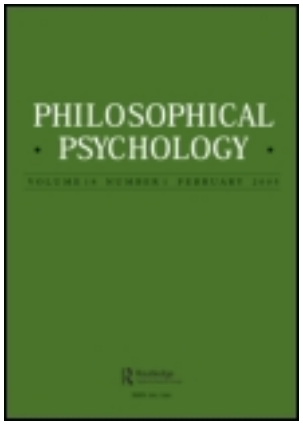

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

Embodied Targets, or the Origins of Mind-Tools

Jan Slaby, Graham Katz, Kai-Uwe Kühnberger & Achim Stephan

Version of record first published: 19 Aug 2006.

To cite this article: Jan Slaby, Graham Katz, Kai-Uwe Kühnberger & Achim Stephan (2006): Embodied Targets, or the Origins of Mind-Tools, Philosophical Psychology, 19:1, 103-118

To link to this article: http://dx.doi.org/10.1080/09515080500471884

Philosophical Psychology
Vol. 19, No. 1, February 2006, pp. 103–118

REVIEW ESSAY

Embodied Targets, or the Origins of Mind-Tools

Jan Slaby, Graham Katz, Kai-Uwe Kühnberger, and Achim Stephan

Philosophy of Mental Representation
HUGH CLAPIN (Ed.)
Oxford, England: Oxford University Press, 2002
332 pages, ISBN: 0198250525 (pbk); $35.00

1. Introduction

In the cognitive science era, in which philosophers frequently make reference to research in the fields of neurobiology, artificial intelligence, and cognitive psychology to motivate the plausibility of their arguments, one might expect a philosophical discussion of mental representation to busy itself with talk about the kinds of representations hypothesized by people whose job it is to model the mental. And, in fact, when philosophers truly focus on mental representation, they do tend to do this. But the concepts associated with mental representation—cognition, intentionality, aboutness, objectification, commitment, situatedness—are so appealingly distracting, that it is possible to read well over one hundred pages of a book entitled *The Philosophy of Mental Representation* before the discussion turns to the eponymous theme at all. The book in question, edited by Hugh Clapin (2002), is a record of the proceedings of a six day workshop held in rural Maine in August 1999.[1] It is a fascinating and thought-provoking document, not least because the participants in the workshop were Andy Clark, Robert Cummins, Daniel Dennett, John Haugeland and Brian Cantwell Smith, all eminent philosophers of Cognitive Science with much to say and no fear of saying it. Part of the reason that the discussion ranges so widely,

Correspondence to: Achim Stephan, Institute of Cognitive Science, University of Osnabrück, D-49069 Osnabrück, Germany. Email: acstepha@uos.de

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/06/01000103-118 © 2006 Taylor & Francis
DOI: 10.1080/09515080500471884

of course, is that many of these philosophers are dubious about the central role of mental representations in cognition, and are eager to talk about the things that they do find central.

Aside from the extensive overview introduction and integrative concluding chapter contributed by the editor, each of the book's core five chapters is devoted to the selected writings of one of the invited philosophers. The philosophers do not present their own work, but rather provide commentary on each other's. So, Haugeland comments on Clark's work, Clark comments on Dennett's, Dennett comments on Smith's, Smith comments on Cummins', and Cummins comments on Haugeland's. The subject of the commentary gets to respond, of course. These commentaries and responses, along with an edited transcript of the wide-ranging discussion accompanying the presentations and a brief introduction by the editor constitute the chapters.

As to be expected, the reader gets an exciting package of ideas, theses, positions, as well as some controversy and some expected and some unexpected agreement. But, and also not surprisingly, the unusual form of the book has an effect on its content. What we end up with are mostly *advertisements* for the participants' theories, but not the theories themselves—the book manages to excite, to awake interest, to make one rush to the bookstore or the library, but it doesn't provide material that stands on its own. Fortunately, the relevant texts of each discussant are referenced at the beginning of each chapter. Reading them is not merely an option—it's *mandatory* if you want to come to a sufficient understanding of the topics discussed to follow the discussion. The topics raised include—beside the main theme ''representations''—intentionality, normativity, evolution, computation, external scaffolds, mind tools, learning, the problem of error (and many more). We are treated not only to a variety of subjects and perspectives on them, but also to a confrontation between different *conceptions of philosophy*: the clash between methodological naturalism and a contemporary version of transcendental philosophy.

The self-understanding of most philosophers working in the vicinity of the cognitive science enterprise is *Quinean*: Philosophy is seen to be on a par with empirical science, not as an exclusive domain of foundational reflection. Methodological naturalism has replaced transcendental philosophy. A sort of consensus scientific realism seems to make exercises in metaphysics or ontology unnecessary, and philosophers happily join the more empirically-oriented scientists on Neurath's ship. Dennett is a prototype of this approach to philosophy, as he draws on all sorts of empirical results from various disciplines in order to present an integrated and largely ''metaphysics-free'' picture of mental content and consciousness. Clark and Cummins work under this conception of philosophy as well.

