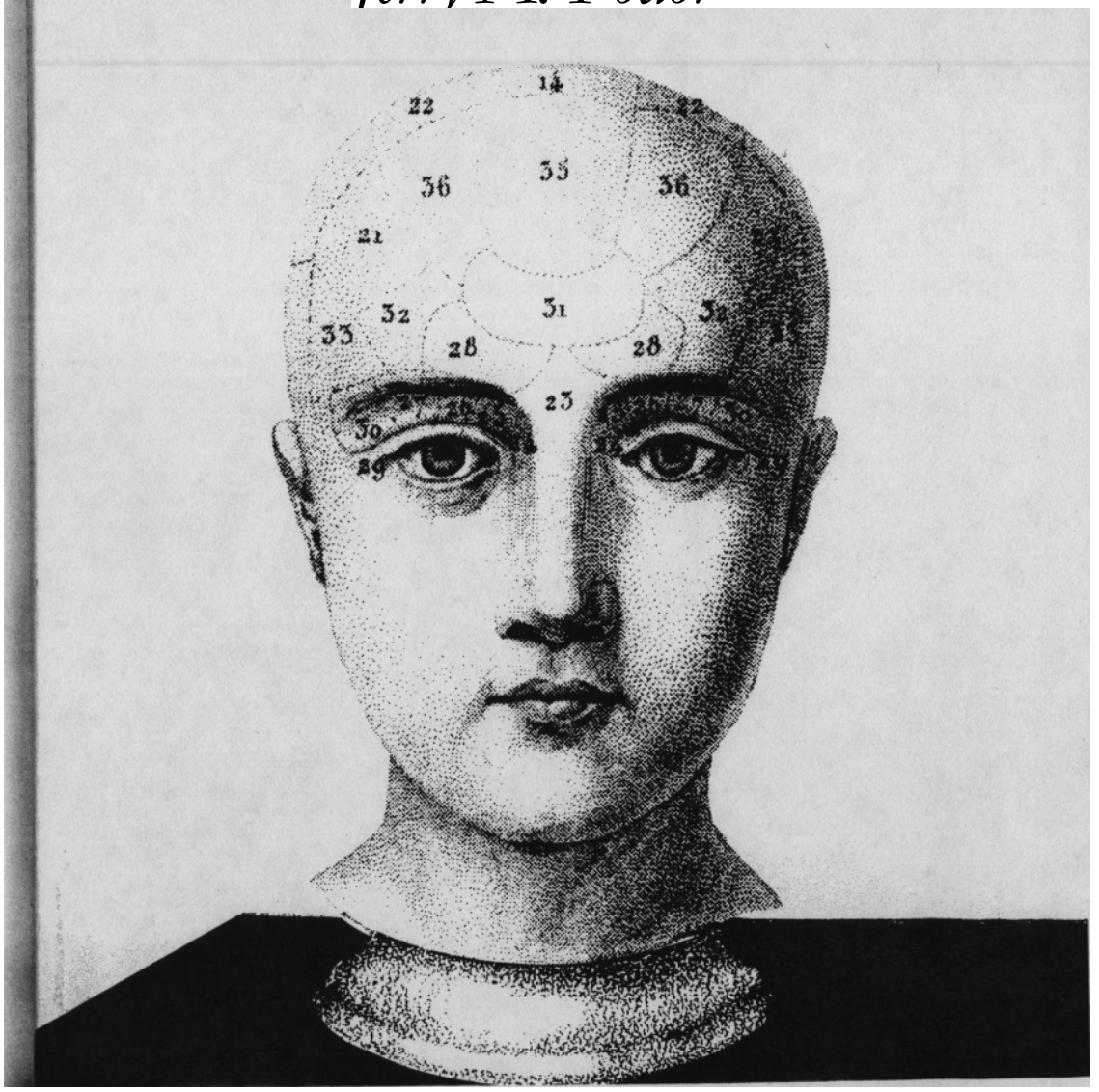
THE MODULARITY OF MIND

Jerry A. Fodor



FACULTY PSYCHOLOGY is getting to be respectable again after centuries of hanging around with phrenologists and other dubious types. By faculty psychology I mean, roughly, the view that many fundamentally different kinds of psychological mechanisms must be postulated in order to explain the facts of mental life. Faculty psychology takes seriously the apparent heterogeneity of the mental and is impressed by such prima facie differences as between, say, sensation and perception, volition and cognition, learning and remembering, or language and thought. Since, according to faculty psychologists, the mental causation of behavior typically involves the simultaneous activity of a variety of distinct psychological mechanisms, the best research strategy would seem to be divide and conquer: first study the intrinsic characteristics of each of the presumed faculties, then study the ways in which they interact. Viewed from the faculty psychologist's perspective, overt, observable behavior is an interaction effect par excellence.

This monograph is about the current status of the faculty psychology program; not so much its evidential status (which I take to be, for the most part, an open question) as what the program is and where it does, and doesn't, seem natural to try to apply it. Specifically, I want to do the following things: (1) distinguish the general claim that there are psychological faculties from a particular version of that claim, which I shall call the *modularity thesis; (2)*

enumerate some of the properties that modular cognitive systems are likely to exhibit in virtue of their modularity; and (3) consider whether it is possible to formulate any plausible hypothesis about which mental processes are likely to be the modular ones. Toward the end of the discussion, I'll also try to do something by way of (4) disentangling the faculty/modularity issues from what I'll call the thesis of *Epistemic Boundedness:* the idea that there are endogenously determined constraints on the kinds of problems that human beings can solve, hence on the kinds of things that we can know.

I shall, throughout, limit my brief to the psychology of cognitive processes, that being the only kind of psychology that I know anything about. Even so, this is going to be a rather long and rambling story, a fault for which I apologize in advance. My excuse is that, though I think the revival of the faculty psychology program has been enormously helpful in widening the range of serious options for cognitive psychologists to pursue, and while I also think that some version of the modularity thesis is very likely to prove true, still the atmosphere in which recent discussions have taken place has been on the steamy side, and a number of claims have been run together that are-or so I'll argue-conceptually distinct and unequally plausible. Moreover, there is quite a lot of ground to cover. A proposed inventory of psychological faculties is tantamount to a theory of the structure of the mind. These are serious matters and call for due expatiation.

PART I
FOUR ACCOUNTS OF MENTAL STRUCTURE

Behavior is organized, but the organization of behavior is merely derivative; the structure of behavior stands to mental structure as an effect stands to its cause. So much is orthodox mentalist doctrine and will be assumed throughout the discussion on which we're now embarked: Canonical psychological explanations account for the organization of behavior by appealing to principles which, they allege, explicate the structure of the mind.

But whereof does the structure of the mind consist? Not, to be sure, the clearest of questions, but nonetheless a pregnant one. I propose, in this section, to consider faculty psychology as one sort of answer that this question can plausibly receive. (Strictly speaking, I shall regard it as *two* sorts of answer, as will presently emerge.) The primary object of this exercise is to delineate the character of faculty theorizing by contrasting it with several alternative accounts of the mind. My way of carving up these options departs, in some respects, from what I take to be standard, and perhaps the eccentricities will edify. Anyhow, I should say at the start that the positions about to be surveyed need not be understood as mutually exclusive. On the contrary, the view ultimately espoused will be, in a number of respects, quite shamelessly eclectic.

1.1. Neocartesianism: the structure of the mind viewed as the structure of knowledge

As practically everybody knows, Descartes' doctrine of innate ideas is with us again and is (especially under Chomsky's tutelage) explicitly construed as a theory about how the mind is (initially, intrinsically, genetically) structured into psychological faculties or "organs." I am inclined to view this Cartesian revival as very nearly an unmixed blessing. However, I think it is important to distinguish the Neocartesian sort of faculty psychology from other, rather different versions of the doctrine with which it is easily confused and whose rhetoric it has tended to appropriate. In fact, most of this essay will defend a notion of psychological faculty that is rather different from Chomsky's "mental organ" construct, and of which Descartes himself would quite probably have disapproved. The following discussion is by way of sorting out some of these strands.

