

"Minds, Brains and Tools" (with a response by Daniel Dennett,) in Hugh Clapin (ed) PHILOSOPHY OF MENTAL REPRESENTATION (Clarendon Press, Oxford, 2002)

Minds, Brains and Tools

(Comments on Dennett for Hugh Clapin's "Workshop on mental representation",
Maine, August 1999)

Andy Clark

*Philosophy/Neuroscience/Psychology Program
Department of Philosophy
Washington University
St. Louis, MO 63130*

e-mail: andy@twinearth.wustl.edu

The texts were:

D. Dennett "Things About Things"
(draft for Lisbon conference on Cognitive Science, May 1998) (TAT)

D. Dennett "Making Tools for Thinking"
(for D. Sperber (ed) volume on *Metarepresentation*) (MTT)

D. Dennett "Styles of Mental Representation"
In his *Intentional Stance* (MIT, 1987) 213-236 (SMR)

1. Reading Dennett

The selected texts for this discussion were two recent pieces by Dennett (“Things About Things” and “Making Tools for Thinking” – henceforth TAT and MTT respectively) and one oldie-but-goodie (“Styles of Mental Representation”, henceforth SMR). What was most striking, to me, was the way these three small texts seemed to fit together and, collectively, to greatly illuminate the shape of Dennett’s whole corpus, from the work on stances all the way to the work on consciousness and personhood. I found this kind of exhilarating, and I hope my enthusiasm – tempered though it is by a few doubts and worries – shows through: I had a ball.

Here’s how I plan to proceed. I’ll start by presenting what I take to be Dennett’s position. I will do this, however, in the most blatant, outrageous and caveat-free way possible, as I want to home in fast on some key issues (I am also keen to see if Dennett can endorse these deliberately provocative formulations). I then raise a somewhat obscure question, but one that is, I think, ultimately important and revealing: the question is “Could it really be mind-tools *all the way down?*”. Having tried to answer in the affirmative, I turn to the issues concerning aboutness and internal representations, and explore the relations between Dennett’s emphasis on skills and tools and his rejection of certain ideas about contents and vehicles. In the final section I offer a tentative illustration of the big picture, using recent work on numerical cognition as a concrete case-study. There is a brief conclusion rehearsing four outstanding questions for the tool-based vision of mind.

2. What Minds are Made of.

One way to say what I think Dennett (TAT, MTT, SMR) is up to is by a simple triplet of questions and answers. Showing what these questions and answers *mean* is the real goal of this section. The triplet goes like this:

Q/ What are minds made of?

A/ Tools for thinking.

Q/ Who or what *uses* the tools to *do* the thinking?

A/ No-one, nothing. The tools-are-use.

Q/ Intentionality, aboutness, content and consciousness: can all these really be brought into being by grab-bags of useless tools?

A/ Yes.

The first point to notice, given the topic of the present workshop, is the constant emphasis on tools rather than on internal representation. Traditional (“intellectualist” – see SMR) cognitive science had it that minds were made of internal representations. Dennett, as far as I can see, simply denies this outright. Anything in there (in the head) or out there (in the world) worth calling a representation is not, according to Dennett (SMR) fundamental. What is fundamental are the skills bequeathed by the tools that build know-how tacitly into the system. Dennett, in SMR (and inspired by Ryle (1949), with perhaps a whiff of Wittgenstein), offers a neat little argument designed to show why this *must* be so.

(The Rylean Argument)

First, let’s distinguish three ways on which information may be incarnate in a system. The information may be explicit, tokened as syntactically distinct, reliably interpretable symbols (SMR p. 215). Second, it may be implicit: logically (actually, I suspect “logically” may be unnecessarily strong here: why not just “derivable”?) implied by whatever is stored explicitly (SMR p. 216). Third, it may be tacit: “built into the system in some fashion that does not require it to be represented (explicitly) in the system” (SMR p. 218). The pocket calculator, it is said (SMR pp. 221-222), represents number explicitly (on screen and in buffers). But arithmetical truths and axioms are nowhere tokened in the machine, nor are they logically implied by what *is* tokened (sequences of numbers). So such truths and axioms are at best tacitly ‘represented’ – the rules are incarnate in the hardware in much the way the laws of hydrodynamics are reflected in the bodily form of a fish.

Now let’s ask in virtue of what any explicit representations have the contents they do? Imagine a pattern of electrical activity, in the pocket calculator, which explicitly represents the number 7 in binary code (SMR p. 221). Now keep the token of the number 7 intact, but re-write the hardware so that the device embodies no tacit knowledge of the rules of arithmetic. Does it still carry the content we previously identified? There is doubtless some room for maneuver here, if we import additional considerations concerning the history and design of a device. But

the basic intuition to be pumped is the negative one. To quote Dennett, “Explicit representations, by themselves [are] quite inert as information bearers...They *become* information bearers only when given roles in larger systems” (SMR p. 217). And how do you get an appropriate role in a larger system? Only, it seems, if the system is set up so as to actually *do* things – to move and act in the world (in our case) or at the very least to go from one representation to another (as in the case of the pocket calculator). And what *this* requires, on pain of the famous infinite regress of rules and recipes, is to be part of a system that has some *tacit* knowledge (“know-how”) which can put the representations to work without requiring more layer so representation to do so (SMR p. 218). Conclusion: tacit knowledge (non-represented skills and know-how) is more fundamental than inner tokens and internal representations.

Now the need for some such underlying structure of tacit knowledge is, I thin, accepted on all sides. Fodor, in his (1987) *Psychosemantics* concedes that “a computer in which the principles of operation are *only* explicitly represented is just like a blackboard on which the principles are written down...when you turn the thing on, nothing happens” (Fodor 1987, p. 23). What is interesting and powerful about Dennett’s treatment is not the insight itself – which is familiar enough – but the way Dennett puts it to work as the cornerstone of a different way of looking at intelligence and meaning. What we find in the two more recent treatments (TAT and MTT) is, precisely, a way of following through on the Rylean argument which fixes genuine points of disagreement with “intellectualist” cognitive science and which paints a very different picture of what *matters* about minds and persons: 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. Dennett thus ends up rather startlingly close to a Heideggerian vision of the nature of cognition, though this is not a theme I am competent to explore.

