

and microstructural constitution. My wedding ring is composed of gold and platinum, with a certain amount of impurity. Its having this composition does not depend on anything outside the ring itself, and hence it is an intrinsic property of the ring. Similarly, my favorite coffee mug has the property of being composed of an arrangement of particular ceramic crystals. This property is intrinsic to the cup and involves nothing external to it.

Relational properties often have intrinsic counterparts. Corresponding to the relativistic property of mass, there is rest mass. Rest mass is the mass of a body when at rest, as it would be measured by an observer who is at rest in the same frame of reference. The rest mass of a body is its intrinsic contribution to its relativistic mass, relative to this or that frame of reference. Analogously for being hard or brittle or flexible: these properties might be understood relationally, in terms of being resistant to certain impacts, breakable by certain impacts, or prone to alter shape when subject to certain forces, respectively. But in each case there is an intrinsic counterpart, for example, an object's proneness to resist certain impacts depends on how its components (molecules or whatever) are bound together, as does an object's proneness to break or to alter shape.

It should be obvious that a good understanding of the nature of a property requires knowing whether it is relational or intrinsic. One would be in the dark as to the nature of avuncularity, for example, if one did not know the kinship relationships involved. Similarly, humans were somewhat in the dark as to the real nature of mass, shape, and so on, prior to Einstein's discovery of Special Relativity. Equally, one would be somewhat in the dark about the property of being made of gold if one did not know that it was intrinsic, a matter of molecular constitution, rather than, say, etiology or market value.

The concern of this book is whether certain psychological properties are relational or intrinsic. Specifically, it will be concerned with what might be called the “cognitive properties” or “cognitive content” of psychological states. By “cognitive properties” I mean those properties that account for the role of these states in typical psychological predictions and explanations. Suppose, for example, that Yogi believes that orangutans are omnivorous, and that all omnivores like chocolate. We might then predict that, if he considers the matter, he will come to believe that orangutans like chocolate. Or, after the fact, we might explain his believing that orangutans like chocolate by citing the other two beliefs. Similarly, if Yogi himself wants to buy some chocolate and believes that in order to buy some chocolate, it is necessary to go to the shops, then we might predict that he will go to the shops. Again after the fact, we might explain why he went to the shops in terms of his desire and his belief.

Psychological explanations of this sort evidently draw on a specific range of properties of the states they cite. The properties appear to be specified by the embedded complement clauses of propositional-attitude attributions, the “that” clauses of “believes that p ,” “doubts that q ,” “hopes that r ,” etc. These clauses give the contents of states they ascribe, in the sense that they specify what is believed, doubted, hoped, and so on.

The term “content” used in this loose, intuitive sense, is rather vague and ambiguous. Suppose, for example, that each day Abraracourcix believes that the sky will fall on his head tomorrow. Does he have the same belief, that is, a belief with the same content, from day to day? In one sense, it appears not, since Monday’s belief has different

truth conditions from Tuesday's. He does not, then, retain a belief with the same truth conditions. But it is tempting also to suppose that in another sense of content, his belief each day has the same content: that the sky will fall on his head tomorrow. A related but different ambiguity in "content" occurs with examples like believing that water boils at 100 degrees Centigrade and believing that H₂O boils at 100 degrees Centigrade. These appear to have the same referential content: the beliefs predicate the same concept (boiling at 100 degrees C) of the same substance (water/H₂O). But they appear to differ in cognitive content, since the beliefs would play different roles in a person's thinking.

From now on when I use "content" unmodified, I mean "cognitive content" rather than, say, truth-conditional or referential content. I will not make any initial assumptions about the relationship among different notions of content, about whether two or more notions might pick out the same phenomenon, and so on. Nor will I make any initial assumptions about the precise relation between cognitive content and the complement clauses of attitude attributions. I leave it open, for now, whether identity and distinctness of complement clauses correspond directly to sameness and difference of content. These matters will come up for explicit discussion as we proceed.