In a nutshell, the central Neocartesian claim is that "intrinsic (psychological) structure is rich ... and diverse" (Chomsky, 1980, p. 3). This view is contrasted with all forms of Empiricism, by which it is "assumed that development is uniform across (cognitive) domains, and that the intrinsic properties of the initial state (of the mind) are homogeneous and undifferentiated—an assumption found across a spectrum reaching from Skinner to Piaget (who differ on much else)" (ibid.). Issues about innateness will recur, in one or another aspect, through much of what follows. But, for now, I want to put them slightly to one side and try to see what notion

of mental structure is operative in this Neocartesian style of psychological theorizing.

Chomsky likes to speak of mental structures on anatomical analogy to hearts, limbs, wings and so forth. "We may usefully think of the language faculty, the number faculty, and others as 'mental organs,' analogous to the heart or the visual system or the system of motor coordination and planning. There appears to be no clear demarcation line between physical organs, perceptual and motor systems and cognitive faculties in the respects in question" (ibid.). There is, of course, a point to this analogy. It rests largely in the contention (entirely plausible, in my view) that for mental faculties, as for bodily organs, ontogenetic development is to be viewed as the unfolding of an "intrinsically determined process." In particular: "... we take for granted that the organism does not learn to grow arms or to reach puberty.... When we turn to the mind and its products, the situation is not qualitatively different from what we find in the case of the body" (ibid., pp. 2-3). But though Chomsky's point is well taken, his terminology is in some respects misleading; important distinctions are obscured by a use of 'structure' that applies promiscuously to bodily organs and psychological faculties as Neocartesians construe the latter. It is, indeed, only when we insist upon these distinctions that we can see clearly what the Neocartesian account of mental structure actually amounts to.

It turns out, upon examination, that what Chomsky thinks is innate is primarily a certain *body of information:* the child is, so to speak, 'born knowing' certain facts about universal constraints on possible human languages. It is the integration of this innate knowledge with a corpus of 'primary linguistic data' (e.g., with the child's observations of utterances produced by adult members of its speech community) that explains the eventual assimilation of mature linguistic capacities.

It is, perhaps, not very important to this Neocartesian story that what is innate should be, strictly speaking, *knowledge.* After all, knowledge is-or so many philosophers tell us-inter alia a nor-mative notion, having much to do with the satisfaction of standards of justification. Chomsky is himself quite prepared to give up the claim that the universal linguistic principles are innately *known* in favor of the explicitly neologistic (hence sanitized) claim that they are innately

"cognized." (See, especially, op. cit., p. 9.) It is, however, important to the Neocartesian story that what is innately represented should constitute a bona fide object of propositional attitudes; what's innate must be the sort of thing that can be the value of a propositional variable in such schemas as 'x knows (/believes,/cognizes) that P'.

Here is why this is important. As previously remarked, it is the fate of the (presumed) innate information to interact with the child's primary linguistic data, and this interaction is assumed to be *computational.* Now, the notion of computation is intrinsically connected to such semantical concepts as implication, confirmation, and logical consequence. Specifically, a computation is a transformation of representations which respects these sorts of semantic relations. (See Fodor, 1975; Haugeland, 1981.) It is, however, a point of definition that such semantic relations hold only among the sorts of things to which propositional content can be ascribed; the sorts of things which can be said to *mean that P.* The idea that what is innate has propositional content is thus part and parcel of a certain view of the ontogeny of mental capacities-viz., that in cognitive development, what is endogenously given is computationally deployed.

So, Chomsky's account of language learning is the story of how innate endowment and perceptual experience interact *in virtue of their respective contents:* The child is viewed as using his primary linguistic data either to decide among the candidate grammars that an innately represented 'General Linguistic Theory' enumerates (Chomsky, 1965) or to 'calibrate' endogenous rule schemas by fixing parameter values that the innate endowment leaves unspecified (Chomsky, 1982). This sort of story makes perfectly good sense so long as what is innate is viewed as *having* propositional content: as expressing linguistic universals, or rule schemas, or whatever. But it makes no sense at all on the opposite assumption.

Seen from this perspective, it is perhaps the *differences* between endogenous psychological and anatomical 'structure' that appear most striking. It may be that the development of arms and the development of anaphora each critically involves the exploitation of a specific genetic endowment. And it may also be that what is innate can, in each case, be described as 'information' in the relatively uninteresting statistical sense that implies only non-randomness. But there is, surely, no reason to suppose that the

development of arms requires access to innately given *propositional contents*. There is nothing that growing arms requires one to cognize, innately or otherwise. By contrast, as we've seen, that propositions about anaphora (inter alia) are innately cognized is the very burden of Chomsky's plaint; ineliminably so, since it is precisely these innately cognized propositional contents that do the theoretical work in Chomsky's account of language development.

It is, I think, the essence of the Neocartesian style in psychology to assume that mental structure should be explicated largely by reference to the propositional contents of mental states. In this respect, no doubt, the new Cartesianism bears the imprint of Descartes' own largely epistemological concerns. Descartes was, after all, mainly interested in determining what sorts of things can be known, and with what degree of certainty. In his epistemology, the primary explicandum is our ability to recognize certain truths (of geometry, of theology, of metaphysics, or whatever); and the prototypical form of explanation is to exhibit these truths as identical to, or deducible from, propositions that are innately given and self-evident. Where the overriding motive is the explanation of propositional knowledge, it is perhaps hardly surprising that one should come to view mental structure largely in terms of the organization of propositional content.

I say that this strategy is prototypically Cartesian but, of course, it is on display as early as Plato's *Meno,* where the slave boy's ability to answer questions of geometry that Socrates puts to him is explained by reference to "opinions" that were always "somewhere in him."

SOCRATES: What do you think, Meno? Has he answered with any opinions that were not his own?

MENO: No, they were all his.

sOC: Yet he did not know, as we agreed a few minutes ago.

MENO: True.

sOC: But these opinions were somewhere in him, were they not? MENO: Yes.

In Descartes and Plato, as in Chomsky, the nativism is so striking that one is likely to overlook a still deeper consensus: the idea that certain of the subject's cognitive capacities should be explained by reference to consequence relations (e.g., deductive relations) that hold among the propositions that the subject knows (believes, cog-

nizes, or whatever). I say to you: "What's 2 plus 17?" and you, being good at that sort of thing, say "19." Your *behavior* is structured in the relevant sense; what sort of *mental* structure is the psychologist to posit in explaining your behavior? According to the Cartesian, it is inter alia the *deductive* structure of number theory to which the explanation must appeal. You know things about the numbers from which it follows that 2 plus 17 is 19, and this knowledge is somehow recruited-perhaps the deductions are literally drawn-when you answer the question. Similarly, according to generative linguistic theory, your ability to detect syntactic ambiguities, distinguish well-formedness from ungrammaticality, respond selectively to the noun-phrase that has been topicalized, and so forth are to be explained by reference to *what is entailed* by the grammar that you learned when you learned your language. In short, your linguistic capacities explain your verbal behavior, and are them-selves explained by reference to the *content of your beliefs. You* can spot the ambiguity of 'they are flying planes' because, so the story goes, (i) You have learned the grammar of English, and (ii) it follows-deductively-from what you have learned that 'they are flying planes' has two well-formed parsings.

So, to return to ontogenetic issues, when Chomsky says that there is an innately specified "language organ," what he means is primarily that there are truths (about the structure of possible first languages) that human beings innately grasp. When he says that the mind of the child is "intrinsically structured," what he means is primarily that there are innately specified propositional contents. When he says that the theory of language learning is the story of how the language faculty matures, what he means is primarily that the ontogeny of linguistic capacities is the unfolding of the deductive consequences of the innate beliefs in interaction with a body of perceptual data. The moral: Chomsky really is a bona fide Cartesian in ways that go deeper than his nativism; the paradigm for mental structure, in Chomsky's theorizing as in Descartes', is the implicational structure of systems of semantically connected propositions.

