representational causes of actions. But they fall into the class of propositional attitudes, i.e., psychological representations, which is ontologically distinct from the class of actions.

The EIR model of psychological representation I will sketch in this section is an intentions-first rather than a belief-first model. Rather than beginning with a conception of beliefs, it begins

with a conception of intentions. Rather than treating beliefs and desires as more basic than and giving rise to intentions, on the EIR model, intentions are basic, and beliefs and desires are understood as aspects of intentions or as partial intentions. The four modes of activating instructive dispositions described at the end of the previous section constitute different aspects of or partial intentions (acorns). Moreover, intentions are not in general propositional attitudes (though I'll argue below that some propositional attitudes correspond to certain types of intentions,) and they are intentional action acorns and *ipso facto* essentially related to intentional actions (see Chapters 2 and 3).

EIRs do not have alethic success conditions essentially. They have both a mind-to-world and a world-to-mind direction of fit because they are supposed to potentiate an action that will practically fit the world. Of course, we can think of an EIR's appropriateness conditions as the conditions under which an EIR content is "satisfied," but it is important that we recognize that where SRs do articulate their satisfaction conditions because they stand-in-for what they represent, EIRs do not articulate their appropriateness conditions. The content of an EIR is just an action directed at facts. The action does not describe itself as appropriate and the action does not describe the facts at which it is directed because the action itself is not a description (unless the action is one of producing a description of those facts).

So long as we keep this in mind, there's no harm in associating satisfaction conditions or other kinds of alethic success conditions with EIRs. In specifying alethic conditions to associate with an EIR, we may focus on its mind-to-world direction of fit aspect by specifying how the real situation (the actual facts (target/ need) at which an EIR is directed would have to be for the EIR to be appropriate. For instance, we can describe a T-instruction as a veridical or falsidical perception or a true or false belief according to whether its target (activating T-input) is an

239

appropriate target. So, T-instruction for eating that's directed at an apple, i.e., a T-instruction with the content "Eat-THAT [Apple]-THUS," may be described as a veridical perception or a true belief because apples are within the EAT solution-type's domain of application. We can do the same thing with N-instructions and sub-instructions (sub-EIRs). Alternatively, in specifying alethic conditions to be associated with an EIR, we may focus on its world-to-mind direction of fit aspect by specifying the conditions of successful execution (performance) of the action the EIR instructs (of which it is the potentiation), that is, what it would be for the relevant intentional action acorn to fully develop or actualize.[173]

The EIR intentions-first model is robustly methodologically intentions-first in that it starts its theorizing about the nature of psychological representations from thinking about intentions *stripped of their belief-first model interpretation* (akin to the way chapter 3 began its theorizing about representation with the notion of subjective reasons). Since intentions are *prima facie* the closest representations to actions, this approach derives their success conditions -- appropriateness conditions, to be precise--from those of actions. So, intentions on this model are not a species of prediction. The EIR model is also intentions-first in that, rather than pulling intentions out of beliefs and desires, it pulls beliefs and desires out of intentions. Intentions, beliefs, and desires are

---

[173] Note that one *could* adopt an intentions-first EIR model of psychological representation without identifying intentions with intentional action acorns. So even if one is not attracted by the idea of an intentional action acorn, one might still be attracted to an intentions-first psychology. However, since EIRs cash-out the metaphor of intentional action acorns, a non-acorn intentions-first psychology would be a non-EIR intentions first-psychology (unless, of course, one merely wants to drop the metaphor while holding onto EIRs qua realized by activations of instructive dispositions, constituted by applications of embodied instructive concepts, etc.) It's also possible that one could have a coherent intentions-first psychology without EIRs. One could, for instance, begin with intentions understood as propositional attitudes and then understand beliefs and desires as somehow modes or aspects of that propositional attitude. I have no idea what that would look like, but I don't want to rule it out. However, as we saw in Chapter 2 in the context of motivating intentional action acorn, and as we saw in Chapter 3 and will to some extent see again in Chapter 5 in the context of trying to provide an illuminating account of original intentionality, part of the point of starting with intentions is to make room for a more practical conception of the normativity of psychological representations. Intentions on the belief-first propositional attitude model inherit the broadly theoretical species of success conditions from the broadly theoretical conception of beliefs that is their model. So conceived, intentions appear to be a species of prediction.

all EIRs-- all employments of partial employments of embodied instructive concepts. Hence, they all have constitutive appropriateness conditions. As per the PEA analysis (see Chapters 3 and 4), psychological representations provide direct-PEA. So, intentions provide direct-PEA, and the beliefs and desires we pull out of them do too (they are just aspects of the direct-PEA the intention provides). *Constitutive* satisfaction conditions only enter the picture when indirect-PEA comes on the scene, and indirect-PEA does not come onto the scene for most intentions. As we'll see below, there is a subspecies of intentions that essentially involves indirect-PEA as inputs to or outputs of instructive dispositions and thus as aspects of EIRs, however EIRs themselves constitute direct- rather than indirect-PEA.[174]

6.1.3 The Basic EIR Model

On an EIR psychology, true ascriptions of occurrent beliefs/perceptions, desires/emotions, and intentions refer to what I'll call "embodied instructive attitudes." Embodied instructive attitudes are realized by modes of activating instructive dispositions, as these realize modes of

---

[174] One might wish to withhold the term "belief" and perhaps also the term "desire" for states that have constitutive satisfaction conditions. The proposed view is revisionary in that it holds that beliefs and desires have appropriateness conditions essentially rather than satisfaction conditions. That said, it remains possible on the proposed view to *associate* satisfaction conditions with beliefs and desires. With respect to the beliefs and desires humans and animals share, their satisfaction conditions might be understood as the conditions under which the intentional actions of which they are potentiations (or primings) are fully executed. Alternatively, one could identify satisfaction conditions with the conditions under which those intentional actions are both appropriate and fully executed. One can also characterize the way in which T- and N-inputs contribute to an instructed action's appropriateness conditions: a belief (T-activation/ T-instruction) is satisfied if it is activated by (directs the primed action at) an appropriate target, and a desire (N-activation/ N-instruction) is satisfied if it is activated and thus directs the potentiated action at an instance of the relevant sort of need. So understood, satisfaction conditions would be the conditions under which an activating input is in the ID's solution-type's domain of application. In the case of the sorts of beliefs that are distinctive of humans, a T-instruction to produce an SR may be said to have a satisfaction condition corresponding to the satisfaction condition of the SR. In the case of the sorts of desires that are distinctive of humans, the SR (or SRs) an agent treats as standing-in for her plans may be said to have a satisfaction condition. And in the case of distinctively human beliefs and desires, beliefs and desires may also be said to have the standard directions of satisfaction, with beliefs (SRs produced or committed to) aiming to fit the way the world is and desires (SRs serving as plans) aiming to make it the case that the world satisfies their contents. More on this below.

deploying embodied instructive concepts. There are four modes of activating instructive dispositions. These modes realize what I'll call T-instructions, N-instructions, i-instructions, and I-instructions.[175] I'll focus on embodied instructive attitudes that correspond to 1) perceptions, perceptual beliefs, or means-end beliefs; 2) simpler emotions and/or desires and/or affective states, and 3) simpler intentions. However, once I have the basic picture in view, I'll argue in the next section that an EIR psychology may be able to subsume all varieties of propositional attitudes, including more sophisticated types of beliefs, desires, emotions, intentions, and related states.[176]

As noted, EIRs are constituted by deployments of embodied instructive concepts which are realized by activations of instructive dispositions. Embodied instructive concepts might be thought of as "standing intentions" or "dispositional intentions" as opposed to "occurrent intentions." Standing or dispositional intentions are similar to standing or dispositional beliefs. Indeed, I'll suggest that standing or dispositional beliefs just are either embodied instructive concepts (instructive dispositions), clusters thereof, or T-instructions (partial applications of embodied instructive concepts/ partial activations of instructive dispositions). Embodied instructive concepts might also play the role of some of the things we call an agent's "prior intentions", though as I'll explain in §2 below, for the most part, prior intentions either correspond to i-instructions that are

---

[175] Activities of sub-instructive attitudes yield sub-T-instructions, etc. But sub-instructions get swallowed up into the manner of instructed actions. Sub-instructions are part of what ascriptions of beliefs, desires, and intentions refer to insofar as they refer to instructive attitudes, but they are referred to via N-, T-, and i-/I-instructions.

[176] Because embodied instructive representations are more fundamental than surrogative representations, I am committed to the falsity of the propositional attitude model of mental representations in the following sense: I deny that the contents of mental representations *generally* are surrogative representations such as propositions. In a sense, I'm claiming that the *attitudes* of propositional attitudes already have their own contents, namely, embodied instructive contents. An embodied instructed content is a potentiated intentional action that is itself intentionally directed at facts/ states of affairs insofar as it constitutively functions to solve a problem involving them; it is the content of an EIR which is constituted by the application of an agent's embodied instructive concepts (activation of instructive disposition). Surrogative representations do not figure in all or even most intentional actions. So embodied instructive representational (*de actu*) content, as a species of content, does not essentially involve propositional content.

not yet I-instructions, or to plans, which are constituted by sets of EIRs in relation to certain surrogative representations (SRs).[177] An i-instruction is a fully (T- and N- input) activated instructive disposition while an I-instruction is a fully activated instructive disposition that has also been selected by organizing dispositions for performance. In what follows I'll explain how beliefs and desires correspond to the T- and N-input modes of activating instructive dispositions before returning to say a bit more about i- and I-instructions.

T-Instructions (Beliefs/Perceptions)

T-instructions correspond to the T-input mode of activation of an instructive disposition. In **Figure 8** below, an apple visually activates Raquel's "Eat" instructive disposition. In other words, the apple serves as a T-input. This activation constitutes an application or mode of application of her "Eat-[_THAT_]-[_THUS_]!" embodied instructive concept. The application of this concept constitutes an embodied instructive representation to "Eat-[APPLE]-[THUS]." The representation instructs an act of eating that is determinate to the extent that it has a particular terminal target together with a particular realizing manner ("THUS")—a sequence of sub-sub-acts— which is determined by the terminal target and its context. However, in this scenario, Raquel's "Eat" instructive disposition is *only* activated by the T-input; she only has a T-instruction. In the absence of a corresponding N-instruction, Raquel isn't motivated to eat (she isn't hungry).

---

[177] Plans also importantly involve what Andy Clark calls "scaffolding," which describes a technological extension of the mind that "offloads" information from the brain onto technology. (Clark, 1997, 2008). The term "scaffolding" takes inspiration from Lev Vygotsky's concept of a "Zone of Proximal Development," or ZPD (1986, p. 187). ZPD refers to what students can't yet recall by themselves without resorting to educational resources that support learning and memory such as dictionaries and multiplication tables. Educational psychologists call these resources "instructional scaffolding."

So, while Raquel has a visual T-instructive representation with the content "Eat-[APPLE]-[THUS]," she isn't moved (motivated) by that instruction.
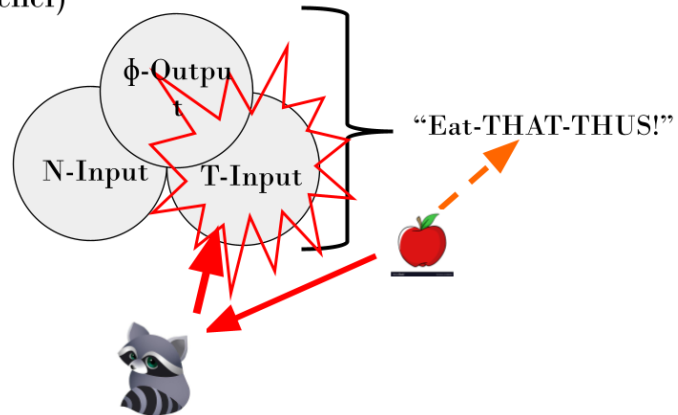

Figure 8 Example of a T-Instruction

The activation of the T-input sub-capacity of Raquel's "Eat" ID constitutes a T-instruction only by virtue of the fact that the T-input sub-capacity is a constituent of the ID such that activating the T-input sub-capacity constitutes an activation (albeit partial) of the instructive disposition. No sub-capacity represents independently of its role in one or more IDs, and when a sub-capacity is activated, it constitutes an EIR. This is compatible with the fact that one and the same T-input sub-capacity may figure in multiple IDs. For example, say Raquel is an inventive and industrious raccoon who suffers from back pain. Seeing the apple may activate the N-input component of both her "Eat" ID and a sub-ID of her "Soothe" ID, that is, an ID that functions to instance solutions to injury-related problems. The apple's N-activation of Raquel's "Soothe-with-[THAT]-[THUS]" ID

constitutes an EIR with the content "Soothe-with-[Apple]-[THUS]" where "THUS" involves rolling her back against the apple. [178]

Some, perhaps most, instructive dispositions have a jointly sufficient set of T-inputs. To explain what this means I'll start with an example. An apple may be a visual T-input to the T-input sub-capacity of Raquel's "Eat" instructive disposition, but the visual apple input alone may not be sufficient to activate the T-input component of her eat disposition. We might imagine that there are some contexts in which apples are unlikely to be edible, e.g., if they are sufficiently rotten so visual contact with the apple won't activate the N-input of Raquel's "Eat" instructive disposition unless Raquel also has olfactory contact with the apple. In other words, in order for the apple to activate her "Eat" instructive disposition, the apple must both be a visual and olfactory T-input to that disposition.

T-instructions are comparable to beliefs (especially perceptual and means-end beliefs). In the example pictured above, Raquel has a T-instruction but no corresponding N- instruction. As we'll see in a moment, N-instructions are comparable to desires. Since there's no N-instruction, this is a mere T-instruction; the T-activation does not on its own yield an i-instruction (comparable to an intention, lowercase "i"), and so of course neither does it yield an I-instruction (comparable to an Intention, uppercase "I"). Accordingly, one way in which T-instructions are like means-end beliefs is that they can occur independently of N- (and thus i- and I-) instructions. By considering a case in which a T-instruction occurs in the absence of a corresponding N- (and i-/I-) instruction, we can also isolate the contribution of the T-input activation (T-instruction) to preparing an agent

---

[178] Alternatively, the instruction might be "Soothe-[Back]-[THUS]" and sub-instructive dispositions that specify the "THUS" in terms of "Roll against-[apple]-[THUS]."

like Raquel to produce a particular instance of a type of action. In this case, Raquel's T-instruction prepares her for an instance of eating that is individuated by the particular apple that is the potential action's target and the way it and its context determine the manner of the potential action.

I've just described the particular way in which T-instructions contribute to the determination of an instructed action. And this corresponds to the particular way in which T-instructions contribute to the appropriateness conditions of an instructed action. For if an instructed action has an inappropriate target (a target outside the action type's domain of application), the action will not (non-accidentally) succeed. If, for instance, apples are not nutritious for Raquel, or if the particular apple she's looking at is actually a wax apple, then the instruction to eat that apple is inappropriate. After all, if she were hungry, and if she were to eat it, it would not in fact nourish her. So, the act of eating that particular apple would not constitute a solution, and thus it would not be appropriate. Therefore, the T-instruction to do that act is inappropriate.

The way T-instructions contribute to the appropriate conditions of instructed actions provides another point of similarity between T-instructions and beliefs. For just as the success (truth, satisfaction) of a belief that an apple is edible depends on whether the apple is in fact edible, so the instruction to EAT-[that-apple]-[THUS] is appropriate iff the apple is in fact edible. Similarly, just as the truth of a means-end belief "if want Y do X" depends on whether the relevant X is really a means to the relevant Y, so the appropriateness conditions of T-instructions depend on whether the targets they incorporate into the potential actions they determine and thus instruct render those actions appropriate solutions. In other words, the success of a T-instructions depends on whether the T-input is an appropriate target, i.e., whether it is within the solution-type's domain

246

of application. Raquel's T-instruction "Eat-[apple]-[THUS]" is appropriate iff eating the apple (thus) would, if executed, solve a nutritional problem if she had one.[179]

Finally, T-instructions and beliefs seem to support the same sorts of counterfactuals. If Raquel believes that the apple before her is edible or would be a means to her end of eating (satisfying her hunger), and she's hungry, then she'll eat it if nothing stops her. The belief that the apple before her is edible positions or prepares her to eat it should she suddenly become hungry. If Raquel's belief is false and she eats the apple, she'll need to eat something else, since the apple won't satisfy her hunger. Likewise, if Raquel T-instructively represents the apple, then, then, if she's hungry, she'll eat it (assuming nothing stops her-- including an organizing disposition that's responding to a stronger need). The T-instruction "Eat-[apple]-[THUS]" positions or prepares her to eat the apple should she suddenly become hungry. If the instruction is inappropriate, Raquel will need to eat something later, since the apple won't solve her nutritional problem.

N-Instructions (Desires/Emotions)

N-instructions are constituted by N-input activations of instructive dispositions. As was the case with T-inputs, N-input activations constitute N-instructions by virtue of constituting partial activations of the full instructive disposition. In **Figure 9** below, Raquel's "Eat" instructive disposition is activated via its N-input sub-capacity/component by her hunger (the "rumble" in her tummy). Unlike T-instructions, N-instructions "motivate": the N-instruction Raquel receives

---

[179] It's possible for an instructive representation to involve a chronically defective T-input sub-capacity—a T-input sub-capacity that is either hyper- or hypo-sensitive. If a T-input sub-capacity is hyper-sensitive, it will frequently instruct actions with inappropriate targets. If it's hypo-sensitive, the instructive disposition of which it is a part will frequently fail to be T-activated. Chronic defects may be mechanical in nature, or they may stem from the influence of defective organizing dispositions. However, the mechanisms that realize application-accountability should in principle address such defects.