I will merely assume that there is such a thing as cognitive content, that it drives standard psychological explanations, and that we use attitude attributions to get at it in some manner or other. I will use "content" to refer to properties and items as they would be individuated in a true psychological theory. So questions about sameness or distinctness of contents are questions about the taxonomy of a correct psychological theory. I will also assume, except where I explicitly say otherwise, that psychological items, states and events, at least cognitive and representational

ones, are to be individuated by their contents. So questions about sameness and difference of beliefs, concepts, etc., are questions about sameness and difference of contents.

Although I will focus almost exclusively on ordinary commonsense psychology and the propositional attitudes that are its main concern, I intend all the main arguments that I offer to extend to any branch of scientific psychology that recognizes contentful states and to all such states, including perceptual states, states of the Freudian unconscious, tacit cognition of language, neonate cognition, animal cognition, and so on.

1.3 TWIN WORLDS

Are cognitive contents relational or intrinsic? Suppose, for example, that Zowie believes that her engagement ring is studded with diamonds. Does her having a belief with that content essentially involve any relations to anything beyond Zowie herself? The easiest way to get a handle on this question is to consider in what kinds of environments it would be possible to have a belief with just that content. Could Zowie have such a belief in a world in which her engagement ring did not exist? (Poor deluded Zowie, driven insane by love unrequited.) Could she have such a belief in a world in which diamonds did not exist? Or in a world with no other humans? (Zowie, an artificially constructed brain in an extraterrestrial scientist's laboratory.)

A particular kind of thought experiment introduced by Hilary Putnam (1975a) is very useful for rendering such questions vivid. Thought experiments of this kind involve imagining or conceiving of what we can call "twin" subjects in "twin" worlds. Twin subjects are microstructural duplicates of each other: they are structurally identical in respect of the elementary particles, the atoms

and molecules, the nerve cells and their interconnections, the neural structures, etc., that make them up. Twin worlds are also microstructural duplicates, except in one specially selected respect. The thought experiment involves assessing whether the difference between the worlds entails a difference in the psychological properties of the subjects.

We have our Earthly subject, Zowie, believing that her ring is studded with diamonds. Let us suppose that Zowie lived in the seventeenth century, prior to the discovery that diamonds are made of pure carbon. We now imagine a twin Zowie on a twin Earth. Twin Zowie is a microstructural duplicate of Zowie. At any given moment, her brain, central nervous system, and everything else within her body are in exactly the same configurations as Zowie's. Twin Earth is exactly like Earth, except in respect of diamonds. On Twin Earth there are stones that are just like diamonds in all superficial respects: they are very hard, when of good quality they sparkle enticingly in the light, etc. They are so like diamonds that nobody on Earth or Twin Earth could have distinguished the two in the seventeenth century. These stones are regarded on Twin Earth just as diamonds are on Earth. Twin Earth counterparts of English speakers call them "diamonds," but we can call them "twin diamonds." Twin diamonds differ from Earth diamonds in their internal constitution, being made up not of carbon but of a kind of aluminum oxide.

saying is that her engagement ring is studded with diamonds. If the stones are indeed diamonds, what she says is true. And if what she says is what she believes, then what she believes is true too. Had the stone on her ring been a twin diamond, aluminum oxide rather than carbon, she would have said and believed something false. For twin diamonds are not diamonds: they just look like them.

Of course, Zowie does not know that diamonds are made of carbon. So she doesn't know that if her jewel is not made of carbon, then it is not a diamond. But that makes no difference. Many contemporary English speakers don't know that diamonds are made of carbon. Yet when they say "diamond" they still mean *diamond*. If one of them pointed to a twin diamond and said, "That's a diamond," he would be saying something false. If he believed what he said, he would have a false belief as well. And so it is with Zowie.