Transcendental philosophy is the attempt to reconstruct, from within, the most general constituents of empirical thought and knowledge. The leading question is: What needs to be in place, on the side of experiencing subjects, in order for their experience to count as that of an objective, mind-independent world? And since the days of Kant many transcendental philosophers agree that ontology and metaphysics—the delineation of the basic structure of reality in the broadest sense of

the term—can't be done without regard to a prior reflection on the *conditions of possibility* of knowledge and experience.

So, when Smith, after years of designing software for various sorts of practical applications, presents a manifesto in ontology called *On the Origin of Objects* (1996), or Haugeland, after acutely charting the field of AI, concludes that we need an account of how the *existential commitment* to *constitutive standards* grounds objectivity and truth in a community of skillfully engaged agents, this unexpected revival of substantive metaphysics can't just be waved off as merely nostalgic.

At this point, the notion of *representation* comes into play. The leading cognitive science paradigm maintains that cognition essentially involves or is based on the capacity of agents to *represent* their world. And whether practitioners are aware of this or not—this essential dependence of cognitive science on representation has deeper philosophical implications. That's because 'representation' is just the now-fashionable, scientifically approved term for the ultimately significant relation between experiencing agents and the world they experience. Whatever else you may add to the core idea, when it comes to the representation of a putatively objective, representation-independent world, you are already moving within the domain of transcendental philosophy, whether you call yourself a cognitive scientist or not.

Smith claims that cognitive scientists approaching their domain of inquiry without the least ontological concerns are prone to commit what he calls ''inscription errors'' or ''pre-emptive registration.'' Pre-emptive registration is what results from the common-sense realist's naïveté: the projection onto ultimate reality of what is in fact an artifact of some later, higher-level development (Smith, 1996, p. 50). So instead, Smith claims, we have to open our eyes to the very process of registration itself—the process by means of which intentional beings *abstract* objects out of ''the infinitely rich ur-structure of the world'' (p. 241). This is nothing less than the strong recommendation to rethink intentionality in a very fundamental way—a way that includes taking neither ''subjects'' nor ''objects'' simply for granted.

Haugeland agrees with this basic agenda. He thinks we miss out on a specific feature of *human* cognition when we abandon all transcendental concerns: Only humans, Haugeland claims, are capable of a particular sort of objective knowledge, one that is based on a special kind of appearance/reality distinction. His idea, inspired by Heidegger's notion of ''*Seinsverständnis*'' (understanding of being), is that this distinction is grounded in our capacity to resiliently *commit* to a set of standards that putatively govern what can *possibly* be real. Once an instance of experience seems to violate these standards, we outrightly reject that instance as impossible. Only when the standards are continuously breached in a systematic way, will we eventually agree to adjust them (p. 36).

So Kant and Heidegger were probably onto something that is of utmost relevance to a proper understanding of the human capability for objectivity and truth-telling. Since the chapter devoted to Haugeland's proposals, which are perhaps the most original and exciting in the book, is largely hijacked by Cummins, we'll start with a more detailed reconstruction of them, and then relate other important issues at the appropriate places.

2. Haugeland: Constitution and Existential Commitment

In his recent work, John Haugeland tries to combine two tasks: One is epistemological and rather traditional, the other is metaphysical and more closely tailored to current tendencies in cognitive science.

> **Epistemological task**: How can experience be of a world that is *external* to our representational schemes, but at the same time *intelligible*, so that it can serve as a *normative* constraint on our representational practices, not merely as a causal force that brutely impinges on our representational systems from the outside?

Haugeland thinks that this task is best approached from the perspective of another task, one that emerges from recent attempts in cognitive science to capture the essential *embodiment* and environmental *embeddedness* of mind and cognition:

> **Metaphysical task**: How can we specify the relation of mind, body, and world without falling back into problematic dualisms, but also without shunning relevant distinctions? How can the *intimacy* of mind, body, and world be spelled out in a sane way?

The commonalities between these two agendas are easy to spot: Both search for an alternative to an all too common philosophical framework that places mind and world, subject and object far adrift on opposing sides of a gulf which it then tries to bridge somehow. The Cartesian conception of the mind, centered around the idea of representation, has been standardly invoked as the ultimate means for connecting the two poles. Haugeland's alternative starting point is the Heideggerian concept of "skillful engagement"—structured and rule-governed activities that are essentially *world-involving*. While it always seems to be an open question whether putative representations do in fact refer to something that is mind-independent, it is hard to pose a similar challenge for the notion of skillful engagement. The very idea of skillful engagement *presupposes* an interdependence of agent and world, of subject and object.

The epistemological task is to show how objects can stand as independent, external constraints on mundane activities, without having to view them as standing outside an outer boundary of what is intelligible. Transformed into the context of the philosophy of mind, this reads: How is the directedness to genuine objects—intentionality—possible? Haugeland's most important move, at this point, is to bring in a specific kind of *constitutive rules* and a corresponding, specific kind of *rule-following*: What he calls "constitutive standards" and "constitutive commitment." 'Constitutive' here means that these rules define or make up the phenomena that they govern, in other words, that these phenomena would not make sense independently of those rules (whereas *regulative rules* are those that govern phenomena that already make sense independently of the rules).