The key to Dennett’s alternative lies in the first of our opening question and answer pairs: What are minds made of? Answer – tools for thinking (“minds are composed of tools for thing”

MTT, p.4). This is a theme that is increasingly prominent in Dennett's work (see especially the closing chapters of Dennett (1996) and chapters 12 and 13 of Dennett (1995). It is best unpacked, I want to suggest, in terms of three overlapping but independent sub-themes:

1. External tools augment and transform biological cognition
2. External symbols somehow pave the way for 'florid representing'
3. Florid representing is what distinguishes mere 'believers' from genuine understanders

I do think that Dennett, in MTT and elsewhere (Dennett (1996)), is effectively arguing for (3). But although there are plenty of sentences in Dennett cognate with (1) and (2), there is nothing that affects (3) quite so boldly and blatantly. Let's creep up on it then, by way of the less provocative sub-themes.

The idea that external items can augment and scaffold both behavior and thought is pretty evidently true, and arguably of deep importance. The sailor armed with hooey and alidade can achieve feats of navigation that would baffle the naked brain, as Hutchins (1995) exhaustively documents. And - perhaps more importantly for this discussion - the way such tools work is by affording the kinds of inner reasoning and outer manipulation that *fit* our brains, bodies and evolutionary heritage. Our visual acuity and pattern-matching skills, for example, far outweigh our capacities to perform sequences of complex arithmetical operations. The slide rule is a tool which transforms the latter (intractable) kind of task into a more homely one of visual cognition. Tools can thus reduce intractable kinds of problems to ones we already know how to solve.

A big question about tools, of course, is how did they get here? If tools are tricks for pressing increased functionality out of biologically basic strategies, what kinds of minds can make the tools that make new kinds of minds? This is an issue that Dennett touches on in MTT in his discussion of "found objects". The idea here, I think, is to reveal certain kinds of tool-discovery-procedure as falling entirely under the umbrella of blind trial and error learning. Köhler's chimps, in initially playing around with sticks and boxes, were not (MTT pp. 5-6) trying to solve the food acquisition problem. Instead they were just "familiarizing themselves with objects in their environments". In this process, affordances may well be spotted (the stick

knocks off a banana and so on). One of Dennett's major themes is thus that the initial path and successful tool use need not involve a process of design in which thoughts about the tool/problem fit guide a search for good tools. Instead, recognition of tool/problem fit, if it comes at all, may well come after the event of successful use.

In fact, we can go further and observe that not just the discovery of tools, but also their evolution, improvement and refinement, can proceed with little or no deliberate design activity on our part. The trick here - as laid out by Dennett (1997) and by Deacon (1997, ch. 4) - is to recognize the tools as themselves a class of replicating entities whose "success" (widespread replication) or "failure" (extinction) depends on the extent to which we adopt them. Tools in general (and language in particular, more on which later) may thus be seen (Deacon (1997) p. 112) as rather like viruses - incapable of reproducing on their own, dependent on a host's metabolic and reproductive systems, yet susceptible to processes of variation and differential reproduction sculpted by their success or failure at invading host organisms. By seeing tools as entities with their own selective histories, we make room for what Deacon calls "a flurry of adaptation.... going on outside the brain" (op cit p. 109).

In focusing our attention on the role of external tools, Dennett is thus concerned to avoid what I have elsewhere termed "the paradox of active stupidity". This would be the idea that making the moves that sculpt the environment so as to allow *cheap* problem solving itself requires *expensive*, advanced, design-oriented cogitation. The nasty upshot being that only *clever* brains could make their worlds smart so that they could be dumb in peace - a result which would deprive the tool-based scenario of its appealing role in explaining the *origins* of advanced, reflective thought and reason. The observations about found objects, trial and error manipulation, tools as replicators, along with the clear potential for tool-based bootstrapping (using the slide rule to design a digital computer and so on), are meant as a response to this natural worry.

The most potent bootstrapping resource - the tool of tools - is surely language itself. And Dennett's long-term corpus is littered (that sounds bad - I mean "liberally sprinkled") with

suggestions and speculations about the way "words do things with us" (Dennett (1991) ch. 8). In the papers under discussion, what words are said to do for us is to open up the space of thinking about thinking, and to thus enable what Dennett (MTT) calls "florid representing". Representing is florid, we are told, when it is "deliberate", "knowing", "self-conscious" - when you either do, or at any rate could, appreciate that you are manipulating object that represent. This kind of florid representing is contrasted, by Dennett, with the sense in which a state of the visual system may "represent" the presence of a food item, and even with the sense in which a young child or chimpanzee may use representational tools (symbols) as part of a cause-effect communication system (Gauker (1990)) to achieve a goal. It is, in short, one thing to learn that using the word "ice cream" or touching a symbol for banana will help assuage hunger, and quite another to appreciate the word or symbol *as* having a representational role: as being an object that is about other objects, as "thing about things" (Dennett TAT). Notice, finally, that according to Dennett an agent might even use meta-representations without genuinely appreciating the notion of a representational role. The chimp or child may represent the caregiver as believing that the food is in the red box (might "have a theory of mind") yet not exhibit "the *knowing* competence to *use representations*" (MTT p. 5, original emphasis).

I was not quite sure now best to understand this last idea, which nonetheless seemed quite central to the project of MTT, so let me pause to review the claim. What is at issue, it seems, is the "florid" or "witting" (MTT p. 5) use of representations. This is tied to the notion of "thinking about thinking" - a notion I stress in some of the work Dennett cites. The point there is that once a thought or argument is rendered as an external symbolic object (a string of typed words or repeatable sounds) it is itself available as an object for further thought or scrutiny (and see Jackendoff (1996)). Dennett's idea, I think, is that the external symbolic objects thus allow us to *unwittingly* (at first) treat a token of thought as an object for further thought. If we then realize that that is what we did, we have stumbled into the ranks of the *florid* representers - beings who are aware of the power of using things to represent things, and can self-consciously exploit this

power in cultural practices, the delineations of explicit norms, and so on. At this point, one may say, the sky (hook) is the limit.

Dennett's suspicion is that witting representational practice depends - contingently but crucially - on the prior unwitting use of external objects as representations, and that this, in turn, is just a special instance of the use of found objects (including, perhaps, involuntary cries or facial gestures - see Dennett (1991) ch. 7) as tools.