With Twin Zowie, things are reversed. When she uses the word "diamond," she doesn't mean *diamond*. She has never seen a diamond, nor has she met anyone who has seen one. In fact, she has had no contact with diamonds at all, no matter how indirect. When she says "diamond," she is using it to refer to what Twin English speakers normally refer to when they use the same word: twin diamonds. If her jewels are genuine twin diamonds, then what she says is true. And if she believes what she says, what she believes is true as well. What she believes is something we might approximately express by saying that her engagement ring is studded with twin diamonds. That is certainly different from what Zowie believes, so their beliefs have different contents.

Zowie and Twin Zowie are identical in all intrinsic respects. They differ only in their relationship to diamonds. So if the above reasoning is along the right lines, it shows that cognitive content depends partly on factors

external to their subjects and so are partly relational. However, one might be tempted to doubt the conclusion and hold that the contents of the twins' beliefs are the same. After all, it is hard to see how the difference between diamonds and twin diamonds, a difference of which the Zowies are quite unaware, can make any difference to how the world appears to them or to how they think and reason.

Throughout this book, I will discuss arguments on both sides arising from this kind of Twin Earth experiment. For now, I will use it as a way to introduce some of the main ideas that will feature as we proceed.

1.4 SUPERVENIENCE

Twin-world experiments are fundamentally about supervenience. Philosophers have refined a number of useful senses of “supervenience” (see Kim 1984 for discussion). But I will just stick with a simple and rough one: a set of properties *B* supervenes on a set of properties *A* if and only if (iff) any two objects identical in respect of *A* properties must be identical in respect of *B* properties too. Weight, for example, supervenes on mass and local gravity: any two objects of the same mass, subject to the same local gravity, must have the same weight. Weight does not supervene on size, however, since two objects of the same size may have different weights. If the twin Zowies differ in respect of the contents of their beliefs, then these contents fail to supervene on intrinsic properties. (Remember that, by hypothesis, the twins have identical intrinsic, microstructural properties.) By contrast, if any possible twin of Zowie, no matter what her external environment is like, must share all her cognitive contents with Zowie,

then cognitive contents do supervene on intrinsic, micro-structural properties, or to use a common abbreviation, they are “locally supervenient.”

The question of whether content is locally supervenient is not quite the same as whether it is intrinsic to the subject. There are at least two reasons for this. The first is that putting the issue in terms of local supervenience on microstructure leaves no comfortable place for a Cartesian dualist to enter the discussion. In effect, it assumes that the subject of cognitive contents—the object to which cognitive properties are or are not intrinsic—is a physical thing, in the minimal sense of being made out of atoms, molecules, and so on. This is not an assumption the Cartesian would share. A Cartesian who believed in the intrinsicness of content would not need to hold that Zowie and Twin Zowie are cognitively exactly similar: whether they are is a question about their immaterial souls, not their material brains and bodies. The Cartesian might or might not believe in the local supervenience of the mental on the physical. But that would not bear directly on the question of whether mental properties are intrinsic properties of their immaterial subjects. (It is often claimed that Descartes himself believed that the mental is intrinsic; whether he did or not is a question that I refrain from addressing.)

The second reason why the questions of local supervenience and intrinsicness come apart is that a property might be relational, at least in a weak sense, and yet be locally supervenient. This is rather obvious if one thinks abstractly: there does not seem to be any principled reason why some relational property *R* should not be locally supervenient. That would just mean that any twins would necessarily be identical in respect of *R*: they would either all have it or all lack it. As an illustration, not to be taken

too seriously, suppose that height supervenes on micro structure. Then any twins are the same height. But having a given height automatically puts one into a relation with a given number: being six feet tall, for instance, puts one in a relation with the number six, specifically the relation “is ____ feet tall.” Being six feet tall is thus locally supervenient—it is shared by all twins—but also, in this weak sense, relational.