Constitutive standards are those standards that govern all the phenomena in a given domain of objects. Haugeland illustrates this by discussing games although his ultimate target is empirical science, conceived in a somewhat Kuhnian way. The standards specify what's possible and what's not possible in a given domain

of objects. They say both how the players (or the scientists) and how the objects involved are supposed to behave: In order to engage in a game of baseball, you need players that behave in accordance with the rules of play, and, in addition, proper equipment that fulfill certain standards, otherwise the game isn't playable. So constitutive standards are essentially more than the constitutive regulations of the players' or participants' behavior; they are the sum of the *conditions of possibility* that have to be fulfilled in order for a certain mundane activity to be engageable at all (cf. Haugeland, 2000, pp. 48–49).

Now consider science. A scientific paradigm can be thought of as a set of constitutive standards that putatively govern a domain of objects. The scientist has to have the ability to reliably *tell* whether given phenomena in fact do accord with the standards or whether they are breaches of those standards. That's the point of experiments: Check whether phenomena that are strictly excluded by the constitutive standards in fact don't occur. But as all practitioners in science are painfully aware, good, valid experiments are a tricky business. They are hard to devise, things can go wrong—consistent results are hard to achieve, since there's virtually always interference and noise. On what grounds, Haugeland asks, do scientists persist in the face of such empirical recalcitrance? Here's his answer: What holds the constitutive standards firmly in place is the practitioners *commitment* to the overall theory they currently hold. Only on the basis of this theory they can actually "triangulate" on their objects via distinct and disparate experiments (p. 35).

It's unfortunate that Haugeland doesn't elaborate much upon this idea in the book, since the essential connection to the issues of objectivity and normativity doesn't become particularly clear in those scattered passages. The connection seems to be this: Such a commitment to standards doesn't boil down to a kind of cognitive blindness, an inability to see things differently than one's current theory claims they are. Quite to the contrary. Commitment entails *responsibility*, and here, particularly, the responsibility to change or abandon the standards in the face of consistent, clearly localizable breaches. Through the very commitment to constitutive standards, objects are finally granted their much sought-after *normative authority*:

> The normative authority of objects, by virtue of which they can stand as criteria for the correctness of mundane results—and thus as binding on judgments and assertions—devolves upon them from the commitment to the standards in accord with which they are constituted. (Haugeland, 1998, p. 343)

This completes the epistemological task, which would have deserved to find more resonance in Clapin (2002).

3. Clark: Animals With and Without Norms

Andy Clark is widely known for prospectively describing and actively promoting the paradigm shift in cognitive science from good old fashioned artificial intelligence (GOFAI) to embodied and embedded cognition. GOFAI had its merits, to be sure, but in general cognition has to be viewed as essentially world related, otherwise we

might miss out on what's most important. The trouble with the research agendas of classical cognitive science and classical AI is, according to Clark, that its high-level problems (e.g., theorem proving, playing chess, translating language) tend to distract attention from the nature and function of most real life cognition. That's because the basic function of cognition is the advantageous adjustment of behavior to environmental circumstances, not high-level math or chess matches against the world champ. And since *homo sapiens* evolved from the anthropoids not so very long ago, his native endowments (or "wetware") can't be all that different from theirs. Thus, it makes much sense to study animals and their cognitive capacities—more sense, at least, than trying to figure out how to create a computer that writes poetry.

Clark's picture has (higher) animals and humans roughly on the same level when it comes to what's "in the head," while humans differ mainly in their ability to create and maintain a large variety of what he calls "external scaffolds"—i.e., external structures, mostly artifacts, that aid cognitive processes to an extent that makes them genuine components of cognition themselves. Examples are language, maps, social institutions, but also pen and paper for doing calculations, as well as various other tools and items that assist or facilitate cognitive processes. Clark's basic point is: If an evil demon took all these external augmentations away, there wouldn't be much left worthy to be called 'human cognition'—there's not much the naked brain could do without its external props and aids.

Haugeland's commentary is an illuminating overview of Clark's position, and he takes issue with the one aspect that stands out to be fundamentally at odds with what he himself claims, namely the thesis that there is no great *biological* divide between human and animal cognition, and that the fundamental difference is rather of *cultural* origin. Unfortunately, both Haugeland's criticism and Clark's reply fail to present and discuss these issues with a sufficient degree of clarity. As will emerge presently, it's not just the *hard* questions that are left open.

Here is Haugeland's point. It is because of Clark's strong affinity to animal research that he—and, for that matter, many of his colleagues in cognitive science as well—overlook an all-important constituent of especially *human* cognition: The pervasive and essentially human capacity *to abide by communal norms*. And it is this very capacity that does in fact open up the "great divide" between humans and all the other animals that Clark and many others so firmly deny (cf. pp. 29–30).