There are aspects of mental organization for which Chomsky's version of the Cartesian story is, in my view, extremely persuasive. But, precisely for that reason, it is important to emphasize that there are other, quite different, sorts of things that a theorist may have in mind when he talks of endogenous psychological structures.

For example, consider memory. If one is going to postulate innately specified faculties, memory is, surely, a plausible candidate. Yet *memory isn't a faculty in the Neocartesian sense of that notion.* Having a memory isn't a matter of having one or another set of beliefs, and if memory is an innate capacity, that couldn't be because there is some set of propositions that organisms are born cognizing. There isn't, in short, the remotest temptation to identify the structure of memory with the inferential structure of a body of propositions. Memory is, so one supposes, some sort of *mechanism,* analogous to a hand or a liver or a heart. Viewed hypostatically at least, memory really does seem to be a kind of mental *organ* in ways that the putative language faculty, even viewed hypostatically, really does not.

The difference between these two notions of psychological faculty will be fundamental to much of what follows; perhaps an example will make the distinction clear. Suppose one believes the doctrine of George Miller's famous paper about the 'magical number seven' (Miller, 1956). Roughly, the idea is that there is a fairly constant limit on the number of unfamiliar, unrelated items that one can cope with in a task that demands immediate recall. (So, if I ask you to repeat a list of nonsense syllables, then the longest list you'll be able to manage on a first presentation will be on the order of seven items, give or take a bit.) Now, one can imagine a Neocartesian treatment of this phenomenon along the following lines: there is a certain mentally represented proposition to which one gives tacit assent-viz., the proposition that, when presented with a list of *n* things to learn, one should indeed learn the first seven and there-upon forget about the rest. (Perhaps this principle is not just cog-nized and adhered to, but also endogenously specified; for present purposes it doesn't matter.)

I said that it is possible to imagine a Neocartesian story that runs along those lines, but I doubt that any Neocartesian would take it seriously; and I'm sure that nobody else would. The sort of treatment that Miller's data cry out for is not the postulation of an innately cognized rule but rather of a psychological mechanism-a piece of hardware, one might say-whose structure somehow imposes limitations upon its capacities. To put it with all possible crudeness: the picture is that there's a box in your head and when you try to put more than seven things in it, some of the things start to fall out.

Perhaps it goes without saying that I'm not endorsing this picture; in fact, I'm not even committed to Miller's idea that there is an item-bounded short-term memory. The point is rather to emphasize a distinction between two quite different accounts of what mental structures-endogenous or otherwise-might be like; one account elaborated around a notion of propositional content and the other around the notion of a psychological mechanism. The former view of mental structure is typically Neocartesian; the latter, however, is not.

I remarked at the outset that the various notions of faculty psychology that I'll be reviewing aren't necessarily mutually exclusive. A Neocartesian could—in my view, a Neocartesian *should*—perfectly well take the line that mental-organs-qua-propositional-structures are only part of the story that faculty psychologists have to tell, much of the rest of the story being involved with the postulation of mental mechanisms. Indeed, it's hard to see how this suggestion could reasonably be resisted. That you say "19" when I say "7 + 12, please" is, no doubt, partly to be explained by reference to what you know about the numbers. But there must be more to it since, after all, knowledge doesn't eventuate in behavior in virtue of its propositional content alone. It seems obvious that you need mechanisms to put what you know into action; mechanisms that function to bring the organization of behavior into conformity with the propositional structures that are cognized. This is the problem of 'performance' in one of Chomsky's uses of that notion. Performance mechanisms do for Chomsky some of what the pineal gland was supposed to do for Descartes: they are invoked to answer the question "How does the structure of behavior come to mirror the propositional structures that one cognizes?"

Equally pressing for a Cartesian, however, is a subtler and prior question-one which I think Descartes himself never faced—viz., "How does the structure of *thought* come to mirror propositional structure?" According to the Cartesian account, you can figure out that 7 plus 12 equals 19 because you know things about the numbers from which it follows that 7 plus 12 equals 19. But, surely, this explanation is an enthymeme; it must be short for something like "You can figure out ... because it follows from what you know about the numbers *and you have some way of figuring out* (some of) *what follows from what you know about the numbers."*

In short, even assuming the Cartesian story about endogenously cognized propositions, we need answers for questions of the form: "Given that so and so entails such and such, in virtue of what psychological mechanisms is the organism able to infer from cognizings of so and so to cognizings of such and such?" Psychological faculties may well be invoked to answer this sort of question; faculties which mediate, for example, the representation, retention, retrieval, and inferential elaboration of the cognized propositions. These faculties-patently not mental organs as Neocartesians understand that notion-would nevertheless count as bona fide mental structures and might well themselves be innately specified (or, if they are not, then their ontogeny has to be accounted for, just as the ontogeny of propositional knowledge does). The point is, once again, that this sort of mental structure does not consist in the internal representation of propositions, and a nativism of such structures would not be a theory of innate *beliefs*. The Neocartesian appropriation of the terminology of mental faculties, organs, and mechanisms to express what is, in fact, a nativism of propositional attitudes tends to obscure this difference; but alertness to it is essential to understanding the range of options available for theory construction in cognitive science.'

1.2. Mental structure as functional architecture: horizontal faculties

We turn, then, to a different notion of mental structure, one according to which a psychological faculty is par excellence a sort of mechanism. Neocartesians individuate faculties by reference to their typical propositional contents (so that, for example, the putative language organ is so identified in virtue of the information about linguistic universals that it contains). By contrast, according to the present account, a faculty is individuated *by reference to its typical effects,* which is to say that it is functionally individuated. If there is a language faculty in this sense of faculty, then it is whatever piece of (presumably neurological) machinery functions to mediate the assimilation and employment of verbal capacities.

One way to appreciate this distinction between faculties-cum-belief-structures and faculties-cum-psychological-mechanisms is to notice that even theorists who are blatantly Empiricist in respect

of the former may nevertheless be (anyhow, closet) Nativists in respect of the latter. This was, in fact, John Locke's position according to some authorities.

> ... Locke thought too obvious to mention explicitly in the *Essay* ... the existence of natural faculties such as perception, understanding and memory, and innate mental powers like those of abstraction, comparison and discernment. The 'white paper' metaphor is meant to indicate that the understanding (and hence the mind) is originally empty of *objects* of thought like ideas; but it has whatever apparatus is necessary to acquire them through experience, and then to derive knowledge by comparing and contrasting them with each other.' [Harris, 19771

So, then, the (noncartesian) faculty psychologist is per se interested in the analysis of mind into interacting component mechanisms.' However, the history of this kind of faculty psychology exhibits two variants of the doctrine according to the axis along which the mind is sliced. According to the most familiar version-which I shall call 'horizontal' faculty psychology-cognitive processes exhibit the interaction of such faculties as, e.g., memory, imagination, attention, sensibility, perception, and so forth; and the character of each such process is determined by the particular mix of faculties that it recruits. However, the character of mentation is more or less independent of its subject matter; the faculties are supposed to be invariant from one topic of thought to the next."

For example, traditional accounts of the mind often acknowledged a faculty of *judgment,* whose characteristic function was supposed to be the recognition of identities and differences among mental contents (in one terminology among Ideas). A very refined judgment is one which can distinguish between even very similar Ideas (in the manner, say, of John Austin distinguishing a mere accident from a full-blooded inadvertence). Judgment found work to do in (e.g.) perceptual recognition, where the categorization of current sensory data is supposed to require comparing it with information from memory; but the details needn't concern us here.

Now, this faculty of judgment might get exercised in respect of matters aesthetic, legal, scientific, practical, or moral, and this list is by no means exhaustive. The point is that, according to the horizontal treatment of mental structure, *it is the self-same faculty*

of judgment every time. The discrimination of identity and difference among aesthetic ideas is thus performed by precisely the same psychological mechanism that distinguishes, as it might be, weight from mass or torts from misdemeanors. On this view, then, aesthetic judgment is simply the application of the faculty of judgment to the process of drawing aesthetic distinctions. It follows that there is no such thing as a faculty-of-aesthetic-judgment per se. A fortiori, there is no such thing as an aesthetic faculty.