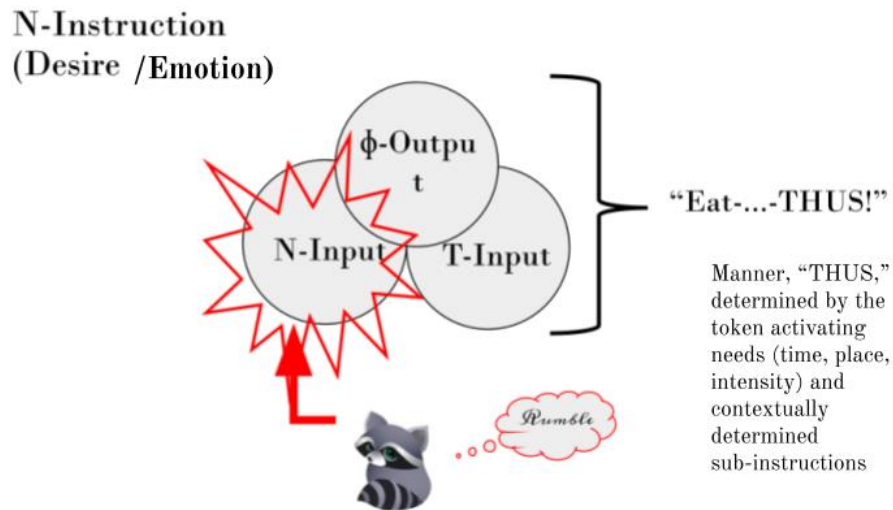from her stomach may be sufficient to "move" her to eat. N-instructions are thus comparable to desires.

Figure 9 Example of an N-Instruction

EIR instructed (potentiated) actions are often extended, meaning that they often involve a sequence of subroutines or sub-acts required to secure a terminal target. The context in which an N-activation (N-instruction) occurs determines the initial manner of the instructed action. This is the situation pictured in **Figure 9** above: Raquel receives the "Eat" N-instruction when there's no food (terminal target) in sight. Here, the instructed act of eating is individuated not only by the token need fact at which it is directed, but also by the context in which that problem is instanced, as the context contributes to determining the manner of the instructed action. The manner is a matter of the sub-actions [THUS] involved in securing a target for eating where those sub-acts will be guided by the targets available to the relevant sub-instructive dispositions in her initial context.

248

For example, Raquel's current context may activate sub-instructions to locomote in the direction of a nearby location where she's likely to find food. The context in which the N-instruction is instanced will not, however, fully determine the manner. Rather, much of the manner awaits the actualization of the action, as it will be determined dynamically. In short: N-instructions token-individuate the actions they instruct by determining the token need at which the action (solution) is directed and by partially determining the manner of the action, insofar as the context of the N-instruction determines the initial conditions for the action.

The primary way N-instructions contribute to the appropriateness conditions of the actions they instruct is by determining the need or problem instance at which the action is directed. It is possible for an instructive disposition's N-input sub-capacity to be activated by a "pseudo-problem" --something outside of the solution type's domain of application that activates the N-input sub-capacity. When this happens, the N-instruction is inappropriate, for if there's no problem (need) to be solved, the action won't be an appropriate solution.[180] So although N-instructions are desire-like insofar as they are motivating, they are also belief-like in that their appropriateness turns on whether they are sensitive to and thus direct us towards (need) facts of the right type.

Finally, N-instructions and desires seem to support the same sorts of counterfactual conditionals. Generally, if an agent desires X, she's likely to take the means ϕ to achieve X. Exactly what the probability is of her doing ϕ depends on what else she desires at the time, what beliefs she has, and the relative strengths of her desires. Likewise, if an agent has an N-instruction to ϕ, she's likely to (try to) perform ϕ. Exactly what the probability is that she will perform ϕ depends

---

[180] It's possible for a chronic defect in a N-input sub-capacity to cause an agent to regularly respond to instances of one type of problem as if they were an instance of a different type of problem. However, such a defect should in principle be addressed by the mechanisms that realize application-accountability.

on the T-instructions available to her as well as the other N-instructions she has and on their relative strengths[181] (together with the modulating influence of organizing dispositions).

i-instructions (pre-performance intentions) & I-instructions (decisions)[182]

An i-instruction is realized by an ID that has been both N- and T-activated either by a terminal target or by its absence so long as its absence N-activates sub-instructions that are themselves T-activated (**Figure 10**). So typically, an N-instruction will also be an i-instruction. An i-instruction is locally appropriate iff its corresponding N- and T-instructions are appropriate.
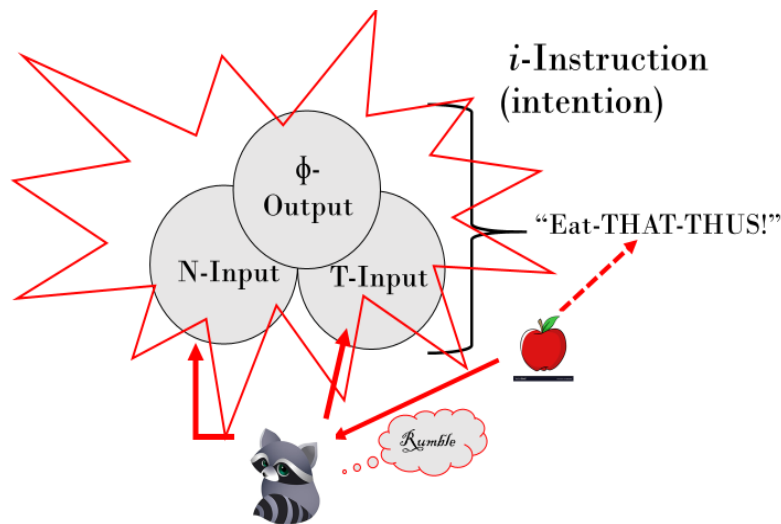

Figure 10 Example of an i-Instruction

I-instructions (**Figure 11**) are just i-instructions that have been selected via ODs. Since an I-instruction reflects the influence of ODs, it is evaluable not only in terms of local appropriateness but also global appropriateness.
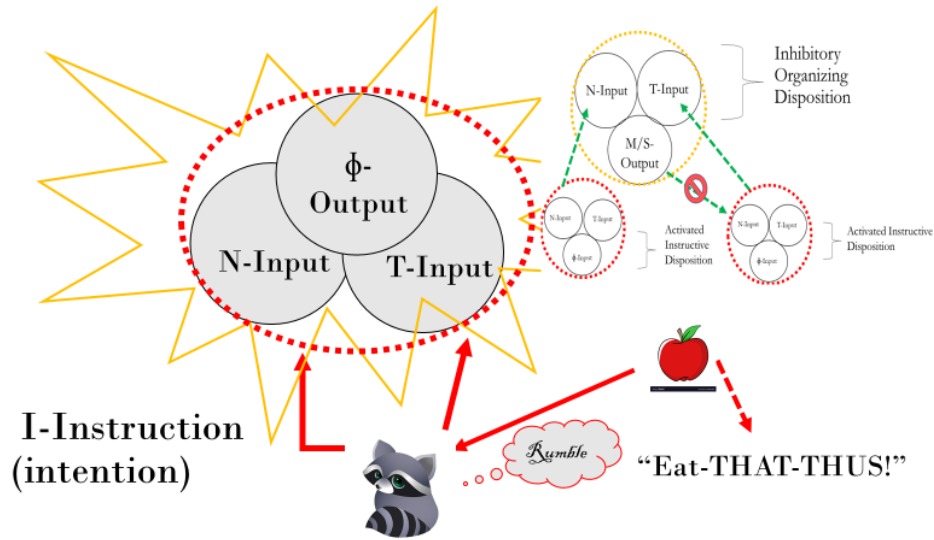
Figure 11 Example of an I-Instruction

EIRs & SRs

As noted in Chapter 2, practical problems can be basic or derived and they can involve different aspects of flourishing. Accordingly, EIRs can instruct all sorts of actions -- they can instruct as many types of solutions as there are types of practical problems. An EIR may instruct an act that essentially involves SRs. This occurs when:

(i) an EIR involves an SR as target (T- inputs), i.e., when it generatively represents a response to an SR *as such*, and/ or when:

(ii) an EIR generatively represents an action that essentially involves the *production* of an SR.

When (i) occurs, the appropriateness of the EIR depends on the contents of the SR; whether the instructed action is appropriate depends on what the content of the SR is. When (ii) occurs, the appropriateness of the EIR depends on the appropriateness of the SR. I suggested in the previous chapter that SRs solve proximity problems only if they are "true to" "accurate of" "satisfied by" "correspond to" or "are sufficiently similar to" to the targets for which they stand-in. If this is right, then EIRs that consist in or essentially involve the production of SRs are appropriate iff the SRs they instruct one to produce are true (veridical, satisfied, accurate, etc.), where truth (etc.) must always be evaluated in the context of the larger practical problems they function to help solve. In the next section, I'll suggest that some EIRs that involve SRs in these ways are propositional attitudes in a full-blooded sense.

Table 3 Embodied Instructive Representations

Instructive Attitude	Comparable Propositional Attitude(s)	Corresponding State of Instructive Disposition	Misrepresentation
T-Instruction	Belief (esp. means-end); Perception	T-Input activation	T-activation by *inappropriate target* s.t. φing-*THAT-THUS* would not non-accidentally solve an instance of the relevant type of local problem.
N-Instruction	Desire; Emotion	N-Input activation	N-Activation by *pseudo-need* s.t. φ-THAT-THUS would not non-accidentally solve an instance of the relevant type of local problem.
i-Instruction	Intention (prior to decision)	N- & T- Input activation	*Either* T-activation by *inappropriate target or* N-activation by *pseudo-need or both* s.t. φ-THAT-THUS would not non-accidentally solve an instance of the relevant type of **local** problem.
I-Instruction	Intention (post decision)	N- & T-Input activation $+$ organizing disposition activation	Not a matter of misrepresentation but rather a miscalculation (may result from misleading "force" ("!") of EIR); the EIR that's selected by organising dispositions does not solve the relevant ***global problem*** (promote overall flourishing in the context).

6.2 Subsuming Propositional Attitudes

I have argued that EIRs have properties that render them analogous, in certain important respects, to beliefs, desires, and intentions. In this section I want to argue that propositional attitudes may be understood as a special class or sub-species of EIRs. First, let me clarify what I mean by this. I do not mean that we can recover propositional attitudes, in special cases, merely in

the non-substantive, minimalist sense that EIRs can be interpreted as propositional attitudes, or that they entail sentences that we treat as true that ascribe propositional attitudes. Rather, I want to distinguish cases in which what we call propositional attitudes are just EIRs to respond to the facts that *we* (humans, theorists) represent (or might represent) by means of certain propositions (SRs) from cases in which EIRs involve genuine SRs. EIRs are intentional attitudes or representations of facts generally. They're attitudes towards propositions only when certain conditions are met.

A first condition is that a vehicle that is in fact an SR may function as a target of an instructed action. However, this is not sufficient for an EIR to count as a propositional attitude in the substantive sense I'm after, for this may occur without any comprehension of the surrogative representation as such. For instance, a bird may intentionally fly away from a sound produced by a human, where the sound produced happens to be a sentence (SR). This is a quite trivial way of having an attitude to a surrogative representation or proposition. The substantive conception of propositional attitudes I'm after involves something far less trivial. For agents who know how to use SRs, they may figure as targets of intentional actions *as such*-- for agents may act on them in solving proximity problems. And agents may act on them in solving proximity problems either by producing them for this purpose or by responding to SRs that were produced by others for this purpose. So, where a bird flies left as a result of hearing a man shout "Go left!" simply to get away from the alarming sound, such that the content of the SR is only accidentally related to its action, a human may turn left precisely because those sounds have the relevant type of spatial significance for her by virtue of activating leftward-orienting instructive dispositions.

However, I believe that this is still insufficient for an EIR to count as a propositional attitude proper. This is because humans are not the only animals that employ SRs. For instance (to borrow a favorite example of Millikan's), a beaver's tail flap stands-in-for the input to the EIR to

254

flap the tail (danger signal) for the recipient beaver insofar as it serves as an input to the RUN/AVOID/ TAKE COVER EIR(s) for the recipient (and directs them roughly to location of the input to the sender's EIR). The flap solves a proximity problem of helping fellow beavers/ socially coordinating protection-- it enables beavers to take cover from something they haven't themselves directly sensed. Humans no doubt also employ SRs like this. Accordingly, the EIRs that are involved in producing and responding to these sorts of SRs are not the kind that distinguish human agency and cognition. While beaver tail flaps count as a kind of shared perspective in the sense that the beaver that seems to perceive danger produces the flap as a way of alerting conspecifics to that danger and in that way shares its practical understanding or interpretation with them, beavers do not respond to the fact that it is understanding or interpretation.

The sorts of SRs that individuate the EIRs that are properly considered propositional attitudes stand-in-for states of affairs (facts) in a more robust way. An example is a scientific model or photograph. Here the point of the stand-in is to enable a kind of indirect interaction with the relevant states of affairs that depends on standing-in-for/ conveying more than time and place-- other features (more intrinsic to the fact). What gives it this point is the fact that humans are engaged in genuinely epistemic activities, the aim of which is something along the lines of the improvement and growth of understanding. For this reason, I believe that propositional attitudes proper are EIRs and SRs that are involved in critical activities-- activities of testing, evaluating, asking after the justification for, seeking understanding. We are not just interested in activating EIRs, but revising and modifying our EIRs, acquiring new and improved EIRs, and to do this we must develop SRs that stand-in-for the facts to which our EIRs need to be (better) fitted.

Consider another example: A parrot learns to produce the sound "avocado" when presented with an avocado. EIRs (activations of IDs) account for the sense in which the parrot's action of

255

producing the sound "avocado" when presented with an avocado is a manifestation of a kind of understanding on the part of the parrot (the parrot, unlike a mere machine, does not just have what Brandom (e.g., 1994) calls " reliable differential responsive dispositions" or RDRDs): the avocado activates an ID to produce the sound as a way of getting a treat or some positive attention that satisfies a need (for food or positive attention). So, the parrot understands the avocado she sees in that she's positioned to produce this sound as a way of acquiring food or comfort or some such. In other words, the action of producing the sound is just a type of eat-type solution or soothe or socialize type-solution (albeit highly unusual for parrots qua wild animals). The production of the sound does not serve any descriptive or ontologizing purpose for the parrot. This is so even though it might solve a kind of proximity problem in that the parrot, being in a cage or whatnot, must produce certain sounds in order to get certain things from her humans. By producing the sound, she makes public a need (or what she interprets as a need) for a human-mediated interaction-- a need to eat some food that, in those circumstances, only her humans can provide, or the need for some social engagement that, in those circumstances, only her humans can provide. The sound stands-in-for the parrot's need insofar as she produces it as part of an indirect (human-mediated) solution to her need. But the sound is not, for the parrot, a descriptive classification of her need. Her need is simply understood in terms of the EIR to produce the sound.

In contrast, the human who has a genuine propositional attitude about the avocado-- who can be said to believe that it is an avocado, that it belongs to avocado *kind*-- is a human who at least potentially grasps the avocado by means of a cluster of IDs for producing and responding to a variety of SRs which are related to each other in particular ways by particular ODs. For instance, a human who can genuinely believe that something is an avocado in the relevant sense plausibly has to have an ID to point or gesture at a visually perceived object (such as an avocado) and ask

256

what it is. In other words, an avocado must be able to figure as the T-input to an ID for solving the problem of needing more information about something (perhaps for the sake of solving some larger problem); avocados must be able to function as the targets of acts of inquiry.[183] Relatedly, I think it is plausible that one must have an ID for producing the label AVOCADO in response to an inquiry in which someone points to any (perhaps of a familiar variety) avocado in standard viewing conditions and asks what it is. Moreover, activations of IDs to apply the label AVOCADO must inhibit the activation of IDs to apply incompatible labels to the same target. I think it is also plausible that one who properly believes *that* something she sees *is an avocado* must also have IDs to apply other descriptive labels to avocados, including: GREEN (as opposed to RED, YELLOW, etc.), OVAL (as opposed to other shapes), FRUIT/ VEGETABLE/ PLANT (as opposed to ANIMAL or ROCK etc.); [RELATIVELY]SMALL/LARGE; EDIBLE. And the IDs to apply these other labels should inhibit IDs to apply incompatible labels. And these IDs to apply descriptions of avocados (or to draw or point to pictures of them) should be activated by someone asking what they look like and by someone requesting one to explain why one has attributed the label AVOCADO to something any time one does so. And, of course, all of the IDs mentioned must be modifiable if they are in fact exercises of abilities for application-accountable solutions to practical problems. This, I believe, at least goes some way towards capturing the intuitive idea that beliefs or commitments to propositions-- to things being some way, e.g., to an object belonging to avocado kind-- impose a rational requirement to update beliefs or commitments in response to new (relevant) information. In short, the propositional attitude of belief corresponds to EIRs that belong

---

[183] Why think this? It seems to me that the kind or ontological category AVOCADO is essentially an epistemic category -- a construct of inquiry.

to such complex and interconnected clusters of IDs (and ODs) and that are involved in solving particular sorts of problems-- problems characteristic of existing in the space of reasons.

The proposed EIR psychology will account for full-blooded propositional attitudes of desires and intending in a way similar to its account of full-blooded beliefs. A desiring propositional attitude might correspond to a need and/or need-activated ID that a subject can represent to herself, via an SR, as a need, that is, as something that she has an ID to employ in practical reasoning in that role, and something that she has an ID to employ in justifying why she does something that would be justified by that need. I mentioned earlier that what are called "prior intentions" either correspond to i-instructions that are not yet I-instructions, or to plans. Like propositional attitudes, I posit that plans are constituted by sets of EIRs that are governed (or controlled) by commitment to some future-directed SR, i.e. an SR that stands-in-for a possible fact one is committed to (has various dispositions for or is in the process of acquiring dispositions for) bringing about.[184] To be clear, I do not intend for these to be final accounts of the propositional attitudes, but merely sketches of what I think they're likely to look like in an EIR psychology.