Which brings us to the third and final sub-theme, the real point of MTT, and the motivation behind a lot of Dennett's recent work. Or so I claim. This is the idea (which I admit Dennett does not state in quite so many words) that florid representing distinguishes mere "believers" from genuine thinkers and understanders. This may come as a shock to any who think of Dennett as a mad-dog instrumentalist who places Einstein, the lectern, and the thermostat on a simple continuum of intentional-stance-worthiness. A case could be made that such a depiction was always demonstrably mistaken, but I shan't pause for that. Instead, just savor a few remarks culled from the more recent corpus:

"I am tempted to say that even if they [chimpanzees] do have beliefs about beliefs, they may well be incapable of *thinking about thinking*. They may, indeed, not really be capable of thinking at all (in some florid but important sense of "thinking").

MTT p. 4 (original emphasis)

"thinking - our kind of thinking - had to wait for talking to emerge"

Dennett (1996) p. 130

"In order to be conscious - in order to be the sort of thing it is like something to be - it is necessary to have a certain sort of informational organization...[one] that is swiftly achieved in one species, our, and in no other...My claim is not that other species lack our kind of *self-consciousness*...I am claiming that what must be added to mere responsivity, mere discrimination, to count as consciousness *at all* is an organization that is not ubiquitous among sentient organisms"

Dennett (1998) p. 347

"our kind of consciousness is...in surprisingly large measure, an artifact of our immersion in human culture"

Dennett (1998) p. 346

Florid representing, consciousness, and thinking about thinking are thus tied together, with the whole bundle depicted as historically dependent on the emergence of a special kind of mind-tool: speech or other forms of linguistic object. What such objects do is position us to acquire the *idea* of representation, thus priming the cultural explosion of mind-tools (notations, slide-rules, laws, norms, advice, education) that sculpt plastic neural circuits and co-constitute human intelligence. Such, at any rate, is my gloss on suggestive, but occasionally cryptic, assertions such as this:

"It is because the lions can't talk to each other that they also can't come to use tokens of other sorts to represent, non-linguistically. And it is by that indirect route, I think, that we come to construct our minds"

MTT p. 7

Whatever instrumentalist leanings Dennett has or may have had, it thus seems clear that he *also* insists on genuine and important organizational differences. Differences so important, indeed, as to render certain kinds of adaptively potent organization (in the chimps and lions for example) unable to support 'real thinking' (in the florid sense) or real consciousness. Since chimps and lions remain rather wonderful candidates for the intentional stance, this organization dividing line must be drawn between grades of *believer*: some believers (the florid representers) are real thinkers while others are not. Seeing just how (and indeed if) this realism about floridity can sit beside the instrumentation about belief would be a nice project, though not one I will pursue (though see note 3 later on). Instead, I want to focus on the positive story, and try to clarify and explore these claims about tools, language and minds. Time, then, to turn up the focus.

3. Could It Be Tools *All the Way Down*

A good place to start is with the second opening question: Who or what uses the tools to do the thinking? The answer here, which I suspect is deeper and more difficult than it initially appears, is no-one. No-one *uses* the mind-tools. The mind *is* the collection of mind-tools (Mind-tools-R-Us). It is at this point that the issues about tools also phase into Dennett's perennial concerns about persons and about consciousness. So there are lots of issues here, and I shan't attempt to get to grips with them all. But here's one which gets pretty quickly to the heart of things.

Consider language, Dennett's "tool of tools". One question that can be asked is this: if language is to be thought of as a tool, what becomes of the user? Relatedly, is language a tool used *by* thought processes, or does the tool *constitute* the thinking? Like Dennett, I am deeply committed to some kind of "cognitive involvement" (Carruthers (1996)) of public symbolic codes in thought – that is to say, I don't see the public codes as merely vehicles for the communication and external encoding of thought, but as active contributors to the processes of thinking. But I think it is useful to distinguish two (at least two) possibilities concerning the *kind* of involvement at issue.

To focus this, take a simple case reported by Henser (1999). Japanese/English bilingual speakers will sometimes use a phrase such as this: "I just wanted to say that I feel really *moshiwakemai*¹ about it". This is an instance of what Henser calls "code-switching": suddenly jumping from one linguistic system to another. A common example of code-switching is when counting. A speaker who is absolutely fluent in English may very well switch back to their native tongue for sotto voce calculation, and even bi-linguals tend to have a preferred language for counting. Why switch codes? It is not that the ideas cannot be expressed in the other language. But it may well be that certain cognitive routines are easier to organize, or better practiced, using internal representations of specific public symbols. Henser thus introduces the notion of a "lingpack" – a set of mental items (perhaps in mentalese) which hang together and are packaged as a manipulable item for use in propositional thought (op cit p. 28). Different languages select different lingpacks, and a lingpack need not be "unzipped" (exploded) to be

used. ‘Moshiwakemai’ is a lingpack in Japanese, but not in English, hence the cognitive economy afforded by code-switching.

Such a story is tempting, but it looks to remain uncommitted on an arguably crucial issue. It does not distinguish between:

1. The idea that we (sometimes) *think in* (say) English
- and
2. The idea that we only *think* in some special inner code or codes (say ‘mentalese’) but that while we do so we also (sometimes) use inner representations of actual words in some specific language (e.g. “moshiwakemai”) to help organize, focus and recall ideas and sequences of ideas.

On the face of it, this is a real and perhaps important distinction. Consider an analogy. I often use pen and paper to help me organize, focus and recall ideas. But it does not seem to follow that I ‘think in’ pen and paper...though (on the third hand) it *is* true that the pen and paper form part of an extended cognitive and problem-solving system with my biological brain. This is, perhaps, the distinction between ‘thinking in’ and ‘thinking with’... In the case of the “moshiwakemai” code-switching example, the evidence seems compatible with either option. What the results suggest is that the specific ‘lingpacks’ available in different languages make some kinds of thinking easier or harder, and hence that code switching can be indicated. But this could equally well be because the subjects, although doing all the *thinking* in mentalese, need internal representations of specific lingpacks to focus, hold and organize the thoughts.

In fact, I have tended to favor this latter option (see for example Clark (1998)). I also think it fits nicely with results concerning the effects of having different number words on arithmetical performance. For example, one Chinese (Cantonese) dialect has very brief words for numbers and speakers show a random digit recall span of 10 as against our ‘magic no 7’ – for a lovely account see Dehaene (1997) pp. 102-103. Code switching would thus be indicated for that task for a bilingual Cantonese/English speaker, even if all that is involved is an internal representation of the phonetics of the words for the digits as a way of aiding recall.