But, to bring home his argument, Haugeland would have to show that and why the capacity for norm-abiding doesn't in fact result from a more general capacity to create and maintain external scaffolds (particularly language). He had to show that, to the contrary, *social normativity* is precisely what enables all that scaffolding. On this, Haugeland offers only a very rough sketch—he merely hints at what he has described in more detail elsewhere (e.g., Haugeland, 1998, pp. 147–156): At the root of all social normativity lies the complex meta-disposition of conformism. Conformists tend to *imitate* and to *sanction* the behavior of others according to whether it conforms to what all or most community members are standardly doing (reward) or whether it deviates (punishment). Conformism creates bodies of communal norms. Haugeland calls this effect "normative gravity."

Not very surprisingly, Clark denies being norm-blind at all. According to him, it is language that makes norms and normativity possible and not the other way around: ''I think that norm-hungriness is a secondary effect of getting language going—and in order to get it going you only need to be as norm-hungry as a chimp'' (p. 54). That's where language comes in: Public language is a prime kind of ''found encoding'' that serves to turn such complex regularities into simpler objects for further thought. Linguistic entities thus function as cognitive objects that significantly enhance the computational capacities of cognitive systems.

The resulting process of *objectification* in which a set of relatively simple tokens stands in for complex features and relations opens up new, quasi-perceptual ''spaces for reasoning''. Finally, Clark applies this idea to the domain of norms and normativity: The space of norms is artificially constructed through the use of such tokens and stand-ins. There's much more to say and explain here, of course, but the upshot is clear. If Clark's story is anywhere near the truth, then there is no need to postulate a brain-based capacity for norm-abiding that would set humans apart from other animals. It would be language (or other kinds of ''found cognitive items'') first, norms and normativity second.

These ideas and proposals are exciting and innovative—both Haugeland's story about commitment and about the peculiar human capacity of ''norm-hungriness,'' and Clark's objectification scenario. However, as presented in the book, neither of them can persuade the critical reader. Haugeland basically owes us an explication of commitment that doesn't contain or presuppose any normative notions, whereas Clark leaves us in the dark about how such higher-order recoding itself might function. Accounts that simply *presuppose* language as given certainly beg all the relevant questions about normativity.

4. Dennett: It's Tools All the Way Down

Dan Dennett's more recent ideas on representations, minds, tools, and mind-tools are introduced by Clark in a fictitious dialogue with Dennett (see p. 67):

> What are minds made of?—Tools for thinking! Who or what *uses* the tools to *do* the thinking?—No one, nothing. The tools-R-us. *Intentionality*, *aboutness*, *content*, and *consciousness*: Can all these really be brought into being by grab-bags of userless tools?—Yes.

The following 23 pages of the chapter ''Minds, Brains, and Tools'' are meant to show what these questions and answers mean. And Dennett is so happy with Clark's presentation of his ideas that he is done after a two-and-a-half page reply.

One of many possible starting points to Dennett's work is his deconstruction of the so-called ''intellectualist myth,'' according to which minds definitely are not made of tools, but composed of such ''episodes as the thinking of private thoughts, the consultation of rules and recipes, the application of general rules to particular circumstances, and the subsequent deduction of implications about those particulars'' (Dennett, 1983/1987, p. 213). Dennett shows up here as a true heir to

his philosophical father Gilbert Ryle, who half a century ago already pointed to the problems we run into if we assimilate "knowing how" to "knowing that." In fact, efficient practice always precedes the theory of it. To Dennett, the classical approach to cognitive science has set itself so defiantly against Ryle that he dubs it "intellectualist science." It "speaks openly and unabashedly of inner mental representations and of calculations and other operations performed on these inner representations" (1983/1987, p. 214). Interestingly enough, positions inspired both by Rylean analytic philosophy such as Dennett's (and Clark's) and by Heideggerian existentialism such as Haugeland's come close together on that topic. And it is again Ryle who had already claimed that proponents of the intellectualist picture will run into a circle if they stick to the intellectualist legend.

A similar argument stems from Dennett. However, before we discuss it we have to pay attention to several distinctions Dennett proposes with respect to the concept of representation. In his "Styles of Mental Representation" (1983/1987), Dennett distinguishes between "explicit," "potentially explicit," "implicit," and "tacit" representations; in his more recent work, the distinction between "florid" and "pastel" representations is added. According to Dennett, information is represented *explicitly* in a system if and only if there exists a physically structured object, which is interpreted as tokening of an element of a system (e.g., a formula), and a device for reading or parsing it. Information is represented *implicitly* if it is implied logically by something that is stored explicitly. It is *potentially explicit* if the system is able to change it into an explicit representation (cf. 1983/1987, p. 216). In contrast, tacit information "has to be built into the system in some fashion that does not require it to be represented (explicitly) in the system" (1983/1987, p. 218), it is "built-in knowledge," i.e., "know-how." As developed in "Making Tools for Thinking" (2000), *florid* representing can be characterized as deliberate representing and knowing representing. Since knowing that we are representing involves having a mental state that is itself *about* the representing we are doing, it is also self-conscious representing, hence, florid representing involves *meta-representation*. Dennett is not so sure what *pastel* representation is in contrast to such florid representing: just perceiving? In such cases, one has a mental or perceptual state that is surely *about* something (*qua* this something), and if that perceptual state plays a role in guiding one in the right way, this state should count as a representation *in use*. This might be representation enough for many.