What about propositional attitudes other than beliefs, desires, and intentions? For instance, what about hoping, wishing, wondering, entertaining, fearing, and imagining? What unifies clusters of IDs that are propositional attitudes on this view is the way in which they involve SRs in fundamentally rational enterprises. Not every ascription of a hope, wish, etc. will necessarily correspond to a genuine propositional attitude. Non-human animals might have some N-instructions that are hope- or wish-like. But paradigmatic hopes and wishes are self-consciously

---

[184] I think plans also essentially involve what Andy Clark (1997, 2008) calls "scaffolding." More generally, I think distinctively human cognition may be essentially extended--at least partly constituted by political and social institutions and myriad products of human labor.

ascribable; they correspond to ways self-conscious agents have of classifying their own varieties of IDs and EIRs in relation to SRs such as plans. And these classifications may well be vague or at least so thoroughly context sensitive that no particularly informative general analysis of these attitudes can be given. But to the extent that it can, I believe that can be done in terms of clusters of IDs and ODs, where the relevant IDs will be for solutions to broadly rational types of practical problems and thus have certain characteristic sorts of inputs and outputs. One might object that this shows the standard (non-EIR) propositional attitude model has an advantage over the EIR model insofar as it offers a simpler account of propositional attitudes. But simplicity is only a virtue where the target phenomenon is simple, and in the case of the attitudes, it isn't obvious that it is. Ascriptions of propositional attitudes seem to entail behavioral dispositions (even if they never get exercised), and which dispositions they seem to entail appears to depend heavily on context and in particular on what you're doing, or what it is you're trying to do.[185]

Moreover, you might recall that in chapter 1 I mentioned that ascriptions of hopes, fears, wants, moods are often of the form "Sally hopes/ is hoping for a raise," "Kevin is afraid of dogs," "Jess wants a dog," "Louise is anxious about her new job." In other words, they don't employ that-clauses, but on their face, they nonetheless describe intentional relations between subjects and objects. Objectualists appeal to such attitudes to resist the view that intentional relations are between subjects and propositions representing objects and to support the view that intentional relations are between subjects and objects. On the view I'm proposing, hopes, fears, wants, and moods correspond to forms of direct-PEA to facts; they correspond to IDs that are activated by

---

[185] This point has been recognized in particular for ascriptions of knowledge and has given rise to various ways of incorporating pragmatic features of context into the semantics of such ascriptions.

and thus directly directed at certain kinds of facts, like the fact of there being a dog. But where humans have such attitudes, they may be properly considered propositional attitudes. And this means that hopes, fears, wants, etc. correspond to clusters of IDs that include IDs for responding to propositions and producing propositions. So, an EIR psychology can both accommodate a rich sense in which human agents can have genuinely propositional attitudes of hoping, fearing, etc., while at the same time it can accommodate the objectualist's observation that these are not fundamentally attitudes towards propositions. In contrast, the standard (substantive) propositional attitude model appears to be committed to the view that psychological attitudes consist in attitudes towards propositions or something very similar.[186]

But what about anxiety? On the one hand, humans may be uniquely capable of some forms of anxiety by virtue of our ability to produce SRs that stand-in-for possible states of affairs. Our anxiety, in some cases, might be explained by EIRs for responding to such SRs. On the other hand, some forms of anxiety might be explained without appeal to SRs. This is perhaps especially the case for the sorts of anxiety that can seem to lack any intentional object. An EIR psychology suggests that this lack of an object is an illusion brought about by a belief-first psychology that, in effect, focuses its attention on T-inputs at the neglect of N-inputs. In contrast, the EIR psychology is an intentions-first psychology, and the heart of intentions are N-instructions, since needs are already teleologically directed at types of objects, while T-instructions should be understood as components of or partial intentions (and thus partial or incomplete representations). Accordingly,

---

[186] As I argued in Chapter 1, although some philosophers who defend the propositional attitude model deny that we should think of propositions as the direct objects of the attitudes but rather as things in terms or through which we represent facts directly, such views appear to be unstable.

the EIR psychology accounts for moods that seem to lack an intentional object in terms of N-instructions in the absence of terminal T-instructions.

To see how this works, consider the example of free-floating anxiety. It is highly plausible that anxiety corresponds to activations of our fight-or-flight IDs and sub-IDs. The facts that are N-inputs to such IDs are often also T-inputs (the same typically goes for the EIRs that correspond to pain), or at least the T-inputs are not terribly remote from the source of the N-inputs. But like any EIRs, fight and flight EIRs can misrepresent. Indeed, it's easy to see how this might happen as a result of modifications to fight and flight solution application algorithms, since when it comes to these responses it is typically better to err on the side of caution (false positives are typically less detrimental than false negatives). But certain learning environments may result in application algorithms that are easily biased-- that tend to over-apply the fight and flight solutions. As a result, a subject will suffer a number of inappropriate fight and flight (and associated) EIRs. In particular, their fight or flight IDs will be need-activated by pseudo-dangers (pseudo-problems). When there's a pseudo-problem, there's likely to be no obvious terminal target-- no obvious thing to fight or from which to flee. In general, when an ID is need- but not terminal target- activated, the EIR will instruct an action whose manner begins with some sort of search-- in the context of fight and flight responses we might call this vigilance. In addition, when fight or flight EIRs are inappropriate because activated by pseudo-needs/problems/dangers, there often will not be any terminal target the fighting of or fleeing from which will result in the elimination of the need and thus the cessation of the EIR. In other words, there will be no (obvious or direct) way to complete the instructed action, so one's free-floating anxiety will tend to persist, and it will feel to one to be objectless.

Finally, one might object that EIRs can't subsume propositional attitudes because propositional attitudes are uniquely fitted to figure in useful models of psychological reasoning.

261

However, as we began to see in the last chapter, problem-solving dispositions can realize certain forms of reasoning. Consequently, the sorts of psychological models in which propositional attitudes figure might be translated into the processes involved in the production and selection of EIRs and sub-EIRs.

First, the activation of an ID may be modeled in terms of an Aristotelian practical syllogism: the premises are "apparent" needs (instructive desires) and targets (instructive means-end beliefs), where "apparent" just means that some facts have activated N- and T- inputs to an ID, and the conclusion is the potentiation of a token of the type of *local* solution (action, activation of Φ-output sub-capacity) the ID is an ability for. Such reasoning is successful iff the conclusion is *locally* appropriate, i.e., if it practically fits the actual (rather than apparent) facts, i.e., if the premises are sound (if the activating inputs were appropriate).

Second, the selection of potentiated local solutions (EIRs, activated IDs) for performance may likewise be modeled in terms of an Aristotelian practical syllogism, only this time the premises are EIRs (activated IDs that are inputs to ODs), and the output is the selection of the winning EIRs, where such reasoning is successful iff the conclusion is *globally* appropriate.

Finally, sub-EIRs (sub-IDs) may be modeled in terms of an Aristotelian practical syllogism, only their need-inputs will be activated IDs, while their target inputs will be context-specific facts, and their outputs will be modifications of the Φ-output of the overarching activated ID (EIR) in response to a dynamic succession of targets. Sub-EIRs may also be modeled in terms of instrumental or means-end reasoning, as an activated overarching or super-ID will, together with the context in which it is activated, recruit sub-IDs (sub-EIRs) as the *means* of performing the overarching instructed action.

Propositional attitudes proper are clusters of IDs that include IDs for using SRs that stand-in-for EIRs or facts that are inputs to or outputs of ID where those SRs can be used in intentional actions or activities of reasoning. In other words, on the proposed view, we must distinguish between two kinds of reasoning. There's a kind of reasoning that is itself an intentional action and therefore something EIRs can instruct us to do, and something we might need to do to activate or bring into existence certain EIRs, but not something that needs to happen to create EIRs in the first place. Then there's a non- or proto-intentional kind of reasoning process-- the process of activating an ID, and thus generating an EIR, and the process of selecting IDs via ODs. This latter, non- or proto- intentional process can be modeled in terms of the intentional reasoning process. But they are importantly distinct kinds of processes. Non- or proto-intentional processes are a necessary antecedent to any intentional actions, while the intentional process, or at least its availability, is plausibly a condition for distinctively rational actions.

Propositional attitude models fit into classical computational models that are used in the psychological sciences. Since EIRs are realized by problem-solving dispositions, connectionist and neural net models are a natural fit for an EIR psychology. But an EIR psychology may also provide a kind of bridge between classic computational models and connectionist models. Problem-solving dispositions are hierarchically organized into types, and classic computational models may usefully describe the nature of the processes that occur between stages. For instance, we've seen that the instructive dispositions activation of which realizes EIRs can be functionally decomposed. This means that traditional practical reasoning-based psychology should still have application in an EIR psychology (see also Chapter 4). In addition, the relationship between EIRs (IDs) and the sub+-EIRs (sub-IDs) they (often dynamically) recruit in determining the particular manner of an instructed action is broadly compositional. The rules of composition for EIRs and

sub-EIRs derive from means-end relations (see also Chapter 4). ODs account for another way in which EIRs are compositional: ODs can dynamically control the performance of EIRs and may combine or interweave the execution of different EIRs in accordance with an agent's changing global needs. Finally, ODs *operate over* IDs and hence EIRs in a way that's sensitive to their embodied instructive representational contents and their logico-practical relations to other such contents (see also Chapter 4). In this way, EIRs have properties that allow them to be understood representationally even from a fairly classical computational perspective.[187]

6.3 Propositional Attitude Problems

In this section I'll argue that the propositional attitude model of psychological representations faces a number of difficulties. I have grouped these difficulties into three types. The first type of difficulty arises in the context of trying to provide an illuminating, broadly naturalistic account of propositional attitudes. The second type of difficulty concerns a tension between providing an illuminating account of propositional attitudes and providing an illuminating account of intentional action. The third type of difficulty has to do with the explanatory power of propositional attitudes. I will focus on teleosemantic approaches to illuminating representation because they're often thought promising and because my positive view shares features with it.

---

[187] Elsewhere I have begun to explore how an EIR psychology can accommodate the sorts of things that are thought to be hard for non-traditional representational pictures to accommodate, in particular, the productivity of thought.

The First Difficulty: Illuminating Psychological Representations

In the context of trying to provide an illuminating, broadly naturalistic account of propositional attitudes, the propositional attitude model faces three challenges: "The Grounding Challenge," "The Alethic Triviality Challenge," and "The Homunculus Challenge."

The Grounding Challenge arises in the context of providing an illuminating account of representation. An "illuminating account," as I understand it, does not necessarily reduce representation into non-intentional concepts. Rather, it reduces *representational intentionality* in a naturalistically acceptable way without going in a circle. On the propositional attitude model, this means grounding a fundamental form of semantic content (truth, veridicality, accuracy, or satisfaction conditions) naturalistically. But attempts to do this often ground only "deflationary" notions of representation, i.e., notions lacking some of the *prima facie* essential features of representation such as the normativity of content. Teleosemantics appeals to practical (broadly biological) functions to ground the normativity of content and is widely considered a promising strategy for grounding original representation. However, Teleosemantic accounts face what Ramsey (2007) calls "the job description challenge." This is the challenge of specifying the right kind of teleo-function, since clearly not just any biological functions will do- not if we're to avoid deflating the notion of representation. Attempts to address the job description challenge result in two further challenges.

First, as Burge (2010) points out, non-trivial alethic success conditions do not simply fall out of practical norms. So, because Teleosemantic accounts attempt to ground representation in information states with practical functions, they face the **Alethic Triviality Challenge**. Importantly, to question whether non-trivial alethic success gets a hold on a purported representation is to question whether genuine attribution is occurring. So where the descriptive

265

thesis is assumed, Alethic Triviality is a problem for establishing the very presence of representation (at least of a "robust" variety).[188]

The **Homunculus Challenge** arises out of a common strategy for meeting the Alethic Triviality Challenge by squeezing alethic conditions (and thus genuine attributive functions) out of practical norms. This strategy appeals to a sub-individual-level version of a Lewisean signaling game, i.e., representation "senders" or "producers" and representation "receivers" or "consumers."[189] The idea is that what's sent between producer and consumer systems doesn't merely causally co-vary with, carry information about, map to, or correspond to something, but has a function in the system the fulfillment of which depends on its bearing that property. Not only that, but the function is properly considered representational in that it involves something like communication between sender and receiver and interpretation at least on the part of the receiver.[190] However, a communicative function just seems to be another name for a representational function. So, claiming that subsystems that can be modeled in terms of Lewisiean signaling games thereby have communicative and therefore representational functions seems to beg the question. At the same time, it's far from obvious that subsystems have genuine communicative functions, since genuine communication is arguably something that takes place between agents. Coordinating sub-systems are not literally agents, we merely model them as such when we employ Lewisian signaling games to describe and predict their activity. Similarly, the idea of interpretation, as Dennett (1978) points out, appeals to a homunculus and in turn threatens to generate a regress:

---

[188] Burge (2010), Springle (2019a), Orlandi (2014), Ramsey (2007). Teleosemantics also faces other challenges I won't address here.
[189] See e.g., Millikan (1984, 2004).
[190] See e.g., Millikan (2004).

Any representation or system of representation…requires at least one user or interpreter of the representation that is external to it. Any such interpreter must have a variety of psychological or intentional traits…: it must be capable of a variety of comprehension, and must have beliefs and goals (so it can use the representation to inform itself and thus assist it in achieving its goals). Such an interpreter is then a sort of homunculus... (122)

A theory that appeals to homunculi with intentional states owes us an explanation of their intentional states, and if it employs the same strategy to analyze the intentional states of homunculi, it generates a regress. The obvious way to avoid homunculi is by denying that there is any actual communication or interpretation going on at the sub-personal level. However, deflating the sense in which "senders" and "receivers" communicate or interpret natural signs (correspondences) deflates the representational functions-- attributive functions and alethic success conditions--those notions were supposed to buy.

In the face of these challenges, one option is to follow Burge (2010) and give up on the task of illuminating representation. [191] According to Burge, representation cannot be reduced to something non-intentional, and intentionality as such is not metaphysically suspicious. Burge might be right on both scores (indeed, I think he is right on both scores). However, it doesn't follow that representations don't cry out for an analysis nor that representation can't be analyzed in terms that don't mention or presuppose representation. Below I'll argue that we don't have to give up on the task of illuminating representation if we instead give up the descriptive thesis and make room for EIRs.

---

[191] See e.g. Ramsey (2007), Hutto and Myin (2013).

The Second Difficulty: Explanatory Power & Structural Constraints

Although non-human animals and human infants are obviously equipped with some cognitive capacities, especially those involved in basic perception and action, many philosophers and scientists are doubtful that they are capable of genuinely propositional thought. This is not to say that their psychological states cannot be stated in terms of attitudes to propositions. In other words, the psychological states of non-human animals is not a problem for non-substantive conceptions of the propositional attitudes. But it is a problem for substantive conceptions that propositional attitudes seem not to apply to nonhuman animals (it's also questionable whether they apply to human infants). Call this **The Non-Human Animal Problem**.

In recognition of this problem, a number of philosophers have developed alternative accounts of psychological representation to accommodate non-human animals, though they're typically also intended to apply to some human psychological states, in particular, perceptual states. For instance, Ruth Millikan discusses "two different kinds of cognitive capacities that we humans seem to have, each of which is at least akin to rationality as Aristotle described it" (1). The first, she believes, "we share with many other animals," while "the second perhaps with none" (1). Millikan suggests that human and non-human animals share a kind of perceptual rationality that's "for immediate use in practical activity," and she claims that "if perception involves concepts, they are in the first instance practical concepts, repositories for procedural rather than propositional knowledge, for storing know-how rather than factual information." According to Millikan, perception involves exclusively "pushmi-pullyu representations" (1996, 405) – representations that are undifferentiated between being indicative and being imperative, between

describing and directing," (4) and for which "the criterion of correct recognition of affording objects or properties lies in practical successes" (10).[192]

Propositional attitudes appear to be unable to accommodate such contents, and the absence of a clear alternative to propositional content makes it hard to make sense of them. For propositional attitudes are the mental analogues of declaratives (or indicatives) and imperatives (or commands). What distinguishes a belief from a desire, or a declarative sentence from an imperative sentence, when these all involve the same propositional content (as they may) is the attitude or force, respectively, attaching to the content. Attitude or force correspond to the way a proposition is being used. So, use and meaning (content) are distinct and declaratives (beliefs) are concerned with being true while imperatives (desires, intentions) are concerned with becoming or being made true. This idea is captured in the notion of a direction of fit (Searle 1980). In the case of an indicative sentence or belief a descriptive content is supposed to fit the world, so it has truth conditions. In the case of an imperative or desire the world is to be made to fit a descriptive content, so the imperative or desire has conditions of satisfaction—it comes to be satisfied when its content comes to fit the world.

The problem with Millikan's pushmi-pullyu representations (and similar proposals, e.g., Streleny 2003) is that they seem to conflate force and content.[193] For while it may be denied that the relevant content is properly or strictly speaking propositional, it's assumed that it's nonetheless

---

[192] Millikan (2017) distinguishes between two kinds of "unicepts"-- the entities with which she proposes to replace psychological concepts: (1) "affording unicepts" which humans and animals share and (2) "substance unicepts" and "attributive unicepts" which are unique to human and correspond to our abilities for attributing properties and membership in kinds and for propositionally articulated thought, that is, the kind of thought that's subject to negation. Although it isn't always clear exactly how unicepts relate to representations, I believe she'd say that exercises of affording unicepts constitute pushmi-pullyu representations while exercises of substance unicepts constitute propositional attitudes, or what she calls "factic representations." In earlier work (e.g., her 2004), she appears to refer to such representations as "dead facts."