I have begun to wonder, however, whether Dennett might perhaps be covertly committed to denying the distinction between “thinking with” and “thinking in”. For suppose someone (like Dennett) hold that content and aboutness get into the picture in virtue of skilled engagement (or the potential for skilled engagement) between agent and world. It could then very well be the case that no *single* mind-tool can support the kinds of flexible, skilled engagement characteristic of what we call “understanding”, “grasp of meaning” and so on. To the extent that thought and understanding thus depend on the activity of multiple *non-privileged* mind-tools, there is no *more* reason to treat internal representations of public language words as “merely derivatively contentful” than there is to thus treat a symbol in mentalese, or a mental image, or any *other* aspect of any one of our myriad mind-tools as only a second-grade content-bearer. Instead, they are all on a par, and *none* of them have the kind of intrinsic ‘aboutness’ that is sometimes posited as a kind of cognitive scientific grail. Strings of words, we might be tempted to say, cannot in themselves constitute a thought. But neither, on this picture, can anything else. Thought is an intrinsically more holistic phenomena, dependent always and everywhere on the action of multiple mind-tools, not all of which *can* (as we saw in section 2) consist in inner tokenings of anything worth calling a representation. Here, then, is a point to ponder: does Dennett’s story imply the breakdown of the distinction between “thinking *with* a tool (e.g. English)” and “thinking *in* a code” (e.g. Mentalese)?

Let’s suppose for a moment that it does. One cost of such a breakdown looks likely to be the consequential lack of any clear distinction between the tools and the user. A user just is a bundle of tools, and no tool is *privileged* – no tool *constitutes* the user in a way other tools do not. Instead, a loose coalition of tools (or “skill-supporting components”) together support the range of flexible engagements and responses characteristic of intelligence and thought.

It is not surprising, from this perspective, to find Dennett unworried by the idea of the environmentally extended mind (see also Clark and Chalmers (1997)) – witness the comments towards the end of MTT (pp. 8-9) concerning the way the distribution of tasks across biological brain and local environment makes “our minds so much more powerful than all other animal

minds”. The idea here seems to be that language uniquely positions us to create a cascade of new mind-tools which literally transform us into more powerful (but extended) cognitive engines – an idea also familiar from Dennettian accounts of the origin of consciousness.

We thus tiptoe into the metaphysically challenging terrain explored by Beth Preston in her work on tool use and cognition. Preston (1998) defends a broad notion of tool use based on Heidegger’s notion of equipment – a notion which avoids the (I think arbitrary – see Clark and Chalmers (op cit)) restriction of tools to items external to the biological organism, and opts for a function-based account in which bodily parts (e.g. hands) and biological cognitive elements (e.g. biological memory) end up on a par with rakes and shopping lists. Dennett’s story, I believe, is similarly (and properly) liberal, and ultimately confronts the same problem – that:

The user has in some sense disappeared in a welter of equipment....if all the bodily parts of an organism, including its mental states...are equipment, you have a situation where you peel away layers of equipment as you would peel away the layers of an onion, ending up with nothing at all in the way of a central core.”

Preston (1998) p. 545

This strikes me as a perfectly acceptable place to end up, though intuitions clearly differ on this. It is, at any rate, hard to see how to *avoid* this without embracing some *equally* problematic story – e.g. anointing some aspect of the inner goings-on as the “real cognitive engine” with all the rest relegated to “mere” support and data-storage (which leads to a kind of unwelcome cognitive *shrinkage*²), or simply *insisting* that the biological agent is the cognitive agent and factoring out the sources of behavioral complexity accordingly.

None of this, however, forces us to give up on the morally and socially crucial notion of *persons* and thinking *agents*. One potential reconstruction might begin (Siegal (ms)) with the phenomenological facts of first person experience. A tool/user divide might then be motivated by facts about how things *seem* to an agent – the pen, the automobile and the blind man’s cane do not typically strike the agent as simply parts of the environment, but as integrated probes and equipment through which they think and act. New tools and ill-matched tools do not lend

themselves to this “knitting in”. The user, on this view, is not any old bag of tools but whatever bag of tools functions, at a given moment, as transparent equipment for thought and action.

Another (more Dennettian) reconstruction might highlight the set of tools which support the so-called “user-illusion” (Dennett (1991, ch. 7,10)(1998b, ch. 24)). It is here that the issues about tools and users phase directly into the ones about consciousness and content, and it may be worth pausing to review the story. Here’s a super-quick, but I hope not *too* misleading sketch.

First move: The intentional stance. The idea here is that a system has a belief just in case its behavior is well-predicted by treating it as a believer. This is, as Dennett (1998, p.331) notes, a “maximally permissive understanding,” which makes no specific claims about inner structure or organization.

Second move: Multiple Drafts. Based on a variety of neuroscientific and cognitive psychological findings, Dennett (see also, Dennett and Kinsbourne 1992) depicts the biological brain as the locus of multiple, quasi-independent processing streams. There is no single, ultimate judgement issued by the brain in response to an input – no decisive moment in space or time where the system settles on a unique definitive content fixing the conscious state. Contrast this with a traditional model in which ‘central processing’ names an area where, in Dennett’s recurrent phrase, “it all comes together”, and a judgement is made whose content fixes how things seem to the conscious subject.