Someone might reasonably hold that content is indeed both locally supervenient and relational if they held that contents are relations to abstract objects, such as properties. One might think, for example, that thoughts about diamonds involve relations to the property of being a diamond, where a property is an abstract object that exists independently of its instances. Then any being thinking about diamonds will stand in a relation to this property, even if there are no diamonds in its environment. Colin McGinn (1989) calls this position “weak externalism.” Whether weak externalism is true is an interesting question, but is irrelevant to the topic of this book.

This book is concerned with whether content essentially involves relations to external, concrete, contingently existing things. The Twin Zowie story illustrated the idea that some sort of relation to samples of the kind of thing a concept represents is necessary for possession of the concept. The story raised the question of whether a relationship with diamonds is necessary for having a concept of diamonds. And many have thought that this is so, that concepts of natural kinds, like diamonds, do require some real relationship with actual instances. It has also been argued, by Tyler Burge (e.g. 1979) in particular, that certain relations to other language users are determinants of content. For these sorts of items—the kinds that concepts represent and other language users—being intrinsic and being locally supervenient coincide.

For if content is locally supervenient, then it will always be possible to conceive of a possible environment in which someone has a state with the relevant content but in which the items do not exist. Take subject *Z* in a certain representational state *S* that represents some kind *K*. We can always imagine a microstructural duplicate of *Z* in an alien scientist's vat and in a world devoid of diamonds (trees, water, tigers, or whatever) and devoid also of other speakers. Or if some glitch comes up with that kind of example, we can imagine that the twin arises as the result of a quantum accident: he or she suddenly emerges in outer space and survives for a short while, floating in the void. If *S* is locally supervenient, then the twin would be in state *S*. And this would entail that *S* is intrinsic, rather than relational, for in the twin's environment there are no *K*s for it to be related to.

My aim is to argue for the local supervenience of content. Given what I have just been saying, and given the minimally materialist assumption that bearers of cognitive properties—humans, animals, cognitive systems of all kinds—are made up out of elementary particles, arguing for local supervenience is a way of arguing that content is intrinsic. The position I will be defending is a version of what is called “internalism” or “individualism,” and I will use both labels to refer to it. However, both labels are vague and should be taken to gesture towards a family of positions rather than any very specific thesis.

The thesis with which I will be mainly concerned is then this: being in a state with a specific cognitive content does not essentially involve standing in any real relation to anything external. Cognitive content is fully determined by intrinsic, microstructural properties: duplicate a subject in respect of those properties and you thereby duplicate their cognitive contents too.

Internalism, the thesis that content supervenes on microstructure, thus embodies two ideas. The first is that content is not relational, does not depend anything outside the subject. The second is that it does depend on microstructure. The latter idea is not without substance, since an alternative would be that content does not depend on anything. Zowie believes that diamonds sparkle. Suppose now that Zowie has a twin who is identical both in microstructure and in relational properties. Perhaps this twin Zowie does not believe that diamonds sparkle. If this were possible, having that belief could still be intrinsic to Zowie; it is just that it would not depend on certain other of her intrinsic properties, the microstructural ones. This is a possibility that some courageous people might accept (see, for example, Crane and Mellor 1990 and Cartwright 1994). This courageous position is closely related to the Cartesian one, although it can allow that psychological properties are properties of bodies or brains. A consequence of my arguments for local supervenience will be that this position should be rejected.

The internalism I will argue for likens content to properties like being hard, being brittle, or being liquid, in their intrinsic versions. One can explain a thing's possession of these properties in terms of properties of and relations among its constituent parts. A diamond is hard because of the way its crystals are bound together. A ceramic cup is brittle because of irregularities at the boundaries of the crystals that make it up. The water in a glass is liquid because of the way its molecules are loosely bound together. This sort of explanation, which could be called "systematic" (loosely to follow the usage of Hauge-land 1978), seems to be fairly widespread. It applies to

functional properties like the ones we've just considered, along with being transparent, being plastic, and so on.