[193] See e.g., Stalnaker (2012).

very much like propositional content insofar as it's descriptive-- it's essentially a matter of having truth or satisfaction conditions. For what else could it be? On the standard picture, what distinguishes indicatives from imperatives is not the kind of content they involve but their force. While Millikan is clear that pushmi-pullyus are neither purely indicative nor purely directive, it's far less clear how they can be both. Indeed, standard models of reasoning appear to depend on the modal separability of propositional contents from force (or attitude). Accordingly, it is not clear how pushmi-pullyus could figure into reasoning.

On the other hand, the essentially practical nature of pushmi-pullyu representations lies in the fact that their indicative aspects *don't* come apart from their imperative aspects. For in having these aspects united so as to put the push and pull (somehow) right into the content, such representations are sufficient to determine an action—they're all a creature needs to know how to respond to its environment. In this way, they differ from the "detached" indicative and imperative representations rational animals use in deliberative reasoning. The detached indicative (or pure belief) floats free of the detached imperative (or pure desire); it is what Millikan (2006) calls a "dead fact"—only potentially practically useful, and only indirectly, as it must join with an imperative to determine an action.

The essentially practical element of pushmi-pullyu representations is thus lost if, in light of the conceptual difficulties accompanying the notion of a pushmi-pullyu, we retreat to "practical" (Stanley 2011, Pavese 2015, 2017, 2019, 2020, 2021) or "pragmatic" (Nanay 2013) representations understood as indicative representations whose contents attribute "practical" (and often subjective or organism-relative) attributes or predicates (e.g. "edible" or "climbable") as opposed to, e.g. natural kind predicates or other predicates corresponding to less immediately practical properties. For while such predicates describe or are about either properties that happen to be practically useful

or relations that are essentially practical, they lack the "push" that's associated with the force of an imperative—the push that's required for action. Nor does it make sense to combine descriptive contents that are about practical or practically relevant properties with an imperatival force, as we do not think that perceiving a nut tells a squirrel "Make it the case that that [the nut] is edible." Perhaps this is why Millikan (1984, 1996) falls into describing pushmi-pullyu representations as mere indicatives whose contents attribute affordance properties like edibility. Call this **The Primary Pushmi-Pullyu Problem**.

There's also a **Secondary Pushmi-Pullyu Problem**. According to Artiga (2014), teleosemantic theories of original intentionality are committed to the view that "simple" representations" -- that is, the most fundamental kind of representations-- are pushmi-pullyu representations. Artiga argues that if we treat their commitment to this view as a core tenet of teleosemantics, then together with their other core tenets, teleosemantic theories entail that "necessarily, all representations are Pushmi-Pullyu states."[194] The issue Artiga raises appears to be related to the Alethic Triviality Challenge in that the problem is that it isn't clear how producers of purely descriptive representations would have evolved. This is because pushmi-pullyus seem sufficient to do the work required by the biological functions that ground original representations. According to Artiga, teleosemantic accounts of representation have yet to provide an adequate account of how decoupled representations-- representations that have only indicative or only imperative contents-- emerge from pushmi-pullyus. Consequently, teleosemantic theories fail to provide an account of the sorts of representations they set out to naturalistically illuminate.[195]

---

[194] In Millikan's (2020a) reply to Artiga, she disputes some of the claims Artiga identifies as core tenets (though not the claim that simple representations are pushmi-pullyus).

[195] Shea (2018) develops a teleosemantic account for cognitive science. In Chapter 7, Shae explains how exploitable relations can play a descriptive role, a directive role, or both, depending on the details of the representation's relation to a set of conditions in explaining a task

As the Secondary Pushmi-Pullyu Problem brings out, Millikan's pushmi-pullyu representations are not only motivated by her desire to account for non-human animal mental representation, but also by her analysis of original intentionality. For Millikan's account of intentionality grants a significant theoretical role to the way a representation is used by a system in the production and guidance of (broadly) intentional actions.[196] This is why the most fundamental representations are "undifferentiated" pushmi-pullyus.[197]

In a similar vein, David Papineau (1984, 1987, 1990 and 1993) develops a "consumer-" or "benefit-side" teleosemantic account of original representation that's thought to be "top-down" in that it applies most directly to whole propositional attitudes. Put another way, Papineau's theory is a kind of use theory-- content is determined by functional use and not mere causal or mapping relations. Papineau defines the truth conditions of a state as those circumstances that guarantee successful action, or successful results. He then defines successful results as those that satisfy desire. Thus, the truth conditions of belief are defined in terms of the circumstances that guarantee the result that is the satisfaction of some desire. (70-76) On Papineau's account, the contents of desires are primary; belief contents are derived from them. Papineau proposes that a desire's satisfaction condition is "… that effect which it is the desire's biological purpose to produce" (1993, 58–59). By "biological purpose" Papineau means "[s]ome past selection mechanism has favored that desire — or, more precisely, the ability to form that type of desire — in virtue of that

---

function. For Shae, descriptive (indicative) representations refer to those cases where explaining a task function depends on a representation producing a certain condition while directive representations refer to those cases where explaining task functions depends on a certain condition already obtaining when the representation causes behavior. I do not discuss his view in the main text because he is explicitly concerned with sub-individual representations.

[196] Strictly speaking, Millikan counts that sub-individual process as representing. See e.g., Millikan (2020).

[197] And this is essentially why Matthen (2005) posits "coercive contents" as the most primitive sensory contents.

desire producing that effect" (1993, 59). So, for instance, a desire may have the function of bringing it about that we have food. Such a desire has the content "that we have food." This is because it was selected for bringing it about that we have food and because if it collaborates with a belief to cause us to grab an apple, the content of the belief is "that the apple is food" if our desire for food would only be satisfied by our grabbing (and eating) an apple if it's true that the apple is food.

Papineau's account of original representation is criticized for appearing to conflict with propositional attitudes. First, it does not treat belief and desire as fundamentally distinct and more basic than an intention-like state. Instead, it seems to suggest that the intention-like state is primary: fundamentally mental representations are action-producing and guiding dispositions that function to satisfy an agent's biological needs. In order to play this role, inputs to these dispositions must figure into the determination of the sorts of actions that are their outputs. Belief and desire seem just to be different ways of describing types of inputs to these intention-like dispositions. Second, Papineau's account of original representation seems to collapse the attitude (force)/content distinction insofar as it is a top-down account of representation, i.e., insofar as his account treats of whole representations-- content and attitude at once.[198] So like Millikan, Papineau seems to be committed to a conception of fundamental representations that, like pushmi-pullyus, fail to conform to the structural constraints of propositional attitudes.

Philosophers have been inspired to posit pushmi-pullyu-like mental representations for reasons not essentially related to teleosemantics. For instance, a number of philosophers have

---

[198] Papineau's account is also criticized for seeming to conflict with the LOT hypothesis. I discuss how an EIR psychology might address this elsewhere.

argued that at least some kinds of perception essentially involve something like pushmi-pullyu representations (e.g., Nanay 2013, Siegel 2014, Watzl 2014, Mandik 2005, Grush 2007, Anderson & Rosenberg 2008.) In addition, a number of philosophers and cognitive scientists have argued that intentional action requires "motor representations" which are non-propositional in that they lack propositional format (e.g., Butterfill & Sinigaglia 2014, Levy 2017, Mylopoulos & Pacherie 2017.) But the assumption that intentions (beliefs and desires) are propositional attitudes has given rise to what's been dubbed the "interaction problem": the problem of accounting for the interaction between propositional intentions and non-propositional motor representations.[199] For motor representations and propositional representations are thought to have different "formats."[200]

Pushmi-pullyu-like representations have also been introduced to accommodate cases of inconsistent and "in-between" beliefs (e.g. Bendana & Mandelbaum forthcoming, Egan 2008, Gendler 2008a, 2008b, Schwitzgebel 2002, 2010, Mandelbaum 2013, 2016, Quilty-Dunn & Mandelbaum 2017, Norby 2014). These cases involve automatic or habitual belief-like psychological states that can be in tension with a person's explicit beliefs. Examples include the psychological states that underwrite implicit bias and irrational fear and disgust. Tamar Gendler (2008a, 2008b) has argued that to accommodate such cases we need to add "aliefs" (and perhaps "Cesires") to the traditional stock of psychological representation (propositional attitudes like beliefs and desires). As distinct from beliefs, aliefs have non-propositional "representational-affective-behavioral content." According to Gendler,

---

[199] Some philosophers have argued that the problem can be resolved by recognizing the ways in which intentions and motor representations aren't entirely different things. See e.g., Burnston (2017), Ferretti & Caiani (2019), Friland (2017), Brozzo (2017), Sutton (2019).

[200] As I understand the concept of "representational format," it is very close in meaning to that of a Fregean sense, though discussions of representational format pay a lot of attention to what might be called "vehicle dependent functional properties": the properties of a vehicle that appear to be required for a kind of representation's operational dynamics, or for the kind of role a type of representation plays in a system.

As a class, aliefs are states that we share with nonhuman animals; they are developmentally and conceptually antecedent to other cognitive attitudes that the creature may go on to develop. And they are typically also affect-laden and action generating. (641)

Pushmi-pullu representations are likewise states humans share with nonhuman animals, developmentally and conceptually antecedent to other cognitive attitudes that the creature may go onto develop. And to the extent that desires tend to be affect-laden, it's highly plausible that pushmi-pullyus-- since they're at once belief- and desire-like-- are typically affect-laden. It's certainly the case that they are action-generating. So, by Gendler's description, aliefs look a lot like pushmi-pullyu representations-- so much alike one might think they're just the same thing.

Speaking of affect-laden states, some philosophers have argued that (at least some) emotions should be understood as pushmi-pullyu representations. For instance, Srinivasan (2018) and Nussbaum (2001, 2016) argue that emotions are belief- or perception-like in that they can be appropriate or "apt"; they have a kind of representational, rationally evaluable content. At the same time, however, they are motivational. So, like pushmi-pullyus, they seem to simultaneously embody qualities of both beliefs and desires. Meanwhile, Scarantino (2014, 2015) explicitly suggests modelling emotions on Millikan's pushmi-pullyu representations.

So there appear to be a number of reasons to posit psychological representations that are non-propositional and something like pushmi-pullyus. But pushmi-pullyus appear to violate the structural constraints imposed by propositional attitudes. In addition, when we introduce a new kind of representational state, we're faced with the question of how they interact with propositional attitudes. So, we may be better off with an account of psychological representations that can accommodate the properties of propositional attitudes where they apply but that doesn't build those properties into the nature of mental representation generally. Below I'll argue that EIRs do just this.

6.4 Advantages of an EIR Psychology

I have argued that EIRs can subsume propositional attitudes. This is because, first, EIRs may be understood as intentions (i/I-instructions), beliefs (T-instructions), and desires (N-instructions), and second because EIRs can essentially involve SRs, and thus propositional contents, in the ways described in §1.4 above. I'll now argue that EIRs are preferable to a non-subsumed propositional attitude psychology by arguing that an EIR psychology avoids or addresses the difficulties for the non-subsumed propositional attitudes that were discussed in the last section.

The First Difficulty: Illuminating Psychological Representations

An EIR psychology addresses the three challenges for substantive conceptions of propositional attitudes. First, it addresses **The Grounding Challenge** because it is essentially connected to the PEA analysis which I argued in Chapter 3 is an illuminating analysis of representation. Importantly, the PEA analysis applies to the genus representation, and suggests that philosophers have been wrong to think of representational normativity essentially in terms of alethic success. Instead, the PEA analysis suggests that appropriateness is the standard of success that applies to robust representations generally, and alethic conditions are just one species of appropriateness conditions and have application only to surrogative representations, i.e., representations grounded in indirect-PEA.

Second, an EIR psychology sidesteps **The Alethic Triviality Challenge** because it rejects the descriptive thesis and thereby the need to ground non-trivial alethic conditions for original representations. On the PEA analysis, representational success is a matter of appropriateness, where alethic success is just one species of appropriateness. Far from trivial, such practical success

conditions are explanatorily ineliminable.[201] To be clear, then, while I agree with Burge that teleosemantic accounts of original representations don't ground non-trivial veridicality conditions, I disagree that this is a necessary condition for grounding robust representation. For non-trivial veridicality conditions (understood as a special species of appropriateness conditions) only get a hold on SRs, i.e., representations grounded in indirect-PEA, whereas original representations are EIRs which are grounded in direct-PEA. What constitutes a robust as opposed to deflationary representation is simply the function of providing PEA, direct or indirect.

Third, EIRs are intuitively robust representations since, by definition, PEA can exist only where intentional actions (qua solutions) can exist. And it does all this without positing any homunculi.[202] Indeed, the PEA analysis not only avoids **The Homunculus Challenge**, but also illuminates why accounts that assume DT tend to generate it: non-trivial alethic conditions attach only to SRs, and SRs depend on agents already having EIRs to produce (send) and respond to (receive) them (SRs in this sense have derivative intentionality).

One might wonder if the PEA analysis is susceptible to some of the concerns about content determinacy that have plagued teleosemantic accounts of representation.[203] Indeed, one might wonder how *facts* are understood on the proposed analysis, and in particular, how fine-grained they are. For instance, does "there's a rabbit in front of me" and "there's a rabbit stage in front of me" pick out the same fact? Similar questions might be asked of targets and needs: Is the target of the frog's tongue-snapping behavior flies or food? Does the frog need to eat flies? Or nutritious stuff? Is this the same need?

---

[201] See also Springle (2019a).

[202] Input systems do not and need not function as homunculi.

[203] Thanks to Frankie Egan for suggesting that I address this worry.

Let's begin with the reasons why indeterminacy is a problem for other teleosemantic accounts of original representation. There are two major reasons. The first is that they are trying to give an account of original representations. Where derived representations can refer to what they refer to because agents have designed them to do so, original representations are not designed by agents. Instead, they are supposed to be the kinds of representations that are essentially involved in agents' intentional activities generally, including the intentional activities that involve producing and employing derived representations. So teleosemantic theories have to account for how original representations come to refer "naturally," i.e., without agents making them do so.

The second major reason indeterminacy is a problem is that other teleosemantic accounts assume the truth of the descriptive thesis. That means that they need to account for "natural" representations with alethic success conditions. As I argued in Chapter 1, alethic success conditions entail a broadly descriptive representational element: something must function as a description, name, or label (or concept, attribute, or category) that properly applies to certain kinds of things but not others. When descriptions, names, and labels (or concepts, attributes and properties) are properly applied, they are true, veridical, or accurate with respect to those things, or those things satisfy the applied descriptions, or genuinely belong to the relevant concepts or categories, or genuinely instantiate the relevant attributes. When descriptions, names, labels, etc. are applied to the wrong kinds of things, they are false, falsidical, inaccurate, and not satisfied. So, in order to give an account of original representations with alethic success conditions, one must effectively give an account of original concepts (Papineau 1987).

Now, the problem is that the resources available to other teleosemantic accounts fail to yield determinate contents-- determinate ways of specifying referents. That is, they fail to establish the descriptions under which referents are represented. This is for several reasons. First,

278

teleosemantic accounts are two-stage accounts. In the first stage, they appeal to natural correspondence (or information or causal) relations to establish a kind of proto-reference relation, sometimes called an "indicator" relation. But such relations are far too ubiquitous to determine a genuine reference relation--if such relations were, fundamentally, all there was to reference, just about everything would refer to just about everything. Any state of affairs (or event) may indicate any number of other states of affairs (or events), so any state of affairs constitutes numerous indicators. There's also nothing normative about mere indication--nothing to make it the case that they are succeeding when referring to something or failing when they are not. In addition, indication is factive-- it is grounded in actual relations of correspondence. So, if representation were just a matter of indication, there could be no misrepresentation. Teleofunctions enter teleosemantic accounts at stage two in order to establish functional indicators. While, at any given time, any state of affairs constitutes any number of indicators, only certain indicators will play teleological (broadly biological) functions. Those that play teleofunctions will be determinate types of representational contents with alethic success conditions.

However, as Millikan (1991) notes, the idea of a functional indicator (or as it's sometimes put, the "function of indicating") has the potential to mislead. Indicators do not function to indicate, they just indicate. In other words, indicating isn't something an indicator functions to do, it is simply something an indicator does. What's needed, rather, are functional indicators--indicators that play some role in a system that depends on what they indicate. Teleosemantics suggests that indicators play certain roles by virtue of becoming involved in some broadly biological functional mechanisms.

Teleosemantic accounts suggests that the biological mechanisms in which functions to indicate arise consist in the coordination of certain kinds of subsystems, namely, "producers" (or

"senders") and "consumers" (or "receivers"). These subsystems work together in that the producers/ senders function to produce indicators that consumers/ receivers use to generate certain effects which are beneficial to both producers/ senders and consumers/receivers. The idea is that the consumer/ receiver only produces the beneficial effects when the producer/ sender produces a particular indicator. In other words, in a producer-consumer system, a producer can produce various indicators including an indicator I which indicates some condition C, and a consumer can produce some effect E which constitutes a beneficial response iff condition C obtains, and hence iff the producer produces I. In such a system, the producer has the function of producing I. Then, if the producer produces some indicator that isn't I, it misrepresents.

However, indeterminacy remains a problem for teleosemantic accounts because there are competing ways to specify indicator contents within this general framework. The easiest way to bring out the disagreement is by means of examples.[204] As a first example, consider the following:

> **Frogs:** Frogs have a visual-motor mechanism that controls their tongue-snapping response. The function of the tongue-snapping response is to catch prey, i.e. flies. But in addition to flies, bibis and other little darkly colored dot-shaped things that are moving horizontally across the frog's visual field will trigger the tongue snapping response. [205]

The idea is that the frog's visual system is the producer system, and its tongue-snapping motor response system is the consumer system. And the question is: what does the producer system function to indicate? Options include flies, food, moving black dot, etc. Which of these

---

[204] I ask the reader to bracket the issue of whether frogs (and below, "Biwis") genuinely represent anything. Most likely they do not-- their behaviors are not manifestations of application-accountable-solutions. These examples are valuable because they are common in the literature and they are simple examples, so they allow me to make the relevant points I want to make in a relatively simple way.