Or consider memory again. A recurrent theme in the traditional literature is the treatment of memory as *a place* where beliefs are stored. Plato has it at one point in the *Theatetus* that memory is like a birdcage; one, as it were, reaches in and pulls out the thing recalled:

SOCRATES:... let us suppose that every mind contains a kind of aviary stocked with birds of every sort, some in flocks apart, some in small groups, and some solitary, flying among them all.

THEATETUS: Be it so. What follows?

SOC: When we are babies, we must suppose this receptacle empty, and take the birds to stand for pieces of knowledge. Whenever a person acquires any piece of knowledge and shuts it up in his enclosure, we may say he has learned or discovered the thing of which this is the knowledge, and that is what "knowing" means.

THE: Be it so.

soc: Now think of him hunting once more for any piece of knowledge that he wants, catching, holding it, and letting it go again.

This sort of architectural analogy is quite characteristic of faculty psychologies in general. The mind has an intrinsic structure, and mental contents have instantaneous locations with respect to this enduring background; things happen in the mind, and what can happen is constrained by the character of the mental layout .[5]

What makes Plato's story about memory a version of *horizontal* faculty psychology, however, is his view about how the birds are kept. The crucial point is that all the memories are in the same place. Or if, as many modem theories would have it, there are *several* memory systems, all horizontal faculties, then presumably each memory may pass through every such system. More precisely, where a given memory is at a given instant depends, perhaps, on how much time has elapsed, or on how much rehearsal there has been. But what it does *not* depend upon is the *content* of the memory. For example, there could not, in point of definition, be a *horizontal* faculty that is specific to remembering 'events' as opposed to remembering 'propositions', or to remembering faces as opposed to remembering tunes. By definition, such content-specific faculties would fail to be horizontal.

As remarked above, more evolved forms of faculty psychology than Plato's tend to think of mental architecture as, at least in the first instance, functional rather than literally spatial. A memory system is thus individuated by reference to its characteristic operations, it being left open whether there are distinct areas of the brain that are specific to the functions that the system carries out. However, the idea of a horizontal faculty survives the abandonment of spatial principles of individuation in favor of functional ones. Instead of speaking of the *location* of a mental content at time *t*, one speaks of the set of mental processes that have access to that content at t-roughly, the set of processes for which it constitutes a domain at t. So, a content that is 'in' short-term memory (but not in long-term memory) at 2:35 on the morning of the 5th is one to which short-term memory processes (but not long-term ones) have access at that date and time. A thoroughly horizontal faculty, functionally individuated, is thus one to which *every* mental content may be accessible at one time or other. Probably nobody believes that there really are horizontal faculties in that very strong sense, but the idealization establishes a useful point of reference.

That's about all that I propose to say about horizontal faculties just now. The character of the construct will emerge in contrast with alternative theoretical options. For present purposes, a horizontal faculty is a functionally distinguishable cognitive system whose operations cross content domains. I shall assume without argument that mental processes are computational insofar as they are cognitive, hence that the typical function of cognitive mechanisms is the transformation of mental representations (see Fodor, 1975). It follows that each distinct cognitive faculty must effect a characteristic pattern of such transformations. I shall also assume that we can make some sense of individuating content domains independent of the individuation of cognitive faculties, since if we cannot the question whether the operation of such faculties cross content domains doesn't arise. I suppose this latter assumption to

be not unreasonable. If, for example, there is some psychological mechanism that is engaged both in the identification of wildflowers and in the balancing of one's checkbook, then we have, prima facie, good reason to suppose that mechanism to be horizontal.

1.3. Mental structure as *functional architecture: vertical faculties*

Horizontal faculty psychology has been with us always; it seems to be the common-sense theory of the mind. By contrast, the 'vertical' tradition in faculty psychology has specifiable historical roots. It traces back to the work of Franz Joseph Gall (1758-1828), the founding father of phrenology and a man who appears to have had an unfairly rotten press.

According to Gall, the traditional census of horizontal mental faculties is largely a fiction. There is, in particular, no such thing as judgment, no such thing as attention, no such thing as volition, no such thing as memory; in fact, there are no horizontal faculties at all. Instead, there is a bundle of what Gall variously describes as propensities, dispositions, qualities, aptitudes, and fundamental powers; of these an aptitude for music will do as an example. (I should emphasize that Gall does *not* himself speak of 'vertical faculties'. I have coined that term to suggest a certain reading of Gall's text—viz., that he agrees with traditional faculty theories that the mind is structured into functionally distinguishable sub-systems, but disagrees about how the divisions between these systems should be drawn.)

From the point of view of a modern cognitive psychologist, Gall's aptitudes constitute something of a mixed bag. Indeed, there is a sense in which aptitudes are a mixed bag from *anybody's* point of view, since the term applies indiscriminately to both *competences* and *proclivities.* An aptitude to commit murder (to mention another of Gall's examples) is a propensity rather than a talent; you're apt to commit murder if you're inclined to kill, however clumsily you carry out your homicides. Compare an aptitude for music, which one lacks unless one is *good* at-not just inclined toward-things musical. This slight tendency of the concept of an aptitude to misbehave may have misled Gall into thinking that his vertical faculties have more in common than in fact they do. Certainly the census of vertical faculties that Gall acknowledges pays less attention to the distinction between cognition and volition than most theorists now believe to be proper.

Anyhow, in the case of what Gall sometimes calls the "intellectual" capacities, it is useful to identify an aptitude with competence in a certain cognitive domain; in which case, the intellectual aptitudes (unlike, n.b., the horizontal faculties) are distinguished by reference to their subject matter. It is of central importance to understand that, in thus insisting upon domain specificity, Gall is not simply making the conceptual point that if music (e.g.) is distinct from mathematics, then musical aptitude is correspondingly distinct from mathematical aptitude. Gall is also claiming that the psycho-logical mechanisms which subserve the one capacity are different, de facto, from those that subserve the other. I take it that this claim is the heart of Gall's theory.

In fact, some of Gall's favorite analogies for aptitudes are ethological. Nest-building and bird song are presumably not to be viewed as applications of a general intellectual capacity to the accomplishment of specific ends; it would thus be a mistake to postulate a horizontal faculty of avian intellect of which competence in singing and nesting are among the manifestations. Similarly with man: "There are as many different kinds of intellect as there are distinct qualities.... One individual may have considerable intellect relative to one fundamental power, but a very narrow one in reference to every other...:, special faculty of intellect or understanding is as entirely inadmissible as a special faculty of instinct" (p. 240) (all Gall quotations are from Hollander, 1920). Intellect per se could not, therefore, be neurologically localizable, any more than instinct per se could be subserved by a specific brain mechanism.

Gall's point is precisely analogous to one that could be made by denying that there is such a thing as *acuity.* There are, no doubt, visual acuity, auditory acuity, and perhaps gustatory and intellectual acuity as well. And one might add that a given individual may have considerable acuity relative to one fundamental power, but very narrow acuity in reference to every other. However, since visual, auditory, gustatory, and intellectual acuity are surely just parameters of vision, audition, taste, and intellect respectively, it follows that there could be no such things as *a faculty of acuity;* that would be the wrong way to carve things up. Acuity, to put it in trendy terms, is syncategoramatic; and so, for Gall, is intellect.