Systematic explanation is the norm when it comes to explaining how artefacts work: when one explains how a car or an espresso machine or a dishwasher works, one cites what its component parts are, what each one does, and how the combined actions of the various parts suffice for the machine to do what it is supposed to do (see Haugeland 1978 and Cummins 1983 for more on this). It applies in the explanation of many different kinds of natural properties. For example, thermodynamic properties of gases are explained in terms of properties of and relations among the component molecules. And biomedical properties of hearts, lungs, and so on, are explained in terms of properties of and relations among their parts (the auricle, ventricle, etc.).

The idea, then, is that cognitive properties, like so many others, can be given systematic explanations in terms of properties of and relations among their bearers' parts. I have largely resisted the temptation to say "physical" parts, since I find that term unhelpful at best. The word "physical" tends to be used by those—physicalists and dualists alike—who think that subject matters of intellectual inquiry divide in some principled way into the physical, the not-yet-shown-to-be-physical, and (possibly) the nonphysical. This is wrong.

I do not know which properties and relations of which parts are the relevant ones for explaining cognitive properties. They might be functional properties of and relations among neurons, of a sort within the descriptive reach of current neurology. They might be computational properties of neurons, of a sort within the reach of current computer science. Or they might be—and I suspect they are—as yet undreamed of properties of some kind of neu-

ral wotsits to be discovered by some future science that develops at the overlap of neurology and psychology.

The advantage of focusing on microstructure, that is, on the level of elementary particles, is that whatever it is that determines content probably supervenes on it. Fix an object's microstructure and you fix its atomic and molecular structure, its neurological and computational properties, and so on. Or so I will assume, anyway, to facilitate exposition.

1.6 THE HARDNESS OF THE SUPERVENIENT "MUST"

The local supervenience thesis—if two beings are identical in respect of their microstructural properties, then they must be identical in respect of their cognitive contents—can be interpreted in different ways, depending on the strength of the “must.” The point calls for discussion.

Internalism is often held to involve a notion of “metaphysical necessity.” Metaphysical necessity was first described clearly by Saul Kripke (1971, 1980). For illustration, consider the (true) identity statement that Hesperus = Phosphorus (“Hesperus” and “Phosphorus” being two names for the planet Venus). It is metaphysically necessary that Hesperus = Phosphorus. It is not logically necessary, since no amount of purely logical deduction could reveal its truth. It is not epistemically necessary, since someone might perfectly well not know that Hesperus is Phosphorus (the ancient Greeks didn't). However, it is necessary in a very strong sense: it could not have been otherwise, no matter what. No matter how different the world might have been in respect of the laws of nature or anything else, Hesperus could not have been a different object from Phosphorus. Since Hesperus is the very same object as Phosphorus, it could not possibly have been

some distinct object, for then it would have been distinct from itself, which is clearly impossible.

Identity statements involving kind terms and certain sorts of statements about a kind's constitution have also been seen as metaphysically necessary. For example, some would hold that it is metaphysically necessary that diamonds are made of carbon. Try to conceive of a possible world in which diamonds are not carbon, and you will fail. You can conceive of a world with twin diamonds, stones that resemble diamonds but have a different molecular constitution, but twin diamonds are not diamonds. So some would argue (Kripke 1980).

An internalist who held that microstructural duplicates must, as a matter of metaphysical necessity, be cognitive duplicates, would hold that all possible twins are in the same psychological states. If Zowie believes that orangutans are omnivorous, then there is no possible world, however different from the actual one, inhabited by a twin Zowie who doesn't share the belief. As I said, internalism is indeed sometimes said to involve a notion of metaphysical necessity. I do not wish to defend quite such a strong thesis.

The main reason for this is that I have a worry about the methodology of assessing claims for metaphysical necessity. The standard way to find out whether a given proposition is metaphysically necessary is to try to conceive of a possible situation in which it is false. But it is not obvious why there should be any very strong connection between what we can conceive of, or imagine, and truths about the world. It is not obvious that if one cannot conceive of something's being the case, it follows that it really could not be the case. Whether the entailment goes through depends at least in part upon what accounts for our capacities to conceive, on how our conceptual faculties function. And this is not a subject about which much is known.