[205] These details probably are not entirely scientifically accurate. But this does not matter for my purposes.

descriptions specifies the relevant indicator (representational) content? According to benefit- or consumer-side teleosemantic theories, what determines the content is what determines the successful response. In other words, food or edible-fly or moving-black-dot. According to producer- side teleosemantic theories, the content should go with what the response is a response to, so fly, or moving-black-dot.[206]

If we say the content is "is a fly," what about a case in which a small edible thing that isn't a fly but that is nonetheless nutritious for the frog triggers its snapping response? Does it really make sense to say that's misrepresentation? On the other hand, what happens if a fly triggers the snapping response that isn't edible? Does it really make sense to call this a successful representation? On the other hand, if we say the content is "is food," that doesn't seem to pick out a sufficiently specific referent. Then what about the case of a black thing that triggers the snapping response that's still food? The most plausible content seems to be edible-moving-black dot. If so, then consumer/ benefit-side teleosemantic theories seem to have the better answer. However, sometimes their answer can likewise seem pretty counter-intuitive. Consider, for instance, the following example, which is a variation of Dretske's (1986) magnetosome example:

Biwis: Oxygen is toxic to an (imaginary) type of bacterium, call it Bibi. Oxygen-levels are higher closer to the surface of the ocean. There's also more light closer to the surface of the ocean. Since light and oxygen-levels correspond this way in the environment in which Biwis evolved, Biwis have light or photon receptors that respond to the presence

---

[206] Benefit/Consumer side teleosemantic theories are advanced by Millikan 1984, Papineau 1984. Producer-side teleosemantic theories are advanced by Neander (2012, 2017), Shea (2018).

and direction of light. When Biwis detect light, this triggers a motor response to swim in a direction opposite the light. This response enables Biwis to avoid swimming into waters with toxic levels of oxygen.

The producer here is the Biwi's light detector system, and the consumer is the motor system that produces swimming behavior. So how should the teleosemanticsist characterize the indicator content the light detector system has the function of producing? Is it the proximal stimulus, namely, light? Or is it the more distal condition, namely, oxygen (level)? On the one hand, Biwis only directly interact with the proximal candidate, i.e., light. Biwis never directly interact with oxygen-- if they do, they die. In other words, the effect that has been selected is a response *to* light, not oxygen. On the other hand, the reason that effect has been selected isn't that light itself is important, but rather the distal condition (oxygen levels) of which it is a sign. Light seems to be *what* gets indicated, but oxygen seems to be *why*.

The producer-side teleosemantics theory will say the indicator content is light, while the consumer-side teleosemantic theory will say that the indicator content is oxygen. What I think this example brings out is that both answers seem sort of right and at the same time sort of wrong. It seems strange to say that the Biwis represent oxygen since it seems strange to say that Biwis represent something they never actually act on. If Biwis represents oxygen, it seems they do so only through representing light. On a selectionist understanding of teleological function, the reason the Biwi's response to light was selected is that historically, those Biwis that responded to light by swimming in the opposite direction didn't make contact with oxygen and therefore survived to reproduce. And more generally, when we are talking about what a creature represents, we seem to be talking about how a creature grasps the things its behaviors are responses to. Since the Biwi's

behaviors respond to light, this is a reason to think that light is what the Biwis represent, and since they don't make any direct contact with oxygen, it's strange to think of them representing it. On the other hand, Biwis don't have any reason to care about light except insofar as it is a sign of oxygen. So, it is strange to think of them *merely* representing light. And as we saw above, in the context of the Frogs example, the consumer/benefit-side teleosemantic way of characterizing content seemed more plausible. In that case, the frog directly responds to edible black dots-- the *what* and *why* of the response don't come apart. But in the Biwi example, they do come apart, and when this happens, the producer-side teleosemantics way of characterizing content seems more intuitive.

In a sense, the PEA analysis is a version of what's sometimes called "consumer-" or "benefit-" side, in contrast with "producer-" or "input-" side, teleosemantic accounts of psychological representation.[207] However, unlike all other teleosemantic accounts, the PEA analysis rejects the descriptive thesis. It holds that original representations are EIRs, not SRs. Rather than using beneficial effects to specify the descriptive contents of indicators, an EIR psychology specifies fundamental representational contents in terms of effects-- specifically, intentional responses-- and what they are directed at. The EIRs are appropriate iff the responses are appropriate, as the responses themselves are supposed to "practically-fit" the facts at which they're directed. EIRs do not have alethic success conditions; their contents are generative (*de actu*) rather than surrogative.

---

[207] On other teleosemantic accounts-- both producer- and consumer-side-- content accords with the descriptive thesis and characterizes some passive causal input. Active causal output just gets involved in the way the input is descriptively characterized. On the proposed view, in contrast, representational content goes primarily with the active causal output which is itself (semi) passively causally shaped by causal inputs and consequently directed at them.

EIR contents capture what both the consumer/benefit-side and the producer-side teleosemantic theories seem to get and without any extra (descriptive) baggage. The Biwi represents its potentiated actions and their targets; it does not represent anything under any descriptions. The content of the Biwi's EIR is "Swim-THAT-Away (from)!" where THAT picks-out the light its detector system detects-- this captures the producer-side teleosemantic insight. But the EIR is only appropriate if that light is in fact a "natural sign" of low oxygen-- this captures the consumer/benefit-side teleosemantic insight.

Where other teleosemantic accounts assume the truth of the descriptive thesis and thus treat original representations as SRs, the proposed account holds that only derived representations are SRs, for precisely the reason that SRs depend on subjects having EIRs. A subject can intentionally use something or produce something to stand-in-for some target at which they are already directed via various EIRs, including the EIR to produce a stand-in for that target. In other words, the sorts of representations that employ descriptions, attributions, etc. are representations that get their referents (ultimately) from the agents who produce or use them. Agents are already intentionally directed at targets when they apply descriptions, attributes, and properties to them in verbal or internalized acts of judgement. In other words, reference precedes description. For the fundamental modes of presentation are generative rather than surrogative.[208] Other teleosemantics accounts think of original representations as SRs that somehow come by their referents "naturally," i.e., without the intentional activities of agents, and this is what gets them into trouble. The PEA account avoids this trouble.

---

[208] One might ask where the descriptions, attributes, etc. agents employ in judgements get their contents. The answer is via the activities (including social activities) of the agents that produce and use them and sometimes from aspects of the surrogative vehicles themselves. I do not think there is a more specific general account of descriptive content available-- different kinds of SRs will need their own accounts.

The Second Difficulty: Explanatory Power & Structural Constraints

In arguing that EIRs can subsume propositional attitudes, I have been careful to maintain a robust, demanding conception of propositional attitudes rather than falling into a deflationary conception. On the view I have proposed, it's not the case that a single, isolated EIR constitutes a propositional attitude. Rather, only EIRs that are generated from IDs that belong to clusters of IDs that include IDs for solutions to certain sorts of practical problems and that take SRs as inputs or produce them as outputs constitute propositional attitudes proper. Consequently, EIR-subsumed propositional attitudes can explain what they are supposed to explain, namely, sophisticated human activities such as complex, planned actions, or acts of judgment, since they correspond to IDs for just such intentional activities. And while we may not be able to attribute propositional attitude constituting ID clusters to non-human animals (and human infants), we can still attribute to them IDs, including IDs of the same generic type (e.g., EAT IDs), and including ID types that belong to those clusters; there's no barrier to attributing EIRs to human infants and non-human animals. Therefore, subsuming propositional attitudes as a species of EIRs makes possible a unified account of psychological representation, and an EIR psychology provides an elegant solution to **The Non-Human Animal Problem.**

EIRs also address the **The Primary Pushmi-Pullyu Problem**. Indeed, an EIR psychology may be understood as modification and development or articulation of consumer-side teleosemantics and pushmi-pullyu representations. EIRs avoid the seeming incoherence of pushmi-pullyu representations while doing justice to their essentially practical nature. Because traditional teleosemantics accounts assume the descriptive thesis, they think of representational *fitness* in terms of *descriptive satisfaction,* alethic success. An EIR psychology (and the PEA

285

analysis of representation more generally) instead thinks of representational fitness in terms of *practical-fitness* (Chapter 3), i.e. appropriateness. Appropriateness conditions (and values) capture the sense in which pushmi-pullyus, qua EIRs, have two directions of fit. For directives and indicatives describe different aspects of EIRs. Directives are a way of describing the actions of which EIRs are potentiations (i.e., ways of describing actions EIRs instruct) with respect to their conditions of successful (complete) execution or performance. Indicatives are ways of describing the appropriateness conditions of the actions of which EIRs are potentiations (i.e., ways of describing actions EIRs instruct) with respect to the facts at which they're directed. Like directives, EIRs are about making things happen. And like directives, EIRs are sensitive to facts (needs and targets) such that they misrepresent (are inappropriate) when they fail to practically fit the facts at which they are directed. But EIRs are not two representations smashed into one pulling in different directions. Quite the contrary, different kinds of mental representations are decompositions of EIRs. However, indicative and directive (or world-to-mind and mind-to-world) directions of fit do not map directly onto these decompositions. All instructive attitudes employ embodied instructive (*de actu*) concepts and hence have constitutive appropriateness conditions: A T-instruction to ϕ-THAT is appropriate iff THAT is an instance of a target-type in ϕ's domain of application and an N-instruction to ϕ is appropriate iff it is prompted by an instance of the need-type that's in ϕ's domain of application (that ϕ functions to solve).

The proposed view also addresses the **Secondary Pushmi-Pullyu Problem**. We can readily concede to Artiga that there is a *sense* in which all representations really are similar to pushmi-pullyus, namely, they all essentially function to provide PEA and consequently they all have constitutive appropriateness conditions (though those grounded in indirect-PEA have alethic conditions understood as a special species of appropriateness conditions). But pushmi-pullyus are

286

EIRs; they provide direct-PEA. SRs, which provide indirect-PEA, are *not* pushmi-pullyus. Pushmi-pullyus instantiate generative aboutness, while non-pushmi-pullyus, i.e., SRs, instantiate surrogative aboutness. The proposed view also meets Artiga's challenge of explaining how detached representations emerge from pushmi-pullyus in that it explains how SRs emerge from EIRs: SRs solve proximity problems and one can exploit an SR to solve a proximity problem only if one has the corresponding EIRs. SRs have alethic success conditions because these measure the properties an SR vehicle must have to potentiate an appropriate indirect action. The view also explains why we seem to be able to distinguish between indicative and imperative psychological representations in terms of the decomposition of IDs and their different modes of activations which constitute the N-, T-, and i- (I-) Instructions and in the ways mentioned in the previous paragraph.

In addition, where it might be unclear how reasoning is supposed to work with pushmi-pullyu representations (or representations with "non-detached" contents more generally), the EIR account comes equipped with descriptions of IDs, sub-IDs, and ODs, and I suggested in Chapter 4 and above that we can make use of traditional models of practical reasoning to describe their activities. If this is right, then EIRs provide an elegant and coherent elaboration of Millikan's pushmi-pullyu representations.

An EIR psychology will also plausibly bypass the interaction problem that arises from the different representational formats of propositional intentions and motor representations. This is because an EIR psychology can subsume propositional attitudes as well as motor representations: Propositional attitudes become clusters of IDs some of whose inputs and outputs are SRs. EIRs and SRs are different kinds of representations and thus have different formats, but it isn't particularly mysterious how they interact: SRs have concrete vehicles that can activate IDs, and EIRs can be for actions that causally produce the vehicles of SRs. Motor representations will

plausibly map onto sub-IDs, which are recruited in determining the manner of an EIR, and thus incorporated into their contents.

Like Gendler's aliefs, EIRs may also account for cases of inconsistent and in-between belief and implicit bias. In fact, since an EIR psychology identifies standing beliefs with IDs and clusters of IDs and occurrent beliefs with the activations of IDS, it *predicts* such phenomena. Consider for instance, Gendler's (2008) example of the inconsistent beliefs many people seem to have about the safety of an all-glass bridge. As Gendler notes, people seem to genuinely believe that such bridges are safe but when they're about to get on one, their behavior suggests otherwise. Now, in many contexts, if a person's IDs are such that she is disposed to judge that the glass bridge is safe when looking at a picture of it, to answer when asked that the bridge is safe, to assure others that the bridge is safe, to disagree with someone who says the bridge is not safe, etc., then that person will count as believing that the bridge is safe-- we will comfortably ascribe that propositional attitude to her. A person can have these IDs and at the same time have certain other IDs such that, when the bridge is right in front of her, she refuses to go on it, or grasps the rail for her life as she reluctantly walks across it or expresses fear for the lives of others who are crossing the bridge. When these are the person's salient dispositions, we are very reluctant to ascribe to her the belief that the bridge is safe. And a person might have all of these IDs activated at once-- she might at once assert confidently that the bridge is safe and refuse, or feel very much inclined to refuse, to cross it.

Now in some of the debates about inconsistent beliefs and related phenomena, one of the objectives is to make sense of such cases in a rationality-preserving way. After all, these inconsistent belief phenomena are pretty common. While some economists appear to be more than happy to embrace a view of humanity as systematically stupid and irrational, most philosophers,

myself included, are much less willing. Although the mixture of IDs described above that dispose a person to refuse to walk on a bridge she swears is safe constitutes a *kind* of practical irrationality, it's not one we should be afraid to admit as a widespread phenomenon. It would be on the model of taking the believing attitude towards "that bridge is safe," and at the same time taking the believing attitude towards its negation, but this is not the EIR model. Nor is the relevant sort of irrationality evidence of some fundamental human stupidity. Quite the contrary, it's a manifestation of our distinctively enlightened nature. For the irrationality consists in the fact that *we recognize* that if we judge that the bridge is safe then there is no good reason to be scared of it. So we recognize that some of our EIRs-- e.g. the one that constitutes our feeling very much inclined to refuse to cross the bridge-- are inappropriate; they're misrepresenting the situation. Recognizing this is precisely what puts us in the position to change our EIRs-- to better embody rationality. But just as recognizing that one is experiencing an illusion will not typically on its own stop the illusion, so recognizing that some of one's IDs are malfunctioning will not typically on its own correct their functioning. Changing our dispositions, especially those that are deeply ingrained, is often extremely challenging work. Indeed, deeply ingrained IDs are often just at the margin between abilities for solutions and abilities for application-accountable solutions. Some of the fear responses one has at the precipice of a glass bridge are not potentiations of intentional actions but rather reflexive responses. What a marvel it is, then, that humans can sometimes take reflexes, i.e., solutions that are not automatically application-accountable, and make them application-accountable by virtue of holding them (and hence ourselves) accountable, i.e. by virtue of making them agent-accountable.

Finally, since EIRs provide a way of making sense of pushmi-pullyu representations, they provide a way of modelling emotions with aptness-- understood as appropriateness-- conditions.

They are particularly well suited to do this because emotions are often thought to be desire-like, but desires, qua propositional attitudes, are insensitive to the facts: a desire succeeds not by means of its content fitting with how the world is but rather by making the world fit with its content. In contrast, N-instructions are sensitive to the facts: an N-instruction to ϕ is appropriate iff it is prompted by an instance of the need-type that's in ϕ's domain of application (that ϕ functions to solve).

6.5 Conclusion

I have sketched an intentions-first EIR psychology spelled out in terms of different ways of activating instructive dispositions (IDs) and I have argued that such a psychology can subsume propositional attitudes. I have also argued that this subsumption is desirable, for it allows us to hold onto a substantive conception of propositional attitudes while avoiding the various difficulties that afflict the standard propositional attitude model. Along the way I hope to have demonstrated the impressive explanatory potential of EIRs.

7.0 Conclusion: Epistemic Dispositions & Rational Actions

This dissertation has tried to do a number of things--quite possibly too many. It has taken on intentionality, at least to some extent, in all its major forms. I have absolutely no doubt that I am guilty of some embarrassing errors and omissions. I hope that, despite this, the encompassing picture of intentional phenomena I've tried to paint has some value.

On this picture, intentionality begins with the needs for flourishing of life forms-- things that have such needs essentially. Needs for flourishing constitute practical problems. Proto-actions are solutions to such problems. Intentional actions are a special form of solution: an application-accountable solution. Intentional states are potentiations of intentional actions, or intentional action acorns. According to my PEA analysis of representation, there are two fundamental species of representation-- embodied instructive representations, which provide direct-PEA and constitute generative modes of aboutness or *de actu* contents, and surrogative representations, which provide indirect-PEA and constitutive surrogative modes of presentation or contents. I believe that embodied instructive representations are a neglected bunch, so I've tried to shine some light on them including the kinds of dispositions that realize them. I've also defended their status as genuine representations, and I've argued that they can do important work in the philosophy of mind, psychology, and action.

By way of concluding, I'd like to address a few worries that I'm only now-- at the end of the dissertation-- in a position to attempt to properly address. I'll begin with concerns regarding my account of representation. Then I'll address some concerns regarding my account of action.

I have defended a kind of teleological behaviorist/ capacity-based account of representation. Concepts, on my view, are abilities (problem-solving dispositions and surrogative extensions of such abilities) for types of direct or indirect intentional actions, and representations are applications of such concepts, i.e., potentiations of tokens of the relevant types of intentional actions (or capacities that consist in such potentiations). In Chapter 4 I noted that there are different views on the relationship between dispositions or capacities and representations, and that some philosophers might object to a view that tries to identify representations with dispositions or capacities on the grounds that dispositions or capacities need to be explained in terms of representations. A version of this objection has been raised explicitly in the context of behaviorist theories of knowledge.