Moreover, what is true of intellect and acuity is also true of memory, judgment, volition, attention, and the rest of the horizontal faculties; on Gall's account they are, one and all, the spectral progeny of misplaced concreteness. "Perception and memory are only at-tributes common to the fundamental psychological qualities, but not faculties in themselves; and consequently they can have no proper centers in the brain" (p. 240). In this respect, the horizontal faculties, which Gall denigrates, are explicitly contrasted with the vertical faculties, which he endorses; the latter correspond to specific brain mechanisms which Gall hoped, sooner or later, to locate:

> Take the musician. He would not be a musician if he did not *perceive* the relation of tones, if he had no *memory* of music, if he could not judge of melody and harmony.... Thus attention, perception, memory, judgment and imagination are nothing else than different modes of action of every one of the fundamental capacities. When the primary mental power is energetic so will these attributes be; when it is feebly developed, there will be a feeble degree of attention, of perception, of memory, a defective judgment and no imagination.... We have to discover` the fundamental powers of the mind, for it is only these that can have separate organs in the brain. [p. 238]

It is perhaps not surprising, since Gall emphasizes the specificity of the neural mechanisms which subserve the vertical faculties, that he should infer from neural specificity that there is what we would call genetic determination:

> The influence of education, instruction, example and of surrounding circumstances acts principally when the innate dispositions are neither too feeble nor too energetic.... The impressions received through our senses from external sources are not the origins of our aptitudes, talents, sentiments, instincts and propensities.... The propensities and instincts, the aptitudes and talents, the intellectual abilities and moral qualities of men and animals are innate. [pp. 250-251]

This style of theorizing, combining nativism with an emphasis upon the domain specificity of cognitive capacities, will seem familiar to those who have been exposed to what John Marshall calls the "new organology."[6]

Much of what follows in this section will be concerned with the elaboration of Gall's vertical faculty idea, since it seems to me that there is much in this notion that modem cognitive science would do well to ponder. First, however, Gall's positive proposals need to be disentangled from a couple of arguments which he thinks show that horizontal versions of faculty psychology must be seriously defective. These arguments were portentous; they go rumbling down the history of psychology, repeated again and again (usually without citation of their source). However, despite their influence in reinforcing the antifaculty bias in much modem psychological theorizing, they actually aren't very convincing.

Gall's major argument against horizontal faculties turns on the idea that if there is only *one* faculty of (say) memory, then if some-body is good at remembering *any* sort of thing, he ought to be good at remembering *every* sort of thing. That is, Gall thinks the existence of a unitary horizontal faculty of memory would imply that an individual's capacity for recalling things ought to be highly correlated across kinds of tasks (across what I have been calling cognitive domains). Similarly, mutatis mutandis, for judgment, imagination, attention, and the rest. "If perception and memory were fundamental forces, there would be no reason why they should be manifested so very differently, according as they are exercised on different objects. There would be no reason why the same, and, in fact, every individual, should not learn geometry, music, mechanics and arithmetic, with equal facility since their memory would be equally faithful for all these things" (pp. 240-241). This is, perhaps, supposed to be a sort of 'Leibnitz' Law' argument: the same faculty cannot be both weak and strong, so if it sometimes happens that mathematical memory is weak and musical memory robust, then the memory that mediates mathematics can't be the same as the memory that mediates music.

If, however, that is the argument, it is clearly fallacious. All that can be inferred, strictly speaking, is that mathematical memory = musical memory; which, though patently true, is quite compatible with mathematical memory and musical memory being exercises of the self-same faculty with respect to mathematics in the one case and music in the other. To put the point slightly less ponderously: there is no obvious reason why the same faculty should not be strong in one employment and weak in another, so long as the employments are not themselves identical.

It would thus be open to a faculty psychologist of the horizontal persuasion to suggest that what is characteristic of each mental capacity is the specific mix of horizontal faculties that it recruits, and to explain the unequal distribution of, e.g., memory across cognitive domains by reference to the interaction effects that different mixes of faculties give rise to. It now seems clear, for example, that the fact that top-level chess players remember distributions of chess pieces better than they remember other sorts of things does *not* warrant the conclusion that there is a specific memory for chess. On the contrary, it turns out that the operative principle is that, quite generally, one remembers what one understands (Bartlett, 1932; Bransford, Barclay, and Franks, 1972). The chess player's ability to remember where the pieces are is thus part and parcel of his grasp of how they might have got there. Witness the fact that it disappears when the pieces are set down in ways that don't make sense (DeGroot, 1965). Spearman (1927, pp. 35-36) remarks that the 'problem of correlation'-in effect, the interaction of the level of functioning of a faculty with the cognitive domain in which it is employed-is the insuperable difficulty for horizontal versions of faculty psychology: " . . . the vital point is the degree of inter-dependence, or, as it is commonly called, the amount of correlation." It is certain that Gall would have accepted this evaluation. Yet it is unclear, in light of the considerations just rehearsed, that a horizontal faculty psychology actually would have to predict the sorts of correlations that Gall and Spearman suppose it would; or that the failure to find such correlations would prove very much one way or the other.

The argument we've just been discussing turns on the claim that the various employments of presumptive horizontal faculties do not correlate *across cognitive domains.* But Gall has a (slightly irritating) tendency to run that argument together with one which emphasizes the failure of mental capacities to correlate *across individuals.* We'll have a quick look at this.

Every faculty psychologist has to find some motivated way of answering the question "How many faculties are there?" One way that Gall seeks to do so is to find the parameters that a psychology of individual differences would need to acknowledge, and then to postulate a distinct faculty corresponding to each such parameter. It is thus among Gall's pet arguments for distinguishing between

a pair of faculties that people can differ in the degree to which they have them. Jones is good at mathematics and awful at metaphysics, and Smith has the reverse aptitudes. So the mathematical and metaphysical competences must be subserved by distinct psychological and neural mechanisms; they must be, in effect, distinct (vertical) faculties.

Now this determination to connect issues about faculties with issues about individual differences is itself something of a departure, on Gall's part, from the beaten paths of the faculty psychology tradition. As Spearman remarks:

> Through the earlier part of ... [the]. .. historical development of the doctrine of faculties, few if any writers were much concerned with the problem ... of the differences between individuals. The purposes for which faculties were first devised, and for a long time almost exclusively employed, had not been to portray the aspects in which men differ, but those which characterize them all alike ... [1927, p. 29]

Nor is it entirely clear what, on Gall's view, reflection upon the existence of individual differences is supposed to add to the arguments against horizontal faculties that we reviewed just above.

The mere fact that Smith and Jones differ in their musical abilities wouldn't seem, in and of itself, to suggest the existence of a specifically musical faculty. Assume that all faculties are in fact horizontal, but that some 'mix' of such horizontal faculties is optimal for musical accomplishment (lots of perceptual acuity, say, a dash of sensibility, and very long fingers; [actually, I don't know much about music, though I do know what I like]). Well, for any such optimal mix of horizontal faculties there will surely be differences in the degree to which people approximate possessing it. If Jones outwhistles Smith, that is because his mix comes closer to the optimum than Smith's does; or so, at least, the proponent of horizontal faculties has every right to suggest, for all the argument to the contrary that we've got so far.

Perhaps, however, what Gall has in mind is this: if Smith and Jones differ in refinement of musical judgment but not, say, in refinement of practical judgment, then it must be true either of Smith or of Jones (or of both) that his musical and practical judgments are unequally refined. But if someone's musical and practical

judgments can be unequally refined (or, indeed, unequally F for any F whatever), then the two kinds of judgment must ipso facto be distinct. If this is what is going on, however, then the individual differences argument reduces to the Leibniz' Law argument previously disapproved of.

Gall's fascination with, and insistence upon, degrees of individual difference is a most striking feature of his writings. Yet it sits badly with another of Gall's favorite themes: the repeated analogizing of faculties to instincts. That Gall apparently didn't feel the tension between these views was perhaps due to a confusion of (to put it very roughly) issues about genetic determination with issues about *species specificity,* the source of the mix-up being that certain sorts of individual differences are inherited just as species-specific psychological traits like instincts are. It may be, for example, that the ability to play really first-class baseball rests on a characteristic bundle of physiological and perceptual-motor endowments. In which case, one wouldn't be absolutely stunned to discover that that ability is inherited to some interesting extent. But of course that would be no reason to suppose that baseball is a species-specific behavior - in anything like the ethologist's sense of that notion. In particular, you wouldn't want to infer from its (putative) heritability that baseball playing has a specific neurological basis, or a specific evolutionary history, or that there are genes for playing baseball. Aptitude for baseball playing, even if inherited, is patently not interestingly like an instinct.'