Third Move: The Narrative Twist. So whence the conscious experience of seeing such-and-such *as* so-and-so, of feeling pain *as* a sharp stabbing in the arm, etc.? This kind of content-fixation, Dennett suggests, is probably a peculiar achievement of *human* biological brains – an achievement made possible not by the presence of some special biologically-evolved circuitry so much as by the cultural imprinting of a kind of “user-illusion”. “*Our* kind of consciousness”, recall (Dennett 1998, p. 346) “is not anything we are born with, not part of our innate hard-wiring, but in surprisingly large measure, an artifact of our immersion in human culture.” Our extraordinary immersion in a sea of culture and language (itself, to be sure, made possible by *some* difference in innate hardware) creates, in the human brain, a new kind of cognitive

organization – a new “virtual machine” – which allows us to make cognitive objects of our own thought processes and to weave a kind of ongoing narrative (about who we are, and what we are doing, and why we are doing it) that artificially “fixes” the cognitive contents. The content is, of course, not *really* fixed (see TAT), because underneath the personal-level narrative stream the more fundamental multiple processing streams are still going like the clappers. But there is, courtesy of the new top-level virtual organization, a striking difference: we now *report* the presence of a specific stream of experiences, a stream, if you will, of *judgings* or *macro-takings*, in which there seems to be a clear fact of the matter concerning the nature of our current subjective state. It is the presence of this serial stream of apparently fixed contents that explains, on Dennett’s account, our tendency to believe in *qualia*. But what these qualia really are now turns out to be nothing but the string of judgements made by the top-level, linguistically-infected, narrative-spinning virtual machine: a machine installed not by nature, but by the almost-incalculable effects, in reasonably plastic human brains, of our early immersion in a sea of words and culture, or more generally by our immersion in a sea of external symbolic items and self-reflective cultural practices.

The upshot is that *believing* is pervasive and fundamental. But human-style conscious awareness requires an extra layer of judgement rooted in a culturally inculcated capacity to spin a privileged report or narrative: “the story you or I will tell if asked (to put a complicated matter crudely)”, Dennett (1998) p. 348. Consciousness *achieved*, not given. Notice then that much of the burden is thus shifted from the notion of consciousness to the notion of personhood. For it is personhood which (via the ongoing narrative) now emerges as the primary culture-driven achievement, and the one most clearly linked to the activity of the linguistic mind-tools we appropriate from our symbol-rich environment.

The point, at any rate, is that there are several ways in which the notions of *persons* and *agency* may be reconstructed despite the image of the user as *nothing but* a bag of (embodied, embedded) mind-tools. Even if no tool or tools are intrinsically privileged, only certain

combinations of tools will yield the user-illusion that, for Dennett, distils consciousness from the flux of adaptive response.³

4.

Internal Representations, Vehicles and Skills.

Dennett's views on internal representation and vehicles of content are best appreciated, I want to suggest, against this backdrop of ideas about mind-tools. In this section I want to sketch these ideas (drawing heavily on TAT), fill in the connections (using the ideas developed in sections 2 and 3) and raise a few questions.

In TAT, Dennett suggest setting aside two idealizing assumptions often made in the study of intelligent systems. The first concerns "how to capture content" and the second "how to isolate the vehicles of content from the "outside" world" (TAT, p. 1). Concerning content capture, Dennett's main point is that we should not assume that wherever there is a genuinely *contentful* state, there is an accurate and exhaustive propositional description of the content of that state. Dennett gives the example of a "piece of cognitive machinery" whose operation does indeed give someone a "thing" about redheads, subtly adjusting their responses to all situations in which redheads are suspected of playing a role. Yet this "thing about redheads", undeniably (?) contentful as it is, and despite its having a perfectly good physical vehicle "in the head", does not seem amenable to accurate and exhaustive propositional specification.

Dennett's suspicion is, I think, that *most* of the inner mechanisms supporting cognitive contents are like that: they are components and circuitry which display *non-propositional aboutness*. One of the interesting things about Rodney Brooks' work, according to Dennett, is thus that in the Cog project, it is "pushing these profoundly non-propositional models of contentful structures into territory that is recognizable as human psychology" (TAT, p. 2). Dennett also suggests, intriguingly, that treating these non-propositional contents as something like *implicit* beliefs can be misleading, as it again invites use to propositionalize and "linguify" contents fundamentally ill-suited to sentential capture (TAT, p.2).

Dennett's emphasis on non-propositional kinds of content is a very natural accompaniment to his stress on skills (section 2 above) and on mind-tools (section 3). Skilled engagement between agent and world is, for Dennett, the root of all content, and this kind of skill-based content is, we saw, explanatorily prior to the kinds of content typically associated with "internal

representations”. It is associated instead with a variety of mind-tools which are, in effect, pieces of bodily, neural or environmental structure tuned and selected so as to support certain kinds of skilled engagement with the world.⁴ Propositional format representations (which may or may not exist inside the head as well as out in the world) are just one such mind-tool, among many, and should not be privileged as the sole genuine bearers of content.

But just how does the emphasis on skilled engagement comport with the emphasis (in MTT and elsewhere) on “florid representing” as (roughly) the mark of “real” understanding? At first sight, there seems to be a tension between these two claims. But in fact, seeing just how they fit together is, I suspect, crucial to understanding Dennett’s real picture.

Florid representing occurs, recall, when there is a *knowing* use of *representations* (see MTT, p. 2), where representations are (at least) some kind of manipulada: objects that bear contents and that can be, let’s say, shuffled, re-organized and recombined in ways sensitive to, and exploitative of, those contents (see MTT, p. 9). The claims then fit together like this. The objects (the manipulada) involved in florid representing bear the contents they do only in virtue of a bedrock of skills and capacities, rooted in multiple non-propositional mind-tools. But florid representing depends on making those skill-based contents into *objects* suitable for the exercise of *other* (non-propositional) skills – skills of combining, shuffling and so on. And it is this “objectification” of certain aspects of content that supports the highly versatile and open-ended range of thought characteristic of (and perhaps uniquely characteristic of) human understanding. Finally, relating all this to other themes in MTT, it is our experiences with public symbols that are said to teach us to make more manipulable objects of our thoughts and ideas. In a very real sense then – and returning briefly to the themes of section 3 – neither type of vehicle of content (the mechanisms supporting skilled response or the manipulable objects which behave more like classical *representations*) is privileged in supporting “real understanding”. Instead, real understanding emerges from the interactions between these various kinds of mind-tool.

Dennett’s take on internal representation is thus that talk of internal *representations* (rather than talk of internal content-supporting mechanisms) become increasingly appropriate as the

inner items become more object-like and more manipulable. In earlier work, Dennett stressed the lack of any “clear dividing line” between devices that really have internal representations and those that don’t (see e.g. Dennett (1987) p. 32). But as the versatility and manipulability of inner (or outer) items diminishes, so too does the usefulness of treating them *as* internal representations (MTT, p. 9).