In "How Knowledge Works" (1999), John Hyman argues for the view that "knowledge is the ability to do things, or refrain from doing things, or believe, or want, or doubt things, for reasons that are facts" (441). According to Hyman, one knows that A iff one's reason for doing something can be that A. Timothy Williamson (2000) has raised the following objection to Hyman's behaviorist theory of knowledge:

> Someone in a Gettier case who believes truly that A without knowing that A cannot do X1 for the reason that A, and cannot do X2 for the reason that A, . . . But a single failure to know explains all these incapacities. If the incapacities constituted the failure to know, the correlation between the incapacities would be an unexplained coincidence. (64)

This objection is supposed to show that it does not follow from the fact that one knows that A iff one's reason for doing something can be that A that one can explain knowing that A as being able to do things for the reason that A. In other words, if one does not know (or believe) that Alison has two cats, then one cannot come to Alison's house on the grounds that (or for the reason that) Alison has two cats, nor can one refrain from going to Alison's house on the grounds that Alison

has two cats. And it's not a coincidence that if one does not know (believe) that Alison has two cats, one has both of these incapacities. Indeed, if one does not know (believe) that Alison has two cats, one cannot do anything on the grounds (or for the reason that) Alison has two cats. Williamson's point is that it is the lack of knowledge that explains these incapacities, and hence not the incapacities that explain the lack of knowledge.

Hyman responds to this objection in "Knowledge and Evidence" (2006):

> . . . it simply is not true that if the incapacities constitute the failure to know, the correlation between the incapacities will be an unexplained coincidence: it will be explained by whatever explains the failure to know—the fact that the person has not seen this morning's paper, for example. Similarly, if the ability to do crawl, or breaststroke, or . . . constitutes the ability to swim, it does not follow that the fact that someone cannot do crawl or breaststroke, or [. . .] is an unexplained coincidence. It is explained by whatever explains his inability to swim— the fact that he grew up in an arid country, for example. (906)

Hyman's response, however, is open to the charge that it misses the point of Williamson's objection. (Setiya 2011) Williamson isn't claiming that if the incapacities constitute the absence of knowledge, " the correlation between the incapacities will be an unexplained coincidence." Williamson's claim is rather that what *explains* the correlation could not be the lack of knowledge, or ignorance. The worry is that we seem to use the presence of knowledge to explain myriad capacities that seem to depend on such knowledge, and we seem to use the absence of knowledge to explain a corresponding myriad of incapacities. The presence or absence of knowledge provides a common explanation of why it is that someone lacks a disjunctive capacity like the capacity to swim-- it explains why one lacks the capacity to do crawl, breaststroke, etc. Whereas my inability to swim, it is supposed, can't be what explains my lacking the disjunctive capacities that would constitute that ability. After all, as Setiya (2011) notes, that would be like "explaining why I cannot write the letter 'a' by saying that I cannot write the letter 'a' or the letter 'b' or any letter." (152)

And what goes for knowledge also goes for belief: "if believing that p were having the capacity to do things—something, anything—on the ground that p, the fact that I am incapable of performing some specific action on the ground that p could never be explained by the absence of that belief. And yet it sometimes can" (152).

I believe that the view I have developed over the last five chapters can answer Williamson's objection and Setiya's elaboration of it. There are, however, several layers to this answer.

First, an upshot of my account is that what we call "beliefs" correspond to IDs, clusters of IDs, and EIRs of different kinds. Seemingly simple ascriptions of propositional attitudes and knowledge and belief are efficient ways of glossing what are actual highly complex capacities. Accordingly, when we use the lack of knowledge to explain the presence of lack of capacities, we are referring to the presence of lack of those capacities, or, as Hyman suggests, whatever explains their presence or lack. If one lacks all of the IDs one who counts as knowing that P must possess, then one does not know that P. Having those IDs is constitutive of having the concepts one must have to count as knowing (or believing) that P, where exercising those IDs successfully is what's involved in having the occurrent knowledge (or belief) that P. To be in a position to do something because one knows (or believes) some fact P is to be in a position to perform (have potentiated) one or more actions that is/are directed at P; one can only do something because one knows (or believes) that P if they're in a position to intentionally respond to the fact that P in some relevant way. One must thus have a capacity (ID) for the relevant type of action and the fact that P must be in a position to activate that capacity (ID).[209]

---

[209] When we attribute an occurrent belief that P instead of knowledge that P because P is false, the belief consists in one or more inappropriate potentiated actions (EIRs) directed at an inappropriate target/need where that/those action(s)/ EIR(s) would have been appropriate had P been the target/need at which it/they were directed.

Accordingly, one cannot do anything on the grounds that Alison has two cats so long as the fact that Alison has two cats has no impact (direct or indirect) on one's IDs. This is just what it is to lack knowledge of the fact that Alison has two cats. The proposed account can also account for a richer sense in which one can or cannot do something on the grounds that P in terms of the presence or absence of IDs/EIRs for reasoning about what is the right thing to do where such IDs/EIRs necessarily relate one to intra-subjective reasons, i.e. SRs that stand-in-for the fact that P for the purpose of enabling one to explicitly reason about what is the right thing to do in some context.

So far, then, my response to Williamson's objection is that seemingly simple and general explanations of capacities or incapacities in terms of knowledge (or belief) or the lack thereof are not actually so simple. They are, rather, convenient abstractions. What they actually refer to is significantly more complex-- the clusters of capacities one would have to have to count as having the knowledge or belief that P in some context. But this response does not yet avoid Setiya's development of Williamson's worry which is that if knowledge or belief consist in such capacities, then explaining capacities and incapacities in terms of knowledge or ignorance amounts to explaining capacities and incapacities in terms of more specific or more general capacities and incapacities.

Two assumptions are implicit in Setiya's worry. One is that explanations of capacities and incapacities in terms of knowledge and ignorance respectively are sometimes informative. The other is that explanations of capacities and incapacities in terms of related but more general or specific capacities wouldn't amount to an informative explanation-- no more than "explaining why I cannot write the letter 'a' by saying that I cannot write the letter 'a' or the letter 'b' or any letter" would amount to an informative explanation (152). So according to Setiya's elaboration of

Williamson's objection, knowledge (or belief) and its absence can't just be explained in terms of capacities and incapacities because the former can informatively explain capacities and incapacities while the latter cannot.

To respond to Setiya's elaboration of Williamson's objection, then, what must be shown is that explanations of capacities or incapacities in terms of the presence or absence of more specific or general capacities or incapacities can in fact be informative, and in the same way(s) explanations in terms of knowledge and ignorance can be informative. To do this, we need to consider what functions epistemic explanations of capacities and incapacities typically play.[210]

There are different reasons for asking why it is that someone cannot do something, that is, there are different reasons for seeking an explanation of the absence of some capacity. One reason is to clarify the nature of the incapacity-- whether it's ultimately a problem with the person or with the context.[211] If someone is not in a position to do something, it might be that they lack the relevant capacity, and hence the ID/ cluster of IDs that underwrite that capacity. If one has the relevant capacity, then one is in a position to perform the actions the relevant ID(s) is(are) for insofar as such actions are potentiated, i.e., insofar as the relevant intentional action acorns (EIRs) exist. But one can be in such a position without being in a position to actually perform the relevant potentiated actions if the enabling conditions for the performance of the action do not obtain. So, if one observes that a person is not currently manifesting a capacity one would expect them to

---

[210] I discuss the major functions that come to mind. Perhaps there are others.

[211] Some incapacities are superficial-- you cannot always tell what a person is intentionally doing from what you observe them doing. The error could be a mistaken judgement (inappropriate EIR) or lack of knowledge (lack of an ID or failure for an ID to activate), but it could also be due to bad luck or sloppy performance (inappropriate sub-IDs or failure for sub-IDs to activate). This is a place where distinguishing between knowledge and skill can be helpful. Someone may know how to do something, but be quite unskilled at it, meaning that they have a lower probability of succeeding at the task than one who does it skillfully. How we intervene on a total lack of knowledge-how is different from how we intervene on lower than desirable skillfulness.

manifest in that context, one may ask whether the reason is the lack of the capacity, or the lack of some of the enabling conditions. And the question "why does S have some incapacity" can be answered either with "because of the lack of enabling conditions" or "because of the lack of the relevant IDs, and hence, because of the lack of knowledge/ belief."

For example, in some cases, whether someone has the capacity to swim can be broken into two components: the person-component and the enabling condition component. So, there are two possible explanations of why someone might lack the capacity to swim. It may be that someone lacks the capacity to swim because there are no bodies of water that, in that context, count as being sufficiently nearby. This would be an explanation of an incapacity in terms of the absence of an enabling condition. Alternatively, if someone does not know how to swim, it may be because that person lacks the IDs that underwrite the capacity for swimming. In this latter case, the explanation for why a person cannot swim is that that person lacks the capacities (IDs) that make up the capacity to swim. So, this is one way in which we can informatively explain the lack of a capacity in terms of the lack of the capacities it subsumes.

What about the case in which we want to explain the lack of a more specific capacity, like the capacity to do a breaststroke? It seems we can explain the lack of that capacity, and likewise the absence of other specific swimming capacities, in terms of ignorance, that is, in terms of the fact that one does not know how to swim. How can ignorance provide an informative explanation of the lack of that capacity if ignorance itself just amounts to the lack of a capacity made up of such capacities?

To make the questions more concrete, let's say that Sally can't swim-- she doesn't count as being able to swim in some context. And let's say that the reason she can't swim isn't that there's no body of water that counts as sufficiently nearby in that context, but rather that Sally

lacks the cluster of IDs that count as having the ability/ knowledge of how to swim in that context. Now suppose that Kevin knows that Sally is a big fan of swimming-- he knows that she knows the name of Olympic swimmers, avidly attends swimming competitions, etc. It is reasonable for Kevin to expect that Sally herself swims, since often people who are passionate about a kind of activity practice or at one point in their lives practiced that activity unless something prevents them from doing so. And Kevin doesn't know of anything that would prevent Sally from swimming. But now suppose that Kevin overhears someone saying, "Sally can't do the breaststroke." Hearing this, Kevin might reasonably ask, "why can't Sally do the breaststroke?" and say the person answers: "because she doesn't know how to swim." What could this mean? How could this answer be referring to capacities?

A natural interpretation of "because she doesn't know how to swim" in this case is "because she never learned how to swim." If Sally never learned how to swim, then she lacks a significant number of the IDs one has to have to count as knowing how to swim. In other words, attributing ignorance to Sally amounts to attributing a lack of capacities to Sally together with something like an historical explanation of her lack. Although humans appear to be born with an innate swimming or diving reflex, we are not born knowing how to swim. Humans have to learn how to swim. So, if one lacks the capacities to swim, there's an historical explanation: the sort of learning that consists in the acquisition of such capacities did not occur in the history of the individual in question. So an informative "epistemic" explanation of the lack of some specific capacity in terms of ignorance may amount to an explanation in terms of the lack of the *learning* required to acquire a body of capacities of which that specific capacity is a member. And this means that an informative "epistemic" explanation of a lack of a specific capacity in terms of ignorance may ultimately amount to an explanation in terms of the lack of a more general capacity, or a larger

cluster of associated capacities. This is especially likely in cases in which the learning of the specific capacity typically occurs in the course of acquiring a more general capacity/larger cluster of associated capacities.

Another reason we might be asking for an explanation of an incapacity is because we are aiming to intervene on a subject deemed to lack the relevant capacities, and in order to do this responsibly or scientifically, we need to have a better idea of what underlies the incapacity. For example, let's again assume that Sally can't swim, and that the problem is not that there's no water nearby but rather that Sally lacks some or all of the members of the cluster of IDs that are required in this context to count as knowing how to swim. Now if we ask why Sally can't swim, an informative answer would list the contextually relevant capacities (IDs) that go into the capacity for swimming. This list lets us know what it is that Sally needs to work on if she wants to count (in that context) as having the capacity to swim. Perhaps Sally has some of the IDs in the relevant cluster. Then the intervention required to make it the case that she knows how to swim will only require us to teach her-- instill in her-- the missing IDs. But it's also possible that Sally lacks all of the relevant IDs. In that case, the intervention will require more work. In other words, it is helpful to find out which specific IDs Sally lacks so that we may focus our efforts on helping her acquire those IDs, assuming Sally wants to know how to swim.

So, this is another way in which the lack of a more general capacity can be informatively explained in terms of the lack of specific capacities falling under that more general capacity. Likewise, it can be useful to explain the presence of a more general capacity in terms of the presence of the more specific capacities (IDs) in which having that more general capacity consists. For such an explanation provides a kind of functional decomposition of the general capacity, and such a decomposition tells us what goes into having that capacity, and this is informative about

what one needs to do to acquire that capacity. Such explanations are often sought after in the contexts of self- (and other-) improvement projects.

So, I think my view has the resources to respond to Williamson's and Setiya's objection to dispositionalist or capacity-based accounts of knowledge and corresponding accounts of belief.[212] This response doubles as a response to philosophers who are inclined to think that psychological representations are needed to account for behavioral dispositions and capacities.[213] So I turn now to a more general worry one might have regarding my account of representation which stems from its relationship to my account of intentional action.

My account of intentionality starts with needs. Needs are a type of biological fact. They're intentional types of facts, but they alone are not sufficient for the kinds of intentionality philosophers typically care about. Needs are practical problems and the kind of intentionality we care about starts with solutions to such problems. That means that the philosophically interesting kind of intentionality starts with intentional action. Of course, on my view, intentional states are just potentiations of or capacities for intentional actions, that means agential intentionality and representational intentionality start at the same time. But regardless, my *analysis* puts actions first. So, much rests on this analysis. In particular, my analysis of representation rests on this analysis. So, worries about my analysis of action are also worries about my analysis of representation. So, I had better be able to show that the theoretical foundation I've built is strong. To this end, I'll address a few worries one might have for my account of intentional action (though I'm sure there's plenty I'll miss!).

---

[212] As far as I'm aware, there's no reason why Hyman could not avail himself of these resources. But I have not spent enough time thinking about how his and my view can fit together to confidently assert that he could.

[213] At least, it doubles as a response to those philosophers to the extent that Williamson's and Setiya's objection captures the reason they're inclined to think that psychological representations are needed to account for behavioral dispositions and capacities.

The problem-solving account of intentional action explicitly extends to both human and non-human animal actions. Some philosophers hold that while both human and non-human animals act *purposefully*, only humans act *intentionally*.[214] Accordingly, some philosophers might protest that I have conflated the notion of purposeful action with the notion of intentional action. I tried to anticipate this reaction in my introductory chapter when I mentioned that I have reserved the term "Rational Action" for the distinctively human (rational) species of intentional action and agency that's typically of primary concern to philosophers. I could be wrong, but I do not think philosophers working on rational action should feel the need to deny that rational action is a species of what I call intentional action (or what they might prefer to call purposeful action). So, I take it that even if what I have given an account of so far corresponds to what they would call purposeful action, this needn't be a problem so long as what they would call intentional action may be understood as a species of application-accountable solution. In Chapter 2 I provided some reasons to think that this is possible: rational actions might be analyzed as application-and-agent-accountable solutions to practical problems, where agent-accountability might be connected with special kinds of rational needs for flourishing.

As I mentioned in Chapter 2, although we can give an analysis of intentional action-- of its means end form-- without appealing to intentional states in that analysis, we may need to appeal to representations or explicitly representational capacities to account for rational action. I'd like to now discuss ways in which my account of representations together with the problem-solving

---

[214] For instance, Harry Frankfurt (1978) notes that non-human animals including spiders engage in purposeful behavior. But human action, or much of it, appears to have a richer psychological structure than this. Not only do humans act purposefully, they're also aware (have a kind of knowledge in acting) that, in acting in goal-directed ways, they are aiming at particular ends which they in a sense have selected as their end. They can hence give an answer to Anscombe's "why?' questions. Frankfurt (1988, 1999) has argued that human agents must act on desires with which they "*identify," where this is inextricably bound up with human freedom.* Velleman (2000, essay 6) and Bratman (1999, essay 10) have developed similar accounts of "full-blooded" human agency.

account of intentional action provides useful resources for distinguishing between merely intentional (or purposeful) action and rational action.

First, there's the quality and quantity of the needs for flourishing and thus practical problems different kinds of agents face. More "primitive" agents may have only a few very basic needs for flourishing. More sophisticated agents may have a larger variety of needs, some of which might be rather sophisticated (e.g., intellectual need, needs generated by culture and society).

Second, some creatures might also have a greater capacity to acquire new solutions. Such a capacity involves capacities for learning and for accessing different varieties of facts. A creature's access capacities are her "senses" or "powers of intuition." The capacities we typically think of as our senses give us access to particulars. But philosophy has a long history of attributing a further intellectual sense to humans that can put us in touch with "forms" or "universals" or other non-particulars. In addition, humans seem to have self-conscious access: a human agent's intra-subjective reasons can in some sense be targets for her intentional actions. Consequently, human agents may have a greater ability to shape and determine themselves; they may be capable of a kind of autonomy that's distinctive of their form of agency.

Relatedly, different kinds of agents may have more or less innate as opposed to learned instructive dispositions, or instructive dispositions that are modifiable through learning, and through different kinds of learning--- e.g., the kind that happens automatically vs the kind human agents undergo intentionally through education and training. Since instructive dispositions determine the types of embodied instructive representations one can have, and these determine an agent's fundamental perspective, the perspective of an agent who has mostly innate and rigid instructive dispositions isn't shaped by that agent's personal history and to that extent may be a less personal kind of perspective. So, the extent to which an agent's instructive dispositions are

302

modifiable through learning may determine where on the scale of agency and mindedness, from most primitive to most sophisticated, a kind of creature intuitively falls. Organizing dispositions are responsible for modifying instructive dispositions through learning, so the number and nature of organizing dispositions a kind of creature has may correlate with where they fall on that spectrum as well.