Further, one should doubt that we always know what we can and cannot conceive. There are those who would claim that they can conceive of diamonds not being made of carbon. Kripke might argue that these people are confused about what they can conceive. Perhaps that is correct. But that would just show that we are not infallible about the content of our own conceptions. Of course, it remains possible that we are reasonably reliable judges of our own conceptions, and that there are cases where our judgements about them can be trusted. But this optimism may be premature. Once again, until we have a better theory of our conceptual faculties, of how they work, we should be cautious about our judgements.

Putting the two points together reveals a gap between what we think we can conceive and what is objectively possible or impossible. There is no sure route from the former to the latter. Since conceivability is the chief method of assessing claims of metaphysical necessity, I think such claims are incautious. Hence it seems that it would be incautious for an internalist to make any claims about metaphysically necessary supervenience.

I will assume that the question of whether a property is relational or intrinsic is a question about natural necessity, or laws of nature. If it is nomologically possible for twins to have different contents, then contents are relational and not fully explicable in terms of microstructure. The thesis that nomologically possible twins must have the same contents is an interesting internalist thesis in itself, and one worth arguing for. For it shows us what content is actually like, no matter what the metaphysical possibilities are.

Here is an analogy. As I mentioned above, according to the theory of Special Relativity, properties like mass, size, and shape turn out to be relative to frames of reference. Special Relativity is an empirical theory that tells us

about the real natures of mass, size, and shape. Nevertheless, it seems conceivable that Special Relativity might have been false. It seems as though one can conceive of a world in which the theory does not apply. But surely this apparent conceptual possibility tells us nothing interesting about what mass, size, and shape are actually like.

There are three reactions we might have to this thought experiment. We might hold that it tells that, e.g., mass might have been intrinsic, although it is actually relational. The relationality of a property would then be not essential to it. Or we might hold that the thought experiment tells us nothing about mass, but reveals the metaphysical possibility of a nomologically impossible world in which objects don't have mass but have instead some counterpart property, *shmass*. *Shmass* might be rather like mass, but it wouldn't be the genuine article. Or we might hold that our thought experiment only presents a world in which Special Relativity would seem to be false, a world in which we would be misled about the laws of nature and the nature of mass. (See Shoemaker 1998 for discussion.)

On all three views, the interesting truth about mass—the fact that it is relational—relates to the natural laws that govern it. The second and third views might allow that the relationality of mass is also metaphysically necessary. But this would only be because the metaphysical necessities flow from the natural, nomological ones. Hence on all three views, the focus of interest would be on natural laws and empirical theories.

So, from now on, I will standardly use the various notions that involve modals like “must” in line with the above, to invoke natural or nomological necessity. Thus local supervenience is the thesis that microstructure nomologically determines cognitive properties, that twins are nomologically possible twins, and so on.

Content that is locally supervenient is often called “narrow”; content that isn’t is called “wide” or “broad.” I will argue that narrow content is genuinely representational, honestly semantic. A number of internalists have argued for varieties of narrow content that are not directly representational but relate to representation proper only indirectly. For example, Jerry Fodor (1987) once argued that narrow content is a function from contexts to broad contents. To illustrate, consider Zowie once again. On that position (which Fodor no longer holds; see Fodor 1994) it is conceded that Zowie and Twin Zowie have thoughts with different semantics: one being about diamonds, the other about twin diamonds. So they have thoughts with different broad contents. But the thoughts also share a narrow content in virtue of the following: if Zowie had been in Twin Zowie’s context, her thought would have been about twin diamonds, and if Twin Zowie had been in Zowie’s context, her thought would have been about diamonds. And indeed, for any context *C*, Zowie and Twin Zowie would have thoughts with the same broad contents had they been in *C*. Thus the thoughts instantiate the same function from contexts to broad contents, and this is what gives the same narrow content.