I was not entirely sure, from Dennett’s most recent treatments, just how big a role he is giving to the *knowing* capacity to manipulate representations. The notion of “witting” and “knowing” uses of representations looms large in the discussions of florid representing. Yet it is hard to imagine an argument that would show witting use as an essential aspect of anything worth calling internal representation. In Clark and Grush (in press), for example, we argue that a robust notion of internal representation requires, indeed, the presence of *something like* Dennettian manipulada, and suggest that the acid test for the presence of such manipulada is whether the organism can use the items to guide appropriate behavior in the absence of the objects or states of affairs the items are “about”. In developing this picture, we are especially concerned to make room for kinds of manipulada that are *not* propositional in format, taking as our key example the re-use, off-line, of circuitry developed to smooth the on-line production of fluent motor action. Such circuitry looks able to support mental imagery and the off-line rehearsal of motor routines. Genuinely non-propositional manipulada such as these, we suggest, may lay the groundwork for similar operations with *internal models of linguistic symbols*, and thus act as a kind of bridge between structures implicated in skilled motor engagements in the here and now and structures capable of “standing in” for distal or non-existent states of affairs. Only these latter structures, we suggest, should really count as internal representations.

Dennett’s picture, it seems to me, must likewise confront the difficult question of how the human species became capable of the prodigious feats of real symbol manipulation in which we engage. Saying that we internalize capacities to manipulate external symbolic items just pushes this question one step back. Still, we need to understand what non-linguistic skills prime the

human brain to thus benefit, so wonderfully, from the provision of external structure. This is a question to which no-one, alas, yet has a compelling answer (see Deacon (1997) for an interesting attempt).

The notion of “bridging manipulada” strikes me as a promising way to fill this gap. But such manipulada will be *neither* propositional in format *nor*, initially, objects of witting, or knowing, manipulation. The transition from unwittingly used, non-propositional, internal manipulada to full human symbolic competence then constitutes (as Dennett notes: MTT, p. 5) an additional mystery, but perhaps one that can indeed be chipped away at by reflecting on the kinds of initially unwitting public linguistic practices that Dennett frequently highlights (see e.g. his (1991) p. 788).

The second assumption challenged in TAT concerns the “isolability” of the vehicles of content from the outside world, including the bio-mechanics of the body. Once again, the story makes good sense in the skill and tool based context we have been exploring. Dennett questions the idea of a pure information processing device neatly bounded at each side by transducers and effectors (TAT, pp. 2-3). Another way to put the point is perhaps this: that once you have a *skill-based* vision of content, you can no longer restrict the psychologically relevant system to the system that trades in internal representations. Instead, the primary unit of analysis (to borrow a term from Hutchins (1995)) spreads to include items such as bodily form (see Dennett’s comments on the bird’s wing, etc., TAT, p. 5) and local context (see the comments on distributed intelligence, MTT, p. 8). Here, then, is a further important consequence of the emphasis on tools and skills: the notion of cognition itself, and hence of the true targets of psychological understanding, is subtly altered to encompass *all* the tools and tricks that promote adaptive success. The special sub-class that involves internal representations (understood as manipulable content-bearing tokens) is again not privileged, and emerges as just another part of the mosaic. (This, I think, may be one way of reading Dennett’s strictures (TAT, pp. 4-5) against the image of the “walking encyclopedia”).

The attack on isolability has a second dimension also. The issues here concern the degree of insulation and “fire-walking” *between* cognitively important processes. Dennett’s point here is that a certain kind of inter-process leakage can be a powerful source of effects which may then be co-opted into new, adaptively potent, strategies. Dennett here talks of features which are not random noise but whose usefulness emerges only when neighboring components learn to use the features to create “new functional structures” (TAT, p. 7). One concrete case that has some of the features Dennett describes, and which I will float for comments and reactions, involves recent work on what has become known as evolveable hardware.

Certain chips, called FPGAs (field programmable gate arrays) contain logic blocks that can be re-organized and reprogrammed in situ, thus falling between our old notions of software and hardware. Using a new, easy-to-reconfigure chip (an Xilinx XC6216), Thompson (1996)(1997) ran a process of simulated evolution (a modified genetic algorithm) to find a chip design capable of distinguishing two sounds (a 1 kHz and a 10 kHz tone). The evolutionary regime was, importantly, working with a real Xilinx chip, repeatedly re-programming (in effect re-wiring) and testing actual physical devices. After 5000 generations, the chip was able to perform the task. What is interesting, from the present point of view, is how it worked. The successful chip used just 21 of the available 4000 logic blocks, making it “one or two orders of magnitude smaller than one would expect from conventional methods” (Thompson (1997) p. 389). And it bought this efficiency, it seems, by exploiting low-level physical features and inter-component leakage in ways that pressed functionality from what a human engineer would have tried to suppress. Some logic blocks, in fact, looked to be disconnected and idle, yet deleting them caused the circuit to fail. The explanation, according to Thompson, is that “they must be interacting by some subtle property of semiconductor physics such as electromagnetic coupling or interacting through the power supply wiring” (Thompson (1997) p. 389).

Here, then, we have a concrete (albeit low-level) case in which inter-component “leakage” and “noise” provides the raw material for a new adaptive strategy, just as Dennett suggests.

Moreover, and this is why I mention the example, there is a serious sense in which the exploitation of this kind of effect challenges the traditional pictures of computation and of software as an autonomous level in nature. Thompson and his colleagues prefer to view the chip as “a dynamical system, not a computational one”, since its success depends so heavily on the “continuous-time....unfolding of the laws of semiconductor physics” (op cit, p. 389). The idea, I think, is that one cannot understand how this system works by thinking of it in terms of an interconnected array of logic boxes, each one performing an isolated function or operation.

Several themes thus come satisfyingly together. Recall Dennett’s comments (TAT, p. 2) about the distortive effects of restricting psychological explanations to operations on propositional contents, instead of really trying to understand how a component accomplishes a task, and his skepticism (TAT, pp. 4-5) about cognitive scientific “boxology” and its imagined separation of software and hardware. The story of the evolved FPGA chip captures, I suggest, the force of these worries and shows how they also bear on issues concerning computation and internal representation. Standard computational and representational approaches, we may well suspect, are overly focussed on one kind of content and one kind of mind-tool, and insufficiently sensitive to the ‘depth’ of the canvass upon which nature sketches its strategies for adaptive success.