Third: Organizing dispositions are responsible for solving global problems; they connect instructive dispositions to one another by taking inputs from some and outputting excitation or suppression on others, and thereby play a kind of unifying role for an agent's perspective. Even within a single agent, some of her instructive dispositions are likely more and others are likely less connected with organizing dispositions and consequently with each other. The more an instructive disposition is connected with other instructive dispositions, and so the more it's connected with organizing dispositions, the more holistically integrated it is. And plausibly, the more an instructive disposition is holistically integrated, the more it contributes to an agent's unified perspective and consequently the more likely are its exercises to be considered robustly representational and the more likely are the actions it potentiates and produces to be considered robustly intentional.

Table 4 Dimensions of Intentional Action
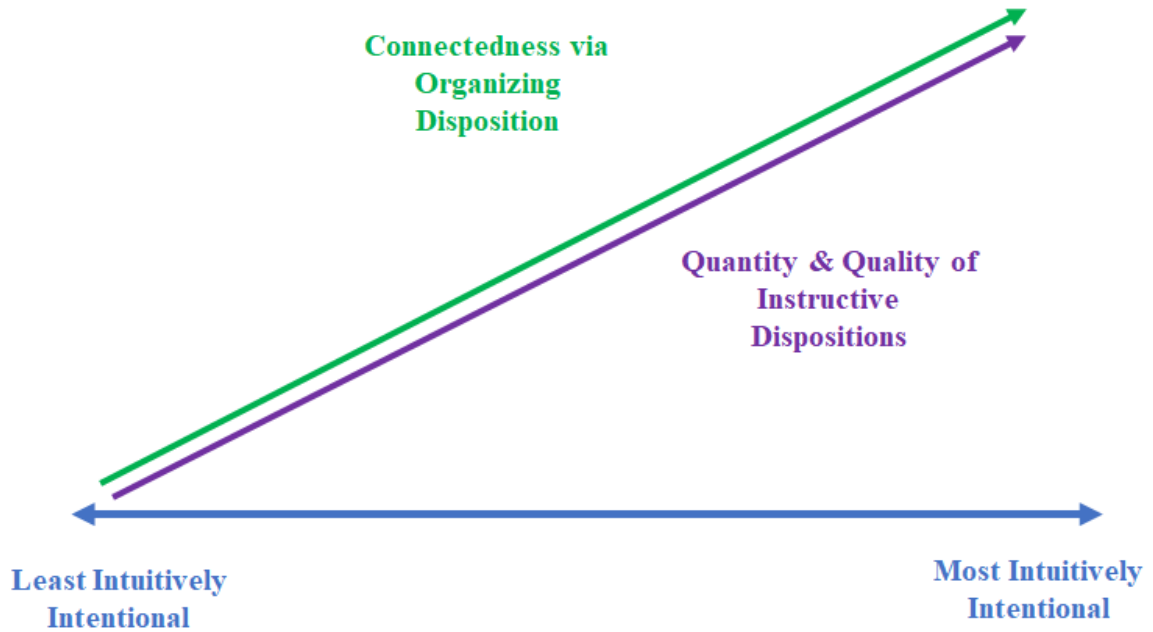
So, the proposed view provides the following kinds of resources for distinguishing the special features of human intentional action, i.e., rational action: 1) special kinds of needs for flourishing, 2) special kinds of solutions, e.g., solutions involving SRs and intra-subjective reasons, and hence linguistic and critical abilities, and 3) special forms of application accountability that may be explained in terms of (1) and (2).

In addition, I take it that intra-subjective reasons (Chapter 3) and propositional attitudes proper (Chapter 5) are essential to the mechanisms that realize agent-accountability. The environments in which we humans live and flourish are environments in which reasons for action (including epistemic actions like believing) are among the facts to be considered; we've made our

home in the space of reasons and/or this home has made us. I'm inclined to agree with Lehrer (1989) that

> what is important about the human mind... is not that it contains sensations, thoughts and desires but that it contains something beyond those states. What is important is that it contains mental states beyond beliefs and desires. First order beliefs and desires are the materials for metamental processing. What is special about human mentality is our capacity for metamental ascent and the conceptually explosive consequences. (19)

By "metamental ascent" Lehrer means the ability to have beliefs and desires whose objects are beliefs and desires; it's a capacity for metacognition. On the proposed account, metamental ascent is a matter of having access to our own EIRs-- our own perspectives-- as objects. In other words, metamental ascent involves intra-subjective reasons (SRs), where propositional attitudes proper involve such reasons. Lehr also suggests an answer to the question of what exactly goes into having abilities for metamental ascent-- what kind of practical problem(s) it solves:

> the function of metamental ascent is the resolution of conflict at the first level. We resolve conflict among our desires for rational decisions in the interests of practice. We resolve conflict among our beliefs for rational acceptance in the interests of truth. The resolution of such conflicts is essential to freedom and knowledge... Metamental ascent is required to articulate the alternative patterns that resolve the conflict. (20)

I am inclined to think that Lehrer's proposal provides at least part of the answer, though I leave it open whether metamental ascent solves other sorts of practical problems, and if so which ones. I think philosophers should be open to the existence of special kinds of intellectual, social, aesthetic, and/or moral needs for flourishing the satisfaction of which requires the construction of increasingly objectified or intersubjective perspectives-- a shared understanding that holds a privileged place in our lives. This specialness is part of what I was trying to bring out by distinguishing intra-subjective reasons from inter-subjective reasons in Chapter 3 and by distinguishing derivative notions of propositional attitudes from propositional attitudes proper in

Chapter 5. I believe that the capacity for this form of surrogative representation is closely connected with self-consciousness and discursive abilities, though at this point I don't dare try to say exactly how they're connected. I also leave open the extent to which the ability to produce SRs that go proxy for EIRs is a cause or effect or simply the two sides of our self-conscious abilities to produce SRs of our own EIRs. All I wish to commit to is that an account of distinctively human intentional action, that is, rational action-- the kind of action that reflects an agent's awareness of the sort of understanding out of which she might act, and thus an ability to be related to her EIRs, to evaluate them as her reasons, and ultimately to do something with them, i.e. to intervene on them-- is compatible with the account of intentional action (what some might just call purposeful action) that I have sketched.

The relationship between rational action and self-consciousness is, I believe, at least a large part of what is at issue when philosophers talk about the relationship between intentional action and self-knowledge. As noted in Chapter 2, Anscombe's knowledge criterion plausibly applies to rational action. However, since intentional action, as I'm understanding it, is supposed to cover actions of both human and non-human (and hence non-linguistic) animals, the knowledge criterion doesn't plausibly apply to intentional action. Merely intentional action does not essentially involve the exercise of *rational* self-conscious capacities. However, I'm inclined to think that intentional actions are always in an important sense self-conscious. In discussing the nature of *de actu* senses/ generative aboutness in Chapter 4, I mentioned that implicit in *de actu* contents is a primitive kind of subject/object distinction in the form of an agent/ target distinction. Agents "show up" in de actu (EIR) contents in that their potentiated actions constitute their fundamental modes of presentation of objects. So, the kind of awareness of the world EIRs afford is at once a kind of spontaneous (generative) self-awareness--a kind of primitive mode of self-consciousness.

This brings me to a final worry I'd like to address: that the problem-solving account of intentional action fails to exclude actions which agents are often unaware of performing and that this is counter intuitive. For example, an agent may be tapping her finger, flicking her tongue, or shaking her leg and not be aware of doing it; if you asked her "why are you doing that?" she'd say "I didn't realize that I was."[215] But on the view I have proposed, if the tapping, flicking, or shaking is an application-accountable solution to a practical problem, it's intentional.

There are different ways to answer this sort of objection. One way of answering this objection is to deny the presupposition that tapping one's finger, flicking one's tongue, etc. are really intentional actions as opposed to merely being things one does. They are not the same as the things one merely does when one, e.g., radiates heat, but the difference might be explained mechanically. They are also not the same as the things one merely does when one trips and falls, but that might be because they involve a pattern of activity. Perhaps those patterns could ultimately be explained in terms of motor capacities associated with our abilities for intentional action without actually constituting exercises of those abilities. It's also possible that these activities are solutions, but not application-accountable solutions.

But let's grant that these activities really are intentional actions. One response to the objection is to say that while these actions are not rational actions they are still intentional actions. They are not rational actions because they fail to meet Anscombe's knowledge criterion (see Chapter 2). But this criterion doesn't apply to intentional actions more generally. On the other hand, above I suggested that intentional actions may always involve some amount of self-

---

[215] See O'Shaughnessy's argument for "sub-intentional acts" (1980 vol ii chapter 10). For a reply with which I'm sympathetic, see Stout (1996 p. 8-11). Thanks to Hong Yu Wong for encouraging me to address this worry.

consciousness. Can this claim be made to square with the kind of case mentioned in the objection? I think it can. For we can understand the agent's tapping, flicking or shaking (and other sorts of "sub-intentional acts") as elements of token application-accountable solutions.

What sort of practical problem might these behaviors function to solve? They strike me as anxious sorts of behaviors, so let's run with that. Anxious behaviors are responses to some kind of practical problem and the agent is aware of being in the midst of trying to solve the relevant problem insofar as she is aware of being anxious. Her tapping, flicking or shaking is a sub-action just as reaching and grasping are kinds of sub-actions. One is aware of doing them insofar as one is aware of doing the overarching action they compose. In other words, awareness goes, primarily, with the activation of instructive dispositions rather than activations of sub-instructive dispositions; awareness of sub-action and sub-instructions (exercises of sub-dispositions) is derivative because they are only derivatively intentional actions. For recall that sub-actions (and sub-sub+ actions), like reaching and grasping, are not themselves solutions to problems but rather solutions to sub-problems (and sub-sub+ problems) that arise in the context of solving overarching practical problems. Agents are often unaware of their sub-actions -- especially when they've been well-trained and largely operate automatically-- except insofar as they are part of the performance of the overarching action (Christensen & Sutton 2016, Friland 2015). We do many things with our bodies in the process of walking to the kitchen to get a drink of which we are not aware, and baseball players are famously unaware of a lot of what they do in the course of e.g., catching a ball (Brownstein & Michaelson 2015). But we don't think that agents are not aware of going to the kitchen or catching balls. Those overarching actions are intentional and self-aware, and they swallow-up the sub-actions that constitute their manner of execution. Likewise, on the view I propose, the agent knows (in the primitive way that's available to animals) that she's tapping,

308

flicking, or kicking insofar as her doing so contributes to the manner in which she manifests the over-arching action that functions to solve the problem she recognizes herself as trying to solve insofar as she recognizes herself as anxious.

<hr />

I have no doubt that there are a number of questions about and objections to the view I've tried to develop in this dissertation that I haven't even begun to address. Believe me, I am, even as I write this conclusion, tempted to continue to rework this manuscript indefinitely, since potential problems, big and small, keep popping up before my mind. At many stages during the writing of this dissertation, I was nearly convinced that I needed to start over -- write a whole new dissertation, and maybe see if I can just do the PhD program over again. Suffice it to say, I wish I could do more -- I wish I could have produced something better thought out, better researched, and better written. But I have done what I can for now, so for now I hope that will do.

Bibliography

Alvarez, M. (2012). The causalism/anti-causalism debate in the theory of action: what it is and why it matters. *Action in Context*, Anton Leist ed., pp. 103-123. De Grruyter.

Alvarez, M. (2009a). Acting Intentionally and Acting for a Reason. *Inquiry: An Interdisciplinary Journal of Philosophy*, 52(3), 293–305.

Alvarez, M. (2009b). How many kinds of reasons? *Philosophical Explorations*, 12(2), 181–193.

Alvarez, M. (2009c). Reasons, desires and intentional actions. In C. Sandis (Ed.), *New Essays on the Explanation of Action*. Palgrave-Macmillan.

Alvarez, M. (2010a). *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford University Press.

Alvarez, M. (2010b). Reasons for action and practical reasoning. *Ratio*, 23(4), 355–373.

Anscombe, G. E. M. (1957). *Intention* (Vol. 57, pp. 321–332). Harvard University Press.

Anscombe, G. E. M. (1963). Critical notices. *Mind*, 72(286), 288–293.

Artiga, M. (2014). Teleosemantics and Pushmi-Pullyu Representations. *Erkenntnis*, 79(S3), 1–22.

Audi, R. (2001). *The Architecture of Reason: The Structure and Substance of Rationality* (Vol. 44, pp. 1–41). Oxford University Press.

Bach, K. (1980). Actions are not events. *Mind*, 89(353), 114–120.

Baier, K. (1958). *The Moral Point of View: A Rational Basis of Ethics* (Vol. 35, pp. 69–70). Cornell University Press.

Baker, L. R. (1995). *Explaining Attitudes: A Practical Approach to the Mind*. Cambridge University Press.

Bendana, J., & Mandelbaum, E. (forthcoming). The Fragmentation of Belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The Fragmented Mind*.

Bennett, M. R., & Hacker, P. M. S. (2008). History of Cognitive Neuroscience.

Bermudez, J. (2009). The Distinction between Conceptual and Nonconceptual Content. In B. McLaughlin, A.

Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford University Press.

Bermúdez, J. L. (1998). *The Paradox of Self-Consciousness: Representation and Mind*. MIT Press.

Bermudez, J. L., & Macpherson, F. (1998). Nonconceptual content and the nature of perceptual experience. *Electronic Journal of Analytic Philosophy*, 6.

Bishop, J. (1989). *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge University Press.

Brand, D. J. (1984). *The Book of Causes* . Marquette University Press.

Brandom, R. B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment* (Vol. 183, p. 238). Harvard University Press.

Bratman, M. (1987). *Intention, Plans, and Practical Reason* (pp. 198–199). Cambridge, MA: Harvard University Press.

Bratman, M. E. (1999). *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press.

Brownstein, M., & Michaelson, E. (2016). Doing without believing: Intellectualism, knowledge-how, and belief-attribution. *Synthese*, 193(9), 2815–2836.

Brozzo, C. (2017). Motor Intentions: How Intentions and Motor Representations Come Together. *Mind and Language*, 32(2), 231–256.

Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.

Burnston, D. C. (2017). Interface problems in the explanation of action. *Philosophical Explorations*, 20(2), 242–258.

Butterfill, S. A., & Sinigaglia, C. (2014). Intention and Motor Representation in Purposive Action. *Philosophy and Phenomenological Research*, 88(1), 119–145.

Byrne, A. (2021). Concepts, Belief, and Perception. In C. Demmerling & D. Schröder (eds.), *Concepts in Thought, Action, and Emotion: New Essays*.

Chomsky, N. A. (1980). Rules and representations. *Behavioral and Brain Sciences*, 3(127), 1–61.

Christensen, W., Sutton, J., & Bicknell, K. (2019). Memory systems and the control of skilled action. *Philosophical Psychology*, 32 (5), 692–718.

Christensen, W., Sutton, J., & McIlwain, D. J. F. (2016). Cognition in Skilled Action: Meshed Control and the Varieties of Skill Experience. *Mind and Language*, 31(1), 37–66.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again* (pp. 138–143). MIT Press.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive* 2 *Extension*. Oxford University Press.

Cohen, J. (2015). Perceptual Constancy. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp. 621–639).

Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, (72), 741–64.

Cummins, R. (1996). *Representations, Targets, and Attitudes* (pp. 175–180). MIT Press.

Cuneo, T. (2007). Divine motivation theory. *Philosophical Books*, 48(3), 252–261.

Cussins, A. (1992). Content, embodiment and objectivity: The theory of cognitive trails. *Mind*, 101(404), 651–88.

Dancy, J. (1995). Why There Is Really No Such Thing as the Theory of Motivation. *Proceedings of the Aristotelian Society*, (95), 1–18.

Dancy, J. (2000). *Practical Reality* (Vol. 78, pp. 414–425). Oxford University Press.

Dancy, J. (2004). Enticing Reasons. In R. J. Wallace, P. Pettit, S. Scheffler, & M. Smith (Eds.), *Reason and Value: Themes From the Moral Philosophy of Joseph Raz* (pp. 91–118). Clarendon Press.

Dancy, J. (2018). *Practical Shape: A Theory of Practical Reasoning*. Oxford University Press.

D'Arms, J., & Jacobson, D. (2014). *Moral Psychology and Human Agency: Philosophical Essays on the Science of Ethics*. Oxford University Press.

Darwall, S. (2003). Desires, reasons, and causes. *Philosophy and Phenomenological Research*, 67(2), 436–443.

Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60(23), 685.

Davidson, D. (1971). Agent, Action, and Reason. In R. Binkley, R. Bronaugh, & A. Marras (Eds.), *Agent, Action, and Reason*. University of Toronto Press.

Davidson, D. (1973). On the Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association*, (47), 5–20.

Davidson, D. (1980). Toward a unified theory of meaning and action. *Grazer Philosophische Studien*, 11(1), 1–12.

Demmerling, C., & Schroder, D. (2020). *Concepts in Thought, Action, and Emotion: New Essays*. Routledge.

Dennett, D. (1994). Get Real. *Philosophical Topics*, 22(1/2), 505–568.

Dennett, D. C. (1978a). *Brainstorms* (p. 367). MIT Press.

Dennett, D. C. (1978b). *Brainstorms: Philosophical Essays on Mind and Psychology* (pp. 115–117). Bradford Books.

Dennett, D. C. (1987). *The Intentional Stance* (pp. 891–892). MIT Press.

D'Oro, G., & Sandis, C. (2013). *Reasons and Causes: Causalism and Non-causalism in the Philosophy of Action*. Palgrave-Macmillan.

Dretske, F. (1969). *Seeing And Knowing* (Vol. 21, p. 82). University Of Chicago Press.

Dretske, F. (1986). Aspects of cognitive representation. In M. Brand & R. M. Harnish (Eds.), *The Representation of Knowledge and Belief*. University of Arizona Press.

Dretske, F. (1995). *Naturalizing the Mind* (pp. 528–537). MIT Press.

Dretske, F. I. (1981). *Knowledge and the Flow of Information* (Vol. 92, pp. 452–454). MIT Press.

Dretske, F. I. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief: Form, Content, and Function* (pp. 17--36). Oxford University Press.