Let's come back now to a sketch of Dennett's argumentation about how the diagnosed regress intellectualist cognitive scientists run into could come to a halt in a finite number of steps. In the end, merely *tacit knowledge* has to terminate it:

P1. Explicit representations as such are not information bearers; they get their content only through roles in larger systems.
P2. An item gets a role in a system if the system is set up in a way as to do things (e.g., to go from one representation to another).
P3. The system's setup is given (i) by rules and recipes that are explicit representations (which would lead to an infinite regress) or (ii) by some tacit

know-how which can put the explicit representations to work without requiring more layers of representations.

C. Tacit "knowledge" (non-representational skills and know-how) is more fundamental than internal representations.

It's this thinking that more or less translates to the claim that minds are made of tools, because tools (i.e., tacit knowledge) are more fundamental for exhibiting mental properties and capacities than internal representations. Know-how is possible without internal representations but not *vice versa*. We, however, are not so sure whether or not there is not a parallel circle for tools and skills. Consider the following alternative sketch of an argument:

P1. A skill/tool as such has no value at all; its exercise does not count as "mindful" if it is not used in the appropriate circumstances.

P2. To figure out what the right circumstances are the system would need (i) some further skills and tools (which would lead to an infinite regress) or (ii) . . .?

C. Tacit "knowledge" (non-representational skills and know-how) does not stop regressing infinitely.

To avoid "tool-regress," Dennett seems to take evolution as the "joker"; or, as Clark puts it: What *matters* about minds and persons is "a picture in which tools and skills take center stage and in which the primary 'vehicles' of content are the embodied capacities of *whole agents embedded in a cultural and ecological niche*" (p. 68, emphasis added). Thus, evolution takes care of the right use of tools and skills in the appropriate circumstances.

The group, then, starts to discuss how mind-tools could differ from the more general class of tools. Sad to say, but we have the strong impression that (analytic) philosophy of mind has seen better days . . .

> *Clark*: The can-opener, is that a mind-tool?
> *Cummins*: Well I know how to open a can because I've got a can-opener, and I know how to run the can-opener. And it *knows* [!] how to open the can.
> *Clark*: Okay, the presence of the tool makes a difference to the knowledge that you have.
> *Cummins*: Yeah, it expands my knowledge. (pp. 95–96)

Although Clapin reminds the reader in his editor's note (p. xiv) that it would not be fair to make the speakers responsible for every detail—the contributions to the discussions were uttered "off-the cuff"—we cannot help wondering why, then, phrases such as the quoted are printed at all. Do they refer to serious positions, or don't they?

Finally, let's focus on the relationship between minds and tools: How could we distinguish a tool that *aids* thought (like paper and pencil in mathematics) from a tool that actually *implements* part of the thinking? Is the concept TOOL adequately used at all in the second case? As we have learned in the short "discussion" at the beginning, no one *uses* the mind-tools. According to Dennett, the mind *is* the collection of mind-tools. However, if language is to be thought of as a tool,

who is the user? Is language a tool used *by* thought processes, or does the tool here *constitute* the thinking? Clark, thus (pp. 75 ff.), questions whether Dennett's story implies the breakdown of the *prima facie* sensible distinction between thinking *with* a tool (e.g., English) and thinking *in* a code (e.g., mentalese). Such a breakdown looks likely to be the erosion of any clear distinction between the tools and the user—the user appears as an illusion, being just a bundle of tools. The consequences of such an approach are far-reaching: As the discussion reveals, Dennett and Clark agree on the disappearance of the user and of persons as we know them (p. 99). We leave it to the reader to treat this result as an insight of recent philosophy of mind.

5. Cummins: Representations at Last

The discussion of Robert Cummins' theory of mental representation suffers less from the brevity that infects the rest of the book. In fact, the theory is summarized at least four times. Cummins uses it to frame his commentary on Haugeland, who responds by re-presenting Cummins' approach. In the chapter devoted to Cummins, Smith reviews Cummins' theory again, to which Cummins responds by again summarizing and explicating his theory. In the end one does begin to understand what Cummins is up to, and where the problems lie. There are in addition pages and pages of discussion, in which Smith, Clark, Dennett, Haugeland and Clapin all prod Cummins' in what appears to be a well-intentioned effort to understand how the theory works. Cummins is frequently forced to clarify (and sometimes obfuscate) and readers of Cummins work are occasionally treated to the "aha!" experience of having questions asked (and even answered) that one has had in the back of ones head since reading *Representations, Targets and Attitudes* (Cummins, 1996).

The problem that motivates Cummins is the problem of representational error. Being an unrepentant naive realist, Cummins thinks that the central problem for a theory of mental representation is to provide a naturalistic explanation for the fact that mental representations stand for things in the world. One simple explanation is causal: A representation stands for the thing that caused the cognitive system to generate it. This sounds reasonable. Typically when we are confronted with cows our cognitive system tokens a mental representation COW, which is why we say that COW stands for cows. But, as is well known, this account is wrong. There are plenty of cases in which we token COW when we are not confronted with cows. These problems of representational error have been discussed extensively by Fred Dretske and Jerry Fodor.