To put it in a nutshell, what is instinctive is genetically determined, but the reverse clearly doesn't have to hold. In fact, if what you have in mind by a vertical faculty is something like what the ethologist has in mind by an instinct, you probably will *not* want to postulate vertical faculties corresponding to parameters of individual differences; not even where such differences are inherited. On the contrary, in the study of instincts, the natural theoretical idealization is to a genetically and neurologically homogeneous population; instincts are forms of *species-specific* behavior. If one takes the analogy between instincts and 'fundamental powers' seriously, one must suppose-precisely contrary to the methodology that Gall endorses-that vertical faculties are to be inferred from the discovery of competences that are relatively invariant across subject populations.

The moral of all this .critical discussion may be only that Gall's theories are sometimes more interesting than his polemics; a situation not without precedent in the history of important scientific innovations. On the other hand, if, as I believe, Gall's arguments against horizontal faculties are less persuasive than his arguments in favor of vertical ones, then the possibility remains open of a 'mixed' model in faculty psychology-one in which some but not all of the mental architecture is vertically arranged. We'll return to this later.

For now, let's put the 'problem of correlation' and the stuff about individual differences to one side. We can then distinguish four major ingredients of Gall's notion of a fundamental power: vertical faculties are *domain specific,* they are *genetically determined,* they are associated with *distinct neural structures,* and-to introduce a new point-they are *computationally autonomous.* The relevant consideration about computational autonomy is that Gall's fundamental powers do not share-and hence do not compete for-such horizontal resources as memory, attention, intelligence, judgment or whatever. This view of vertical faculties as not merely distinct in the functions they perform, but also relatively independent in the performance of their functions, will be important later when we turn to consider the notion of a cognitive module.

Suffice it, for present purposes, to note that his emphasis upon the computational autonomy of vertical faculties is one of the chief points that distinguishes Gall's theorizing from Chomsky's. For example, Chomsky (1980) suggests that there is perhaps a mathematical faculty. But, as one might expect in the light of the discussion in Part 1.1, what he appears to mean by this is only part of what Gall would have meant. Chomsky's claim is primarily that some mathematical information (specifically, the idea that you can generate the natural numbers by adding one indefinitely) is innately specified. Gall would quite probably have liked that, but he would have claimed considerably more. *Qua architectural* nativist, Gall's view would be that the psychological *mechanisms* of memory, judgment, imagination, will, or whatever that mediate mathematical reasoning are themselves innately specified. *Qua vertical faculty theorist,* Gall's view would be that these mechanisms, insofar as they come into play when you do mathematics, are only nominally related to the memory, judgment, imagination ... etc. that are en-

gaged when you talk or commit homocides.[8] And, *qua autonomy theorist*, Gall's view would be that the mental operations that go on when you do mathematics do not much interact with and, specifically, do not much interfere with others of one's mental capacities. That we can, most of us, count and chew gum at the same time would have struck Gall as a fact that offers significant perspectives upon our mental organization.-

It is important to emphasize that innateness and computational autonomy, in particular, are quite *different* properties of cognitive systems, only the first being at play in Chomsky's notion of a mental organ. Suppose, to take an extreme case, that knowledge of Peano's axioms is innate; they are not learned but genetically transmitted. It wouldn't follow, even from this radical thesis, that there is an arithmetic faculty in Gall's sense. For, the hypothesis that arithmetic knowledge is genetically transmitted is—but the vertical faculty thesis for arithmetic is not—compatible with the possibility that the psychological mechanisms that mediate arithmetic reasoning are the same ones that underlie the capacity for abstract thought in general. It is thus compatible with Chomsky's notion of a mental organ, but not with Gall's notion of a vertical faculty, that arithmetic reasoning shares (horizontal) psychological resources with jurisprudential reasoning, aesthetic reasoning, or filling out one's income tax.'

It is worth adding that, just as the innateness thesis for fundamental powers does not imply their organization into computationally autonomous vertical faculties, so the horizontal analysis of a cognitive capacity would not imply that that capacity is learned. Most faculty psychologists have, in point of historical fact, been nativists of the horizontal persuasion. It may be that there is use for the notion of horizontal cognitive organization, particularly in light of the possibility of a mixed model which includes both vertical and horizontal elements. It would not follow that there is much use for (or much sense to be made of) the notion that mental structures are learned. (See Fodor, 1975.) It is thus important to disentagle the horizontal faculty story from any form of Empiricism.

A final word about Gall. It seems to me that the notion of a vertical faculty is among the great historical contributions to the development of theoretical psychology. So, why isn't Gall honored in the textbooks? The story of Gall's posthumous reputation is a

sad illustration of the maxim that the good men do is oft interred with their doctoral dissertations. Gall made two big mistakes, and they finished him: he believed that the degree of development of a mental organ can be measured by the relative size of the corresponding brain area, and he believed that the skull fits the brain "as a glove fits a hand." Phrenology followed as the night the day,[10] and with it all sorts of fraud and quackery, for none of which Gall was responsible but for much of which he appears to have been retrospectively blamed. It is lucky for us that we don't make mistakes any longer; those who do so clearly have little to expect from history or from the intellectual charity of their professional colleagues.

1.4. Associationism (and: Whatever Became of Faculty Psychology?)

I now want to take a brief look at yet a fourth way of answering the question: "How are cognitive capacities organized?" I shall refer to this tradition as 'associationism' (though I do so with some trepidation, contemporary versions of the doctrine having shed much of what the label once implied). Roughly, associationism is related to the claim that there are faculties in something like the way that phenomenalism is related to the claim that there are tables and chairs; you can take them to be incompatible, or you can read associationism as saying that faculties exist but that they have the status of constructs out of some more fundamental sort of entity. On either interpretation, however, associationists denied much of what faculty psychologists wished to assert, so that the ascendence of the former doctrine implied the decline of the latter.

Baldwin's (1911) *Dictionary of Philosophy and Psychology-in 3* volumes, so by no means an insubstantial tome-allows "faculty psychology" a single scanty paragraph. It deserves quotation, since it illuminates the nominal (though not, I believe, the real) cause of the eclipse of that tradition.

> To say that an individual mind possesses a certain faculty is merely to say that it is capable of certain states or processes. But we find in many of the earlier psychologists a tendency to treat faculties as if they were causes, or real conditions, of

> the states of processes in which they are manifested, and to speak of them as positive agencies interacting with each other. Thus persistence in voluntary decision is said to be due to extraordinary strength of will, or to will-power or to the faculty of will. Certain mental processes in man are said to have their source in the faculty of reason, and certain other processes in lower animals are explained by the existence of a faculty of instinct. This mode of pretended explanation has received the name of Faculty Psychology. Locke, in criticizing the phrase 'freedom of the will', has brought out very clearly the nature of the fallacy involved. 'We may as properly say that the singing faculty sings, and the dancing faculty dances, as that the will chooses, or that the understanding conceives....'

This passage contains, by my count, one importantly false statement and two bad arguments. To begin with: it is simply not the case that "to say that an individual mind possesses a certain faculty is merely to say that it is capable of certain states or processes." There are, of necessity, far more mental capacities than there are psychological faculties on even the most inflationary census of the latter. For example, our mental capacities include the ability to add 1 plus 1, the ability to add 1 plus 2, the ability to add 1 plus 3 ... and so on for indefinitely many drearily similar cases. And all these capacities are (presumably) to be attributed to the operation of *one and the same mathematical faculty*. The situation would not be different in any principled way if we were to assume that there is a subfaculty of the faculty of mathematics specially in charge of the addition of finite integers. You still get indefinitely much mental capacity out of each faculty you posit, this being simply a special case of the general principle that every causal agent has indefinitely many potential effects. A census of faculties is *not*, in short, equivalent to an enumeration of the capacities of the mind. What it is instead is a theory of the *structure of the causal mechanisms that underlie the mind's capacities*. It is thus perfectly possible for all hands to be agreed about what *capacities* a mind has and still to disagree about what *faculties* comprise it. Contemporary examples of such disagreements include: whether human maternal behaviors are instinctive; whether the ability to talk is an expression of 'general intelligence', etc.