Others have argued for a “functional role” theory of narrow content. A state’s functional role is given by its causal potentialities: roughly, what would cause it to occur, and what, in turn, it would cause to occur, in various possible circumstances. So two states have the same functional role if their potential role in the causal nexus is the same. Some functional-role theorists hold, as Fodor did, that the representational contents of twins’ states are different, and hence broad. But they construct a notion of

narrow content in terms of functional role and hold that twins' states have the same narrow content because they have the same functional role.

Although I sympathize with much of the motivation for these versions of internalism, I wish to defend a different one. The version I will defend holds that narrow content is a variety of ordinary representation. Narrow content is just content, to be understood in roughly the terms it always has been understood (at least since Frege), such terms as sense, reference, truth, extension, "satisfaction" in the technical, Tarskian sense, "aboutness" in the philosophers' sense, and "intentionality" in roughly Brentano's sense. My view mandates denying that Zowie and Twin Zowie's "diamond" concepts have different extensions. In fact, my view is that both Zowies' diamond concepts apply to both diamonds and twin diamonds, so contrary to what some might initially think, if Zowie pointed to a twin diamond and said "That's a diamond," she would be saying something true in her idiolect.

I will not, however, argue that all representational content is narrow. It is necessary to make an exception for singular, demonstrative concepts such as those expressed by "this ring." The basic idea is simple. Suppose that Zowie and Twin Zowie have exactly similar engagement rings. Each twin points to her ring and thinks *This ring is beautiful*. It is clear that the referents of the demonstrative concepts are different. For reasons I will come to in chapter four, this leads to a legitimate notion of wide content. But it is arguable that this sort of wide content is not properly classed as psychological or cognitive.

I will not offer any "philosophical theory" of content. Trying to develop a philosophical theory of content is rather like trying to develop a philosophical theory of heat or water. Content is a real natural phenomenon. The most we should hope for by way of a "philosophical" theory of

content is one that will tell us what content *is*, where the “is” has roughly the force it has in “Heat is the motion of particles” or “Gold is the element with atomic number 79.” It may be that we are even asking too much when we ask for a theory of content in this sense. Maybe there is no answer to the question “What is content?” After all it is pretty unclear how to give a general account of heat that applies across solids, plasmas, and vacuums.

Be all that as it may, there is certainly a great deal we can discover about the nature of content and about how the organization of matter gives rise to it. The point is that we should not expect to discover too much from the arm-chair. Discovering the true nature of content should be a scientific enterprise (whether we also call it “philosophical” or not). These enterprises usually progress slowly because they involve a great deal of empirical and technical work. It might take many decades of detailed research before we make real progress. So a defense of internalism should not require defense of a theory of content. Proposed theories may well be premature. It should suffice to cast reasonable doubt on externalism, to motivate internalism, and to provide reasons to believe that good psychology is, or could be, internalist.

1.8 TACTICS

I do not claim to have any convincing, knockdown argument for internalism. I do not claim to have conclusive arguments against all varieties of externalism. But I think I can make a decent case for internalism. The basic strategy is to undermine what I think are the most popular and influential externalist theses and to show that an internalist alternative is workable and attractive.

What follows comes in four chapters. The first two offer arguments against two leading externalist theses. In chapter 2, I address externalism about natural-kind concepts and present an argument against it based on the existence of empty kind concepts, ones that lack an extension. In chapter 3, I present an argument against a popular form of “social externalism,” the view that the content of many concepts depends in part on the views of experts, lexicographers, and so on. The message of the first two chapters will be that all general (nonsingular) concepts have a narrow content. The considerations leave open whether general concepts also have broad contents. Chapter 4 considers and rejects the leading two-factor theories that endorse both broad and narrow content for general concepts. Chapter 5 outlines and defends a radical alternative version of internalism, arguing that extension conditions are narrow.