Fodor, incidentally, is clearly the missing man at the table. His work on mental representation (e.g. 1975, 1987, 1990) forms the backdrop against which much of the discussion in the book takes place. In fact, he is cited as frequently as any of the researchers who were present in Blue Hill. Fodor's very particular view on the nature of mental representation—the hypothesis that mental representations are language-like in nature—is the "classical" view against which the embedded, embodied reformers Clark and Smith (and even the later Haugeland) are reacting. And it is

Fodor's account of the problem of representational error—that which makes reference to "asymmetrical dependency"—to which Cummins wants to provide an alternative. What's more, the theory that Cummins develops is in some sense simply an explication of a statement about representational error that Cummins attributes to Fodor.

What Fodor supposedly said was this: "Error occurs when a representation is applied to something it is not true of" (p. 123). So a representation is related to two kinds of worldly objects. On the one hand a representation is related to its "target," the object that a particular representational token is intended to represent (the object that it is applied to). On the other hand, a representation is related to its semantic "content" (the object or objects that it is true of). A distant, foggy horse might be the target of a particular tokening of COW, but—not being a cow—would not be an object that COW is true of. Cummins' insight is that targets are not contents and that the theory of mental representation must be factored into two independent theories: the theory of target fixation and the theory of representational content. The fact that there is a potential gap between target and content is what makes room for representational error.

However, to maintain a gap between target and content, Cummins needs a theory of representational content that is independent of use. In a move that Smith characterizes as "a mite daft," (p. 171) Cummins proposes that representations are imbued with their content via structural isomorphism. His favorite example is that of a city map. A map is available for guiding behavior, and is rich enough to do so. We have a cognitively manipulable representation with inherent content. So far so good, at least on the intuitive level. Cummins is remarkably silent about what structures he is talking about, however. In the case of maps it is fairly clear that some properties of the representation he discusses are not relevant. It is the conventions of map interpretation that tell us that in some maps a cross-hatched line in the map corresponds to two parallel steel rails affixed to the ground, or that the property of being a certain shade of blue corresponds the property of being a body of water. More on the problem of conventions below. The idea that Cummins wants to push, however, is that it is the structure of a representation that gives it content.

Representations—as representations—aren't intentional. In fact, intentionality is the domain of target fixation, which is the result of what Cummins calls "intender mechanisms," and it is only via them that representations have intentional properties. Intender mechanisms, then, are the central intentionality bearers in Cummins' theory, so it comes as a great relief when, in discussion, Clark finally asks straight out: "What is an intender mechanism?" We don't, of course, get a direct answer. But we do learn that "there is a lot more intentionality than is involved in representation." To which Clark memorably replies, "So intentionality is out there in the trees, it's all around, for you." In the course of the discussions (pp. 145–166, 198–217) it becomes clear that Cummins is trying to identify two different notions of what might be called "misrepresentation" that may not be reconcilable. On the one hand we have representations that are like pictures in the mind, reflecting the

structure of the thing they are intended to represent. Unlike the things themselves, these objects can be manipulated by the cognitive system directly to guide behavior. In this case misrepresentation is just how much the picture is not like the thing it represents. The second intuition is that representation is about information. This is the notion of intentionality as indication. When the brake lamp goes on it means that the brake fluid is low, when the oil lamp goes on, it means . . . very bad things usually. Here we can have misrepresentation, in some sense. The fuel lamp might go on because we are going up a steep incline, despite the fact that the tank is half full.

Cummins is stuck between a rock and a hard place. He is forced, on the one hand to say that an indicator light is a kind of representation and the fluid level being too low is the target, and the application of indicator light to target yields a content with the meaning like 'Yes, the fluid level is too low.' On the other hand he wants to say that indicators aren't representations. In the end, Cummins admits that this is a continuum from indication to representation, from indicator lights to strings (which represent piano size) to slotted cards (which represent the path through a maze) to city maps. He never quite states, however, where indicators get their contents, if they have them. Dennett raises the important point that systems that generate Cummins-style representations are probably going to be built up out of many small parts that are, in fact, not representations at all, but the issue is left unresolved as the discussion moves elsewhere. Cummins and Poirier (2004) have addressed this recently, but also inconclusively.

All the participants understand very well what Cummins is getting at and wish him well on his endeavor, but are having great fun testing out the system, seeing where things are muddled and where they are clear, and pushing Cummins into making distinctions he doesn't want to make and suggesting refinements. Cummins himself takes almost a detached empirical perspective on his own theory. He frequently talks as if he, too, is waiting to find out what the results from the lab are: Are there intender mechanisms in the digestive system? Are concepts representational? Can systems have targets that are never hit? Or create targets consciously? Enquiring philosophers want to know!