Dretske, F. I. (1988). *Explaining Behavior: Reasons in a World of Causes* (Vol. 100, pp. 641–645). MIT Press.

Dummett, M. (1993). *The Seas of Language* (Vol. 58, pp. 781–782). Oxford University Press.

Egan, A. (2008). Seeing and believing: perception, belief formation and the divided mind. *Philosophical Studies*, 140(1), 47–63.

Evans, G. (1982). *The Varieties of Reference* (Vol. 137, p. 515). Oxford University Press.

Evans, G., & Mcdowell, J. (1978). Truth and Meaning: Essays in Semantics. *Philosophy*, 53(205), 404–407.

Ferretti, G., & Caiani, S. Z. (2019). Solving the Interface Problem Without Translation: The Same Format Thesis. *Pacific Philosophical Quarterly*, 100(1), 301–333.

Fischer, J. M., & Wilson, G. M. (1984). The Intentionality of Human Action. *Philosophical Review*, 93(3), 483.

Fodor, J. A. (1968). *Psychological Explanation: An Introduction To The Philosophy Of Psychology* (Vol. 80, pp. 108–113). Random House.

Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (pp. 289–293). MIT Press.

Frankfurt, H. (1999). Responses. *The Journal of Ethics*, 3(4), 369–374.

Frankfurt, H. G. (1978). The Problem of Action. *American Philosophical Quarterly*, 15(2), 157–162.

Frankfurt, H. G. (1988). *The Importance of What We Care About: Philosophical Essays*. Cambridge University Press.

Fridland, E. (2015). Knowing-how: Problems and Considerations. *European Journal of Philosophy*, 23(3), 703–727.

Fridland, E. (2017). Skill and motor control: intelligence all the way down. *Philosophical Studies*, 174(6), 1–22.

Gendler, T. S. (2008a). Alief and Belief. *Journal of Philosophy*, 105(10), 634–663.

Gendler, T. S. (2008b). Alief in Action (and Reaction). *Mind and Language*, 23(5), 552--585.

Gibson, J.J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gibson, J.J. (1977). "The Theory of Affordances*"* In R. Shaw & J. Bransford (eds.). *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ: Lawrence Erlbaum. pp. 67–82.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Ginet, C. (1990). *On Action*. Cambridge University Press.

Glüer, K. (2009). In Defence of a Doxastic Account of Experience. *Mind and Language*, 24(3), 297–327.

Goldman, A. I. (1970). *A Theory of Human Action* (p. 180). Princeton University Press.

Green, E. J., & Schellenberg, S. (2018). Spatial perception: The perspectival aspect of perception. *Philosophy Compass*, 13(2), e12472.

Greenberg, G. (2013). Beyond Resemblance. *Philosophical Review*, 122(2), 215–287.

Grice, H. P. (1957). Meaning. *Philosophical Review*, 66(3), 377–388.

Grush, R. (2007). Skill Theory v2.0: Dispositions, Emulation, and Spatial Perception. *Synthese*, 159(3), 389–416.

Grzankowski, A., & Montague, M. (2018). *Non-Propositional Intentionality*. Oxford University Press.

Haugeland, J. (1981). *Mind Design: Philosophy, Psychology, and Artificial Intelligence*. MIT Press.

Hieronymi, P. (2011). Reasons for Action. *Proceedings of the Aristotelian Society*, 111(3pt3), 407–427.

Hurley, S. L. (1998). *Consciousness in Action*. Harvard University Press.

Hursthouse, R. (1991). Arational actions. *Journal of Philosophy*, 88(2), 57–68.

Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism: Basic Minds Meet Content*. MIT Press.

Hyman, J. (1999). How knowledge works. *Philosophical Quarterly*, 49(197), 433–451.

Hyman, J. (2006). Knowledge and evidence. *Mind*, 115(460), 891–916.

Kalderon, M. E. (2017). *Sympathy in Perception*. Cambridge University Press.

Kelly, S. D. (2001). Demonstrative concepts and experience. *Philosophical Review*, 110(3), 397–420.

Kenny, A. (2010). Concepts, Brains, and Behaviour. *Grazer Philosophische Studien*, 81(1), 105–113.

Kim, J. (1989). Mechanism, purpose, and explanatory exclusion. *Philosophical Perspectives*, (3), 77–108.

Lakoff, G. (1999). *Philosophy in the Flesh the Embodied Mind and its Challenge to Western Thought*. Basic Books.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By* (Vol. 40, pp. 208–210). University of Chicago Press.

Lehrer, K. (1989). Metamental Ascent: Beyond Belief and Desire. *Proceedings and Addresses of the American Philosophical Association*, 63(3), 19–30.

Levy, N. (2017). Embodied savoir-faire: knowledge-how requires motor representations. *Synthese*, 194(2).

Lewis, D. K. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50(3), 249–258.

Mandelbaum, E. (2013). Against alief. *Philosophical Studies*, 165(1), 197–211. Mandelbaum, E. (2016). Attitude, Inference, Association: On the Propositional Structure of Implicit Bias. *Noûs*, 50(3), 629–658.

Mandik, P. (2005). Action-oriented representation. In A. Brook & K. Akins (Eds.), *Cognition and the Brain: The Philosophy and Neuroscience Movement* (pp. 284--305). Cambridge University Press.

Mantel, S. (2014). No reason for identity: on the relation between motivating and normative reasons. *Philosophical Explorations*, 17(1), 49–62.

Matthen, M. (1988). Biological functions and perceptual content. *Journal of Philosophy*, (85), 5–27.

Matthen, M. (2005). *Seeing, Doing, and Knowing: A Philosophical Theory of Sense Perception.* Oxford University Press UK.

Mcdowell, J. (1984). De re senses. *Philosophical Quarterly*, 34(136), 283–294.

McDowell, J. (1994). *Mind and World* (pp. 169–181). Harvard University Press.

Mcdowell, J. (2011). "Some Remarks On Intention In Action." *Studies in Social Justice*, 1–18.

Mele, A. R. (1992). *Springs of Action: Understanding Intentional Behavior* ,103(175–178). Oxford University Press.

Mele, A. R. (1997). *The Philosophy of Action*. Oxford University Press.

Mele, A. R. (2003). *Motivation and Agency* (pp. 197–200). Oxford University Press.

Mele, A. R., & Moser, P. K. (1994). Intentional action. *Noûs*, 28(1), 39–68.

Millikan, R. (2006). Useless content. In G. Macdonald & D. Papineau (Eds.), *Teleosemantics.* Oxford University Press.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism,* 14 (51–56). MIT Press.

Millikan, R. G. (1990). Compare and contrast Dretske, Fodor, and Millikan on teleosemantics. *Philosophical Topics*, 18(2), 151–61.

Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives*, 9, (185–200).

Millikan, R. G. (2004). *Varieties of Meaning: The 2002 Jean Nicod Lectures* (pp. 674–681). MIT Press.

Millikan, R. G. (2006). Styles of Rationality. In S. Hurley & M. Nudds (Eds.), *Rational Animals?* Oxford University Press.

Millikan, R. G. (2017). *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford University Press.

Millikan, R. G. (2020). Comment on Artiga's "Teleosemantics and Pushmi-Pullyu Representations." *Erkenntnis*, 1–9.

Millikan, R. G. (2020). Neuroscience and Teleosemantics. *Synthese*, 1–9.

Mylopoulos, M., & Pacherie, E. (2017). Intentions and Motor Representations: the Interface Challenge. *Review of Philosophy and Psychology*, 8(2), 317–336.

Nagel, T. (1970). *The Possibility of Altruism* (p. 82). Oxford Clarendon Press.

Nagel, T. (1986). *The View From Nowhere* (Vol. 37, p. 323). Oxford University Press.

Nanay, B. (2013). *Between Perception and Action*. Oxford University Press.

Neander, K. (1995). Misrepresenting and Malfunctioning. *Philosophical Studies*, 79(2), 109–41.

Neander, K. (2004). Teleological Theories of Mental Content. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*.

Neander, K. (2013). Toward an Informational Teleosemantics. In D. Ryder, J. Kingsbury, & K. Williford (Eds.), *Millikan and Her Critics* (pp. 21--40). Wiley.

Neander, K. (2017). *A Mark of the Mental: A Defence of Informational Teleosemantics*. MIT Press.

Neander, K., & Rosenberg, A. (2012). Solving the Circularity Problem for Functions: A Response to Nanay. *Journal of Philosophy*, 109(10), 613–622.

Norby, A. (2014). Against Fragmentation. *Thought: A Journal of Philosophy*, 3(1), 30–38.

Nussbaum, M. C. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.

Nussbaum, M. C. (2016). *Anger and Forgiveness: Resentment, Generosity, Justice*. Oxford University Press.

Orlandi, N. (2014). *The Innocent Eye: Why Vision is Not a Cognitive Process*. Oxford University Press.

O'Shaughnessy, B. (1980). *The Will: A Dual Aspect Theory* (Vol. 95, pp. 451–455). Cambridge University Press.

Papineau, D. (1984). Representation and explanation. *Philosophy of Science*, 8(51), 550–72.

Papineau, D. (1987). *Reality and Representation* (Vol. 100, pp. 109–111). Blackwell.

Papineau, D. (1990a). Explanation in Psychology: Truth and Teleology. *Royal Institute of Philosophy Supplement*, 27, 21–43.

Papineau, D. (1990b). Reality and Representation. *Behavior and Philosophy*, 18(2), 85–88.

Papineau, D. (1990c). Truth and Teleology. In D. Knowles (Ed.), *Explanation and its Limits* (pp. 21–43). Cambridge University Press.

Papineau, D. (1993). *Philosophical Naturalism. Philosophical Naturalism* (pp. 1070–1077). Blackwell.

Papineau, D. (1998). Teleosemantics and indeterminacy. *Australasian Journal of Philosophy*, 76(1), 1–14.

Parfit, D. (1997). Reasons and Motivation. *Aristotelian Society Supplementary Volume*, 71(1), 99–130.

Pavese, C. (2015). Practical Senses. *Philosophers' Imprint*, 15.

Pavese, C. (2019). The Psychological Reality of Practical Representation. *Philosophical Psychology*, 32(5), 784–821.

Pavese, C. (2021). Practical Concepts and Productive Reasoning. *Synthese*, 1–30.

Pavese, C. (forthcoming). Practical Representation. In C. Pavese & E. Fridland (Eds.), *Routledge Handbook of Philosophy of Skill and Expertise*. Routledge.

Peacocke, C. (1983). *Sense and Content: Experience, Thought and Their Relations* (pp. 581–585). Oxford University Press.

Peacocke, C. (1992). *A Study of Concepts* (Vol. 103, pp. 541–544). MIT Press.

Peacocke, C. (1998). Nonconceptual Content Defended. *Philosophy and Phenomenological Research*, 58(2), 381–388.

Peacocke, C. (2001a). Does Perception Have a Nonconceptual Content? *Journal of Philosophy*, *98*(5), 239–264.

Peacocke, C. (2001b). Phenomenology and Nonconceptual Content. *Philosophy and Phenomenological Research*, *62*(3), 609–615.

Pierce, C. (1883). Studies in Logic. *Mind*, 8(32), 594–603.

Quilty-Dunn, J. (2016). Iconicity and the Format of Perception. *Journal of Consciousness Studies*, 23(3–4), 255–263.

Quilty-Dunn, J., & Mandelbaum, E. (2018). Against Dispositionalism: Belief in Cognitive Science. *Philosophical Studies*, 175(9), 2353–2372.

Ramsey, W. M. (2007). *Representation Reconsidered*. Cambridge University Press. Raz, J. (1975). *Practical Reason and Norms* (Vol. 26, p. 287). Hutchinson.

Raz, J. (1998). Incommensurability and agency. In J. Raz (Ed.), *Engaging Reason: On the Theory of Value and Action* (pp. 110–28). Oxford University Press.

Rosenberg, G. H., & Anderson, M. L. (2008). Content and action: The guidance theory of representation. *Journal of Mind and Behavior*, 29(1–2), 55–86.

Rowlands, M. (2006). *Body Language: Representation in Action*. Bradford Book/MIT Press.

Ryle, G. (1949). *The Concept of Mind* (pp. 23–37). Hutchinson & Co.

Scanlon, T. (1998). *What We Owe to Each Other* (Vol. 66, pp. 169–175). Belknap Press of Harvard University Press.

Scanlon, T. M. (2014). *Being Realistic About Reasons*. Oxford University Press.

Scarantino, A., & Nielsen, M. (2015). Voodoo dolls and angry lions: how emotions explain arational actions. *Philosophical Studies*, 172(11), 2975–2998.

Schellenberg, S. (2010). Perceptual Experience and the Capacity to Act. In N. Gangopadhay, M. Madary, & F. Spicer (Eds.), *Perception, Action, and Consciousness* (p. 145). Oxford University Press.

Schellenberg, S. (2014). The Relational and Representational Character of Perceptual Experience. In B. Brogaard (Ed.), *Does Perception Have Content* (pp. 199–219). Oxford University Press.

Schellenberg, S. (2016). De Se Content and De Hinc Content. *Analysis*, 76(3), 334–345.

Schellenberg, S. (2018). *The Unity of Perception: Content, Consciousness, Evidence*. Oxford University Press.

Schellenberg, S. (2019). Accuracy Conditions, Functions, Perceptual Discrimination. *Analysis*, 79(4), 739–754.

Schroeder, M. (2007). *Slaves of the Passions* (pp. 574–576). Oxford University Press.

Schroeder, M. (2008). Having Reasons. *Philosophical Studies*, 139(1), 57–71. 10

Schwitzgebel, E. (2002). A Phenomenal, Dispositional Account of Belief. *Noûs*, 36(2), 249–275.

Schwitzgebel, E. (2010). Acting Contrary to our Professed Beliefs or The Gulf Between Occurrent Judgment and Dispositional Belief. *Pacific Philosophical Quarterly*, 91(4), 531–553.

Searle, J. (1980). Intrinsic intentionality. *Behavioral and Brain Sciences*, 3(3), 450–457.

Searle, J. (1992). *The Rediscovery of the Mind* (Vol. 44, p. 259). MIT Press.

Searle, John R. (1980). The Intentionality of Intention and Action. *Cognitive Science*, 4(1), 47–70.

Searle, John R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.

Sellars, W. S. (1962). Philosophy and the scientific image of man. In R. Colodny (Ed.), *Science, Perception, and Reality* (pp. 35–78). Humanities Press/Ridgeview.

Setiya, K. (2008). Practical knowledge. *Ethics*, 118(3), 388–409.

Setiya, K. (2011a). Knowledge of intention. In A. Ford, J. Hornsby, & F. Stoutland (Eds.), *Essays on Anscombe's Intention* (pp. 170--197). Harvard University Press.

Setiya, K. (2011b). Reasons and Causes. *European Journal of Philosophy*, 19(1), 129–157.

Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press.

Siegel, Susanna. (2005). Which properties are represented in perception. In Tamar S. Gendler & J. Hawthorne (Eds.), *Perceptual Experience* (pp. 481--503). Oxford University Press.

Siegel, Susanna. (2014). Affordances and the Contents of Perception. In Berit Brogaard (Ed.), *Does Perception Have Content?* (pp. 39–76). Oxford University Press.

Siegel, Susannah. (2010). *The Contents of Visual Experience*. Oxford University Press USA.

Skorupski, J. (2010). *The Domain of Reasons*. Oxford University Press.

Smith, M. (1987). The Humean theory of motivation. *Mind*, *96*(381), 36–61. Smith, M. (1994). *The Moral Problem* (pp. 125–126). Blackwell.

Springle, A. (2019a). Methods, Minds, Memory, and Kinds. *Philosophical Psychology*, 32(5), 635–661.

Springle, A. (2019b). Perception Representation Realism and Function. *Philosophy of Science*, 86(5), 1202–1213.

Springle, A. A., & Humphreys, J. (2021a). Anti-Intellectualism, Instructive Representations, and the Intentional Action Argument. *Synthese*, Minds in Skilled Performance.

Srinivasan, A. (2018). The Aptness of Anger. *Journal of Political Philosophy*, 26(2), 123–144.

Stalnaker, R. (2012). Intellectualism and the Objects of Knowledge. *Philosophy and Phenomenological Research*, 85(3), 754–761.

Stanley, J. (2011). *Know How*. Oxford University Press.

Starr, W. B. (2020). A Preference Semantics for Imperatives. *Semantics and Pragmatics*, 20.

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Wiley-Blackwell.

Stout, R. (1996). *Things That Happen Because They Should: A Teleological Approach to Action*. Oxford University Press.

Thalberg, I. (1984). Do Our Intentions Cause Our Intentional Actions? *American Philosophical Quarterly*, 21(3), 249–260.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard

University Press.

Thompson, M. (2008). *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.

Travis, C. (2013). *Perception: Essays After Frege*. Oxford University Press.

Tye, M. (2005). On The Nonconceptual Content of Experience. *Schriftenreihe-Wittgenstein Gesellschaft*.

Velleman, D. (2000). *The Possibility of Practical Reason* (Vol. 106, pp. 694-). Oxford University Press.

Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Watzl, S. (2014). Perceptual Guidance. *Ratio*, *27*(4), 414–438. 12

Williams, B. (1979). Internal and External Reasons. In R. Harrison (Ed.), *Rational Action* (pp. 101–113). Cambridge University Press.

Williamson, T. (2000). *Knowledge and its Limits* (pp. 460–464). Oxford University Press.

Wilson, G. M. (1980). *The Intentionality of Human Action* (pp. 483–489). Stanford University Press.

Wright, L. (1973). Functions. *Philosophical Review*, 82(2), 139–168.

Wright, L. (1976). *Teleological Explanations: An Etiological Analysis of Goals and Functions*. University of California Press.

Wu, W. (2016). Experts and Deviants: The Story of Agentive Control. *Philosophy and Phenomenological Research*, 92(2), 101–26.