One can, of course, try to re-cast the notions of representation and computation to embrace more and more of the territory Dennett and others are exploring. My own tendency, indeed, has been to do just that (see Clark (1997b)). But what matters most is not, of course, what we call these strategies. Only that we recognize the roles they play in enabling intelligent behavior. In fact, things only get *really* radical when proponents of alternative approaches attack the very idea that psychological explanation must plot relations between states identified, at least in part, in terms of their *contents*. I do not *think* that Dennett, for all his worrying about propositional modes of content capture, wants to do that. But we should at least note the following question: In attacking the fixation with propositional modes of content specification and drawing attention

to the complex roles of body, world and physical hardware, does Dennett envision a content-based but largely non-propositional science of the mind, or does he envision a “withering away” of the role of content altogether? The answer seems to be (see e.g. TAT, p. 2) that he still sees cognitive scientific explanation as essentially content-involving, but that we should be aware of the inadequacy and inaccuracy of most (all?) of our specifications of content. But this is a point that might reward further clarification.

5. Numerical Cognition: A Case Study.

Here’s the big picture. Minds are grab-bags of (userless) tools for thought. Propositional content specifications typically fail to capture the essence of the tools. Content and “aboutness” is determined by the skilled engagements made possible by the tool-complexes. Talk of internal representations is, however, increasingly appropriate as the skilled engagements become more flexible and open-ended courtesy of the creation of more manipulable content-bearing tokens. This process probably originates in experiences of *external* symbol use.

This is a compelling but rather abstract and schematic story, so let’s end with a concrete case-study to help focus the questions and issues that arise.

Dehaene et al (1999) ask: “does the human capacity for mathematical intuition depend on linguistic competence or on visuo-spatial representations?” (op cit, p. 970). The answer that emerges is “both” – what makes us uniquely (as a species) competent in mathematics is the interplay between two distinct sets of mind-tools, with human embodiment and motor-skills acting as a source of what we dubbed (section 4) “bridging manipulada”.

Dehaene (1997) and Dehaene et al (1999) adduce a diverse body of evidence in support of a certain picture of human mathematical skill. The picture has three main components:

- (1) It depicts an innate, biological competence at low-grade approximate arithmetic:
a simple number sense, shared by infants and other animals, and involving the

appreciation of changes in quantity, of relative quantities, and of a handful of absolute quantities (oneness, twoness, and threeness).

- (2) It depicts a culturally acquired capacity to think about exact quantities (other than 1, 2, and 3) as dependent upon verbal and language-specific representations of numbers
- (3) It speculates that the cultural-evolutionary processes that allowed us to develop the symbol systems supporting exact mathematics crucially involved the use of body-parts as stand-ins for numbers.

Preliminary evidence for the “two systems” hypothesis comes from studies (Dehaene et al (1999)) of Russian-English bilinguals. Subjects were taught, in one of the two languages, a set of exact or approximate sums of two-digit numbers. They had to select the correct sum from two candidates. In one condition (the exact condition) they were told to select the answer from two numerically close candidates. In a second condition (the approximation condition) they were told to estimate the result and select the closest candidate. After training, performance in the approximation condition was shown to be unaffected by switching the language, whereas in the exact condition, language switching resulted in asymmetric performance, with subjects responding much faster if the test-language corresponded to the training-language. From this and several related studies, Dehaene et al concluded that the knowledge used for the approximation tasks is stored in a non-linguistic format, whereas the knowledge used for exact arithmetic is stored in a language-specific format. Dehaene (1997, p. 102) observes, in this vein, that different languages and number notations yield different typical numerical skills. English speakers asked to memorize 7 single digit numbers have a 50% chance of failure whereas Chinese speakers nearly always succeed. This is because Chinese number words are so brief. The typical numerical-list memory of speakers of a certain Cantonese dialect is a full 10 digits, revealing digit-span as a “culture-and-training-specific value [that] cannot be taken to index a fixed biological memory size parameter” (Dehaene (1997) p. 103). (Speakers of any language

can, however, improve their skills by the use of new mind-tools such as digit grouping and re-coding).

A second line of evidence for the two systems draws on lesion studies in which (to take one example) a patient with severe left-hemisphere damage cannot determine whether $2+2$ is 3 or 4, but reliably chooses 3 or 4 over 9, indicating a sparing of the approximation system.

Finally, and perhaps most dramatically, Dehaene et al (1999) present neuroimaging data from subjects engaged in exact and approximate numerical tasks. The exact tasks show significant activity in the speed-related areas of the left frontal lobe, while the approximate tasks recruit bilateral areas of the parietal lobes implicated in visuo-spatial reasoning. These results are presented as a demonstration “that exact calculation is language dependent, whereas approximation relies on nonverbal visuo-spatial cerebral networks” (op cit, p. 970) and that “even within the small domain of elementary arithmetic, multiple mental representations are used for different tasks” (op cit, p. 973).

Dehaene (1997) rounds off the story with some cultural-evolutionary and developmental speculation. The cultural-evolutionary question is, of course:

“How did *homo sapiens* alone ever move beyond approximation?”

(op cit, p. 91)

And the answer, in part, is of course “the human ability to devise symbolic numeration systems” (op cit, p. 91): the ability to create external representations, that become internalized as language-dependent mind-tools. This process began, Dehaene suggests, with the use of body parts as stand-ins for numbers. Once you use one, two, and three fingers (say) as stand-ins for the biologically appreciable quantities of oneness, twoness and threeness, it becomes possible to accidentally discover that an additional oneness can be tracked by associating the new quantity with a fourth finger. Historical and multi-cultural studies, reported in detail by Dehaene, show

the body-part roots of many names for numbers. The developmental process is also discussed, and here Dehaene makes some nice points about the need to somehow establish links between the linguistic labels and our innate sense of simple quantities. At first, it seems, children learn language-based numerical facts *without* such appreciation. According to Dehaene, “for a whole year, children realize that the word “three” is a number without knowing the precise value it refers to” (Dehaene (1997) p. 107). But once the label gets attached to the simple innate number line, the door is open to understanding that all numbers refer to precise quantities, even when we lack the intuitive sense of what the quantity is (e.g. my intuitive sense of 53-ness is not distinct from my intuitive sense of 52-ness). What all this amounts to, in Dennett’s terms, is of course the gradual installation, via a route that passes through external symbolic notations, of a new and potent virtual machine in the head: a culturally-incubated mind-tool for exact arithmetic. But this mind-tool, Dehaene insists, retains some of the character of the basic biological arithmetical device. Dehaene approvingly quotes Von Neumann:

“When we talk about mathematics, we may be discussing a *secondary* language, built upon the primary language used by the central nervous system”

(Von Neumann (1958), quoted in Dehaene (1997) p. 236)

This recalls some issues first raised way back in section 2. Are all mind-tools really on a par, or are some privileged – the *real* tools of understanding? What we may dimly glimpse in this example is a way in which each bald assertion contains an element of truth. Humans must, according to Dehaene, ultimately use an analog, and rather fuzzy, number line to represent quantity: a conjecture that explains the so-called distance effect by which it takes us (unlike a digital computer) longer to compare two close numbers than two distant ones. Exact mathematical tasks do, however, require the use of internal representations of number symbols. But our claim to genuine numerical *understanding* seems to rest on the presence of *both* kinds of skill, and (crucially) the presence of certain bridges and links that put our exact calculations in contact with some kind of biological sense of quantity and relative position in an array.

Numerical understanding thus depends crucially on *both* systems, but it is in a real sense “grounded” in the more biologically basic one. The image to *avoid* the, is one of simple translation from a public to a biological code. For, at least on my gloss, no such translation is possible or necessary. Instead, the two resources *collaborate*, and there are crucially important *links and bridges* between them. But seeing that this picture is deeply different from the picture of translation into a non-linguistic code is, I think, the heart of the copernican revolution that Dennett seeks to foment.

So *is there* an internal representation of, let’s say, ‘98’? I think the answer that now emerges is ‘no’! What we have is a genuine internal representation of the *word* and of the *numeral*, and genuine internal representations of rough quantity and relative location in an array. Mathematical understanding depends heavily on the interplay and links between our skills with the representations of number words and our skills with the biological resource. And in the latter case, at least, the skills are not rooted in familiar (digital, discrete token manipulating) kinds of computation but in the operation of a fundamentally analog device.

Conclusions: The Big Lacuna(e)

I’ve tried to show just how radical and challenging Dennett’s emphasis on skills and mind-tools can be, and how it subtly shifts the focus from simple notions of internal representation and computation to the complexities of agent-environment interaction. Along the way, I’ve tried to highlight and explore some themes that are less commonly discussed than (say) Dennett’s view about consciousness and the intentional stance, but that seem to me to be the essential backdrop to all the rest. These include the image of mind-tools *all the way down*, and the idea that no mind-tool or class of mind-tools are essentially privileged in an explanation of thought and understanding.

The rather concrete tales concerning hardware evolution and numerical cognition were then meant to show both (very) much Dennett has got right, and how (very) far there is still to go.

The emphasis on skills, multiple mind-tools, and loops involving external symbol systems all look correct, as do the qualms about simple notions of content and computation, and about the isolability of inner vehicles. But there are trouble spots. It's not *really* clear, for example, what holds the bag of mind-tools together as an agent – can narrative really bring persons into being? Nor is it clear just how to conceive the delicate balance between the idea of privileged (because biologically basic) mind-tools and the idea that real (florid?) understanding requires the co-presence of multiple tools, including the more recent tools derived from public symbol systems. It's not *at all* clear how best to understand the crucial processes of interaction and co-ordination by which the two broad classes of mind-tool interact. And it is not *at all* clear what, in the biological evolution of our species, opened the door to the creation and use of potent external symbol systems and thus got the Dennettian ball rolling in the first place. With lacunae like these, who needs enemies?

¹ 'Moshiwakemai', according to Henser (1999), means an apology made in the context of a relationship whose hierarchical structure renders the apologist unable to "presume on the indulgence" of the other, so it typically occurs when the lower ranking persons wants to apologize to the higher. See Henser (1999) p. 33.

² Herbert Simon took this route, after observing the importance of environmental complexity in determining behavior. He adds that:

a human being can store away in memory a great furniture of information...I would like to view this information-packed memory less as part of the organism than of the environment to which it adapts.

(Simon (1981) p. 53)

The 'organism' for Simon thus shrinks to something uncomfortably small (not unlike the read/write head of a Turing machine or the CPU of a digital computer). But nothing *that* impoverished seems likely to equate with our intuitive idea of an agent, who surely has goals, plans, memories and so on.

³ In thus opting for a clear dividing line between species in terms of the cultural imprinting of a user-illusion, Dennett may create a subtle background tension with his own deflationary account of qualia. For in rebuffing the fans of qualia, Dennett sometimes accuses them of "inflating differences in degree [of richness, control, etc.] into imaginary differences in kind" (Dennett 1997, p. 419). But Dennett *also* wants to claim that humans really are *different*, possessing (courtesy of language and cultural immersion) an informational organization that makes us (and not the lions, etc.) genuinely conscious (recall e.g. Dennett (1998) p. 347 quoted earlier). I find it hard to reconcile this notion of an organizational dividing line *among species* with Dennett's equally firm insistence (see Dennett (1997)) that *within* the human species, various phenomena of response and discrimination mark only differences in degree. For pretty clearly *some* of those phenomena are rooted in phylogenetically old pathways that we share with other animals, while others will hook directly into the kinds of new informational organization created by our "immersion in human culture" and responsible for "our kind of consciousness" (which now looks to be the real kind – the kind associated with florid representing).

⁴ Much of what Dennett has to say here seems nicely compatible with the sort of skill-based approach associated with the work of Gareth Evans and various contemporary apostles of 'non-conceptual content'. Grush (1998), for example, follows Evans in suggesting that many aspects of experiential content are *constituted* by sensorimotor skills. To experience something as *pulsating*, for example, is to be able in principle (i.e. assuming no bio-mechanical breakdowns) to co-ordinate a number of motor actions (swinging a baton, tapping your fingers) with the sound. A being lacking *all such* skills could not, on this account, directly perceive a sound *as* pulsating (see Grush (1998) para. 21); though they might infer that it is. Moreover, the content of such skill-based states is never reducible to a set of propositions believed: instead, the sensorimotor skills are what creates the contents that can later, perhaps, figure in other (propositional) episodes of thought and reflection. Dennett's tendency to view skilled

engagement as the root of all content (see SMR and section 2) seems motivated in a somewhat similar way, though without the emphasis on experiential content.