Of the two bad arguments Baldwin endorses, the second—Locke's—is simply beside the point. No faculty psychologist is in fact required to say that the singing faculty sings, or that the dancing faculty dances, or that the will chooses or any such thing. He can—and should—rather say that the *organism* sings, dances, chooses, or whatever in virtue of the operation of the various faculties that it possesses. As for the understanding, it conceives one's argument only as one's stomach digests one's dinners—viz., synecdochically.

The more important of Baldwin's arguments-at least in terms of historical influence-is the first, which consists simply of a charge of vacuous hypostatization. This claim-that the postulation of mental faculties is ipso facto a form of pseudo-explanation-is practically universal in the secondary sources, the decline of the faculty tradition being attributed to widespread recognition that such postulations are indeed empty. For example, D. B. Mine (1970, p. 374) has this to say: "Subsequent criticism of (Christian Wolfe's) faculty doctrine was an elaboration of the kind of objection raised by Descartes and Locke ... the objection revealed an appeal to faculties to be a question-begging kind of explanation as revealed by invoking an aquatic faculty to explain swimming or a terpsichorean faculty to explain dancing. This is the equivalent of substituting an impressive label for a genuine explanation, as in saying that some salve will heal a rash because it contains a therapeutic ingredient."

Connoisseurs of heavy irony will find much to please them here; for, after all, what this supposedly conclusive objection has against faculty psychology is only that faculties are individuated by their effects—i.e., that they are *functionally* individuated. And it is, of course, this very strategy of functional analysis which, according to the now standard philosophy of psychology, allows the individuation of mental constructs to steer a proper course between the unacceptable ontological alternatives of eliminative materialism on the one hand and dualism on the other. As Ned Block summarizes the doctrine in his excellent introduction to the contemporary functionalist literature (Block, 1980, p. 172): "Functionalists can be physicalists in allowing that all the entities (things, states, events, and so on) that exist are physical entities, denying only that what binds certain types of things together is a physical property.... Metaphysical functionalists characterize mental states in

terms of their causal roles." Not to put too fine a point on it: the functionalist idea is that pain is whatever is the normal cause of pain behavior; and, mutatis mutandis, the language faculty is what-ever is the normal cause of one's ability to speak. Functionalists take this line in full awareness of what Mauler said about dormative virtues; and, in my view, they are quite right to do so. (For further discussion see Fodor, 1965, and 1981b.)

This is not, of course, to say that the tactic of individuating mental entities functionally is ipso facto proof against vacuous explanation. It would be a bad idea (not to say an incoherent one—see above) to postulate a faculty corresponding to each prima facie distinct behavioral capacity and let it go at that. For one thing, not all prima facie distinct behavioral capacities really do differ in their etiology, and theory construction ought to find the causal uniformities beneath the heterogeneity of surface appearances. More-over, some capacities surely arise from the *interaction* of underlying causes; in fact, the more of these, the merrier the theorist, since his goal is to get the maximum amount of psychological explanation out of the smallest possible inventory of postulated causal mechanisms. None of this, however, has anything to do with faculty theorizing per se, since the corresponding remarks apply equally to *all* theoretical enterprises where the postulation of unobservables is at issue. Nor is it true, in point of historical fact, that faculty psychologists were particularly disposed to flout these general methodological canons. On the contrary, as Spearman (1930) correctly points out: "The general intention (in faculty theories) ... is to represent the countless transient mental experiences by a small number of relatively permanent—particularly innate—different principles. The multitudinous actual events are thus governed by very few 'potential' ones. [Vol. 1, p. 108]... The theory of faculties consists essentially in deriving multitudinous processes from a few powers" (p. 155). It's hard to imagine what alternative strategy could rationally be commended.

In retrospect, then, the supposedly decisive methodological arguments against faculty theory were, on the face of them, so silly that it's hard to believe (much) in their historical significance. And, indeed, isolated arguments—like isolated experiments—generally don't alter the course of science. What usually does the job is the emergence of an alternative theoretical enterprise. As I indicated

above, it seems pretty clear that what did for faculty psychology was the promise of an associationistic theory of mind. For just as Empiricist epistemology offered an account of the origin of mental *contents* which dispensed with the Cartesian postulation of innate ideas, so associationism offered an account of the ontogeny of mental *processes* which dispensed with the postulation of innate cognitive architecture-which, in short, dispensed with the need for faculties.

I take it that what an associationist (of either the classical mentalist or the more recent learning-theoretic variety) is prepared to acknowledge by way of explanatory apparatus in cognitive theory is this:

(a) A set of elements out of which psychological structures are constructed. Reflexes are the preferred elements for associationists who take it that psychological structures are behavioral; "Ideas" are the preferred elements for associationists who take it that psychological structures are mental.

(b) A relation of association defined, in the first instance, over the elements. (Only "in the first instance" because the property of being associable is preserved under association; the associative laws can apply to Ideas/Reflexes that are themselves products of association, thereby generating a distinction between elementary psychological structures and complex ones.)

(c) The laws of association. These are principles in virtue of which the character of an organism's experience determines which of its Ideas become associated or (mutatis mutandis) which conditioned reflexes get formed.

(d) Theoretically relevant parameters of the psychological structures and of the associative relations among them; so that, for example, associative relations can differ in respect of their strength and reflexes can differ in respect of their operant level.

Some associationists have been willing to acknowledge a scattering of irreducible horizontal faculties as well: for example, sensibility in the case of all the Classical Empiricists and imagination and reflection in the case of Hume and Locke respectively. But it seems clear that such concessions—often enough equivocal anyhow (see above, note 2)—are best viewed as unwilling. Ideally, according to the main stream of the associative tradition, all cognitive phenomena are to be accommodated by appeal to the very exiguous

theoretical apparatus just described. As Hume says *(Enquiries,* p. 321), association is a form of attraction which "in the mental world will be found to have as extraordinary effects as (gravitational attraction does) in the natural, and to show itself in as many and as various forms."

In consequence, a profoundly reductionistic impulse has characterized much of the boldest-psychological speculation in the Anglo-American tradition. The trick, for an associationist, is to show that there is nothing that faculties are required to explain, all bona fide psychological phenomena being reducible to the objects and relations enumerated in a-d. As usual, the treatment of memory provides revealing examples. So, Hume proposes to distinguish what is actually remembered from what is merely imagined not on logical grounds (you can imagine, but not remember, what didn't in fact occur), nor in terms of hypothesized differences in the underlying causal mechanisms (as a horizontal faculty psychologist would surely do) but rather by reference to the "force and vivacity" of the Ideas being entertained; whatever is remembered is assumed ipso facto to be more forceful and vivacious than anything that is merely conjured up. (Hume explains, with vast implausibility, that this is why history is always more gripping than fiction.) Hume's treatment is surely not attractive, but it exhibits in perfect microcosm the strategy of dissolving presumptive psychological mechanisms into parameters of the association relation or properties of the associated relata.

Curiously, the pursuit of this strategy sometimes led associationists to say things that sound very like Gall, though of course for quite different reasons. Thus Thorndike (of all people) echoes Gall's doctrine that there is no such thing as memory, and he cites Gall's evidence: the variability of recall across cognitive domains. Thorndike's account of this interaction is not, however, that retentiveness is a parameter of the operation of vertical faculties, but rather that it is a parameter of the association relation. "There is no memory to hold in a uniformly tight and loose grip the experiences of the past. There are only the particular connections between particular mental events and others"-which connections can vary in strength from one case to the next. (Quoted by Kline, 1970, p. 662.)