One is tempted to jump on the Cummins bandwagon. Cummins' contrast seems to bring with it some of the attractive features of the familiar subject-predicate contrast and intension-extension contrasts without the headaches. The target-content distinction would appear, e.g., to provide for a straightforward account of classical semantic problems such as Donnellan's (1966) attributive-denotational contrast: The man with water in his glass is the target that the linguistic representation 'the man with the Martini' is applied to, but not true of. Unfortunately, Cummins doesn't go there. He can't go there, in fact, because for him what is traditionally thought of as linguistic representation is not representation at all: It is ''conventional communication,'' a means for encoding concepts from one speaker's mind in a form that allows them to be decoded into concepts in the hearer's mind (pp. 212–213). So if Cummins has trouble with how indicator lamps get their content, it is no surprise that he has trouble with how linguistic representations do, and opts for an alternative. This aspect of the theory is the least satisfying, and it is the only part of

the theory that Smith rejects outright: ''I don't think he's right that words don't represent'' (p. 189).

Smith also accuses Cummins of making the theoretical mistake of confusing non-identity with independence, pointing out that it is perfectly possible for representational content to be *different* from both target and use without being *independent* of them. Perhaps this is why Cummins can't see clear to avoiding the picture theory of representation. Smith suggests that overarching use—at the type level—can constrain content ''without sacrificing the possibility of substantial error'' on a particular occasion (p. 184). Perhaps ''overarching use'' can be identified with ''conventions of use'' or ''norms,'' leaving a loophole for bringing back linguistic representation, while preserving the target-content distinction. Shared conventions may be ways of turning what Cummins calls non-portable indication-like content into portable content. Cummins clearly wants to hold on to his picture theory—he has stood enough ridicule for it—but it is equally clear that if he had a reasonable story about conventional content, in which linguistic items might also be said to represent, his approach would be significantly more appealing.

6. Smith: The Ontology of Rock (and not Rocks)

The fifth part is dedicated to the philosophy of Brian Cantwell Smith, originally a trained computer scientist who worked for many years for the Xerox Palo Alto Research Center (PARC). Smith's philosophical questions are strongly driven by practical problems in computer science, in particular, those that arose while he was developing functional programming languages and reflexive architectures for computer systems. His is not mainstream philosophy of mind and mental representation, and the innovative perspective on metaphysical and ontological questions he has is probably as familiar to cognitive scientists as to philosophers.

The discussion of Smith's philosophy is probably one of the most interesting parts of Clapin (2002). His philosophy constitutes a substantial contribution to transcendental metaphysics which contrasts so strikingly with mainstream methodological naturalism. Smith develops an ontological theory of object generation from an ineffable metaphysical flux by an interactive dynamic relation between region-like entities (his ''s-region'' and ''o-region'') that can be seen as a radical new approach to the ontological foundation of cognitive science. Although Dennett makes quite clear from the very beginning that ''*pure* metaphysics is a played-out game'' (p. 223), he seems to accept Smith's ontological move towards the reengineering (or retooling) of physical objects because of certain fundamental ontological problems in practical applications in computer science: A major motivation for his ontological turn (cf. Smith, 1996) was based on the question of how the reference of symbols, entities, processes, programs, etc., in real-world computer systems and real-world computer architectures can be established. How can reference be understood by taking the ontological basis of symbols occurring in a computer system into account?

Smith identifies as problematic the fact that the classical distinction between numbers and numerals is not sufficient for writing computer programs. His LISP dialect ''2-Lisp,'' for example, was designed to force the programmer to distinguish recursively at least four categories (besides a type-instance distinction): First, numbers which are abstract entities not occurring in the machine itself; second, numerals which refer to numbers; third, notations which are internal entities of the machine notating numerals; and fourth, handles which designate the internal structure of numerals or notations (cf. 1996, p. 39). Although Smith might be right in making this point, it is less clear whether confusing reference of symbols occurring on a screen (or internally in a machine) with numerals that are referred to by these symbols is but a simplification that does not generate any harm in practical applications. Although his point may be correct and remarkable in its own right, it is not clearly different from a situation where we use finite ordinal numbers instead of finite cardinal numbers to measure the cardinality of finite sets. We may apply an erroneous concept, but it simply does not matter.

Motivations aside, what is clear is that Smith is mainly examining the foundations of the individuation of objects, of objectifying the world, and of the detection of structure(s) in the world. He is not primarily concerned with examining the problem of what mental representations are. A natural question is: Is it the case that the ontological considerations of Smith's account have any interesting consequences at all for a theory of mental representation? Dennett hints at how Smith's ontological turn relates to a theory of mental representation. Here are some important points Dennett mentions:

1. The real world is not *per se* divided into objects and relations (i.e., the world does not exist ''pre-parsed'' for the agent acting in this world) rather the subjects themselves acting in the real world are individuating objects, relations, and the like.
2. By ignoring the fact that subjects are the individuating actors we end up with so-called ''pre-emptive registration'', namely an inscription of ontological assumptions on a system by an observer who then extracts these assumptions from the system as an empirical discovery (p. 224).
3. Without subjects objectifying the world, the world is a world of ''ineffable flux'', lacking stable objects, a world of rock not a world of rocks, a world much more like tundra than a tropic rain forest.
4. Obviously, we are at a dead end if we want to know more about this world of flux, because every feature or property we are using to describe this world or apply to this world confronts us with pre-emptive registration errors which we want to avoid.
5. As a consequence, Smith needs to develop a theory of objectification mainly driven by the subject—in other words a theory about the ''origin of objects''—in order to provide an account how conceptualizations are possible without pre-emptive registration.

Smith points out a complicated picture of a local interaction between his so-called s- and o-regions that can constitute ''registration,'' a concept used by him for

something similar to conceptualization (1996, p. 191). This interaction contains actions like tracking, separation, and stabilization of o-regions and therefore establishes an intentional aspect. Smith is not very clear about these actions, but an intuitive interpretation would be that tracking and separation are used to focus on the o-region nevertheless allowing dynamic changes of the connection.

Dennett adds some issues that are, in his opinion, lacking in Smith's philosophical project. The most important one concerns the missing evolutionary aspect of Smith's ontological considerations (p. 230). Dennett claims that Smith's account would be plainly impossible without focusing on (or at least taking into account of) evolutionary aspects. Furthermore, evolution would provide a solution for explaining objectification without assuming pre-emptive registration in the sense that the world is assumed to be pre-parsed. Smith's answer to Dennett's challenge is a simultaneous "Clearly Yes!" and "Definitely No!"—a Solomonian assessment: "Yes," because he accepts that evolution is the means by which we learned to register and "No," because he does not believe that registering is itself an evolutionary process (p. 261). Smith is pluralistic enough to accept such an explanation how we can register the world but he cannot accept the explanatory reduction of registration to a form of evolutionary process.

Smith's theory of objectification (i.e., his theory of registration) provides a basis where we can find a clue for a connection to the problem field of intentionality and mental representation. In order to "establish an object," a form of conceptual directedness towards the world is necessary. While classical theories presuppose given classifications or ontologies, the world, according to Smith, rather presents itself to the observer as a messy collection of concrete as well as abstract copies and versions of one and the same thing, very similar to a document that can occur to the librarian as a template of the document, a reproduction of the document, a certain edition of the document and so on (pp. 252–253). And it is precisely here that the non-dynamic character of this simple type-token distinction comes in: There cannot be a simple given classification of the real world.

Smith clearly does not have a theory of mental representation. What he would need is a theory that is relative in a certain sense, namely, that includes a flexible type-token dynamics, that allows for contextual relativity and contextual changes, that allows for perspective shifts, that models intentions, attitudes, and the directedness of mental processes etc.—without using concepts like types and tokens, contexts, intentions, attitudes and the like, because they are not available in this account. In addition, he would need an approach that makes objectifications concrete and plausible and gives an idea how object generation can be introduced without making the pre-emptive registration error. Faced with these challenges, Smith's project becomes—although extremely interesting—more and more questionable.

7. Conclusion

The Philosophy of Mental Representation is something like the Bible, in the sense that it is highly compressed, but sometimes terribly longwinded, with the same story

being told in a number of different places from different perspectives. It is hard to read through as such, since there is no argument being made, but easy to come back to over and over. Each author has attempted to extract the pearls of wisdom that, from his perspective, another author has unearthed. And the discursive discussions, which at first seem entirely haphazard and informal, contain gems—and not just the humorous ones we've cited here.

Note

[1] Quotations from Clapin's book are just by page numbers; all other citations follow the standard citation format.

References

Clapin, H. (Ed.) (2002). *Philosophy of mental representation*. Oxford, England: Oxford University Press.

Cummins, R. (1996). *Representations, targets and attitudes*. Cambridge, MA: MIT Press.

Cummins, R., & Poirier, P. (2004). Representation and indication. In H. Clapin, P. Staines, & P. Slezak (Eds.), *Representation in mind: New approaches to mental representation* (pp. 21–40). Oxford, England: Elsevier.

Dennett, D. C. (1987). Styles of mental representation. In *The intentional stance* (pp. 213–225). Cambridge, MA: MIT Press. (Original work published 1983)

Dennett, D. C. (2000). Making tools for thinking. In D. Sperber (Ed.) *Metarepresentations: A multidisciplinary perspective* (pp. 17–29). Oxford, England: Oxford University Press.

Donnellan, K. S. (1966). Reference and definite descriptions. *Philosophical Review*, *75*, 281–304.

Fodor, J. A. (1975). *The language of thought*. New York: T.Y. Crowell.

Fodor, J. A. (1987). *Psychosemantics. The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Fodor, J. A. (1990). *A theory of content and other essays*. Cambridge, MA: MIT Press.

Haugeland, J. (1998). *Having thought: Essays in the metaphysics of mind*. Cambridge, MA: Harvard University Press.

Haugeland, J. (2000). Truth and finitude: Heidegger's transcendental existentialism. In M. Wrathall & J. Malpas (Eds.), *Heidegger, authenticity, and modernity: Essays in honor of Hubert L. Dreyfus* (Vol. 1, pp. 43–77). Cambridge, MA: MIT Press.

Smith, B. C. (1996). *On the origin of objects*. Cambridge, MA: MIT Press.