It is, of course, no accident that associationists devoted so much time to showing that the phenomena which faculties had previously been invoked to handle could be adequately explained with more parsimonious theoretical apparatus. Associationism developed in conscious and often explicit opposition to the older faculty tradition, and it was precisely the parsimony of the associationist's theory that was supposed to convince one of its scientific good repute. No Gothic proliferation of mental structures was now to be tolerated. The "how many faculties?" question would receive *a principled* answer at the associationist's hands: If a faculty is a primitive psychological mechanism—a *fundamental* power—then the answer is: "only one; only the capacity to form associations."

Thus far I've been reading the associationist tradition in a way that the associationists would themselves surely have found congenial: as proposing an *alternative* to faculty psychology, one characterized by a notable reduction in the amount of theoretical apparatus to be deployed in the explanation of cognitive phenomena. In recent decades, however, a sort of revisionist reading has developed, in which associationism is viewed less as *replacing* than as *reconstructing* the theoretical mechanisms that faculty psychologists worked with. A little background discussion is required in order to see how this could be so.

As I remarked above, contemporary cognitive theory takes it for granted that the paradigmatic psychological process is a sequence of transformations of mental representations and that the paradigmatic cognitive system is one which effects such transformations. I thus assume, for purposes of this essay, that if faculties *cum* psychological mechanisms are to be acknowledged in our cognitive science, they will be computational systems of one sort or another. Now, it is a major achievement of modern logic to have shown that computational processes of any complexity whatever are reducible to (or, looked at the other way, constructible form) concatenations of surprisingly small collections of basic operations. There are a number of notations in which such constructions can be expressed, Turing machine theory and production systems being among the most familiar. Very roughly, what they have in common is the postulation of a census of computational elements on the one hand, and of combinatorial operations on the other, the output of the theory being generated by the arbitrarily iterated application of the latter to the former.

If you don't mind a little anachronism, it is not impossible to see in this sort of logical apparatus the basis for a refined and purified associationism, the idea of sets of elements with combinatorial operations specified over them being what provides the common ground. Since the logical formalism permits the construction of computational systems of arbitrary complexity, the postulation of even an elaborate population of faculties is tolerable to this new sort of associationism. For, so long as the operation of the faculties is assumed to be exhaustively computational, they can be viewed as mere constructions out of whatever elementary 'associations' the theorist is prepared to acknowledge. Perception, memory, thought, and the rest of the faculty psychologist's brood can then be accepted as distinguishable aspects of mind (specifically, as distinct mental processes) without abandoning the basic associationistic premise that practically all of the mental life is "assembled"—i.e., put together from some relatively simple and uniform population of psychological elements.

There is quite a lot of recent psychological literature which, more or less explicitly, recommends this sort of computational reinterpretation of the associative tradition. A passage from Allport (1980) will serve to give the feel of the thing:

> In the old psychology ... linkages between a calling cue and a particular category of action were called 'habits'. The key idea ... was that actions ('responses') are addressed or evoked by particular calling conditions ('stimuli'). If we undo the restriction that these a-b pairs must be directly observable events, and instead interpret the a's and b's as specific 'states of mind', providing in addition some relatively simple mechanisms for their interaction, then this simple associationistic conception can have surprising power. Its simplest and most direct application in information processing terms can be seen in so-called 'Production Systems'.

Allport is by no means alone in commending this line of thought. *To* consider just one famous example, Miller, Galanter, and Pribram in their enormously influential *Plans and the Structure of Behavior* (1960) are explicit in offering the "TOTE unit" to *replace* the reflex as the element from which complex psychological structures are to be constructed, the constructivist program itself being accepted quite without visible hesitation (or argument).

However, this marriage of concepts from associationism with concepts from computer mathematics gives evidence of being a shotgun arrangement: it's hard to recognize either the theoretical commitments of associationism or the considerations which made those commitments seem plausible, given the computational reinterpretation.

For one thing, in the traditional literature, association was viewed as a mechanical relation *among* mental contents, not as a computational relation defined over them. Hume speaks of associations between Ideas on the model of gravitational attraction between physical objects; Skinner speaks of stimuli as *eliciting* the responses conditioned to them. Now, it is important to understand that this tradition of push-pull talk in associationism is not mere unreflective metaphor. On the contrary, it is part and parcel of the associationist's rejection of mental architecture—of psychological mechanisms whose function it is to 'process' mental contents. Right at the heart of associationism is the idea that you can dispense with such mechanisms in favor of intrinsic, dynamic relations (attraction, repulsion, assimilation, and so forth) among the psychological elements them-selves. This is, in its way, a brilliant-if doomed-idea (influenced, beyond any doubt, by the successes of Newtonian dynamics in physics); but it makes associationism a doctrine that is profoundly different in spirit from the picture of the mind that computational psychologists endorse.

For example, if we are to think of associated mental representations as somehow connected by *rule* rather than by mutual at-traction, then we will need mechanisms to apply the rules and also places to keep them when they are not in use. (Cf. Allport: "some relatively simple mechanisms for their interaction"; no bigger than a man's hand, as one might say.) Even Turing machines exhibit a minimal architecture of tape, executive, and reader; and any remotely plausible candidate for a computational model of cognitive processes would presumably require access to considerably more such apparatus than Turing machines make do with. But this *'functional* architecture' (as it's sometimes called; see Pylyshyn, 1980) is precisely the sort of unreduced mental structure that real associationists wanted very much to do without. The moral is: give up

the idea of dynamic relations among psychological elements in favor of the computational picture and you thereby give up a lot of what distinguishes Hume's picture of the mind from, say, Kant's.

Qualms about computational associationism are, however, by no means restricted to suspicions of historical unauthenticity. Deeper issues emerge if we ask why one should *want* to treat faculties as 'assembled' out of elementary psychological objects, even assuming the logical apparatus for effecting the construction to be available.

One answer that, of course, *won't* do is that you somehow increase the available computational power by treating faculties as constructs. On the contrary; it is a point of definition that you can't tell from the input-output capacities of a cognitive system whether it is, as it were, a primitive piece of mental architecture or something that has been put together from smaller bits. Computationally equivalent (that is, input-output equivalent) systems can, in principle, be built either way; from the point of view of an external device which communicates with them, all such systems count as the same machine. (You may be able to tell them apart because one rattles when you shake it and the other doesn't; but if so, the rattle doesn't count as part of the output.)

Moreover, similarity relations among cognitive systems far stronger than mere input/output equivalence can, in principle, be defined without broaching the issue of whether the systems should be viewed as assembled. Computer theorists, when they want to talk about computational systems in a way that abstracts from the difference between assembled and primitive processors, often speak of identities of *virtual* architecture. Roughly, you establish the virtual architecture of a machine by specifying which sets of instructions can constitute its programs. So, for example, there could be two devices, both of which can be programmed to perform simple arithmetic calculations, which are identical in *virtual* architecture in that both can execute instructions of the form 'add *m* to *n'*. However, it might be that the relation of the virtual architecture of these machines to their more elementary computational organization-and, eventually, to their physical organization-is quite different: for one of them, adding integers is a simple, primitive operation (performed, perhaps by making some measurement on voltages in a circuit); whereas, for the other, addition requires a sequence of mediating computations (as it would if the operations of a pocket calculator were to be simulated by a Turing

machine). For the second machine, then, addition is an assembled operation (and, in consequence, commands to add integers must be "compiled" into the appropriate sequences of elementary operation before they can be executed). The machines may nevertheless be identical (not only in their input/output functions but also) in the set of programs they can run; hence the possibility of identical virtual architecture between machines that are 'hardwired' in the one case and assembled in the other. In approximately this way, a traditional faculty psychologist and an associationist might end up agreeing about the virtual architecture of cognitive capacities, but disagreeing about whether the psychological mechanisms which mediate these capacities ought to be viewed as constructs.

Well, to end this excursis, the present question is why anything except virtual architecture should be of any interest to the psychologist; why, in particular, should anybody *care* whether faculties are assembled? What I think many cognitive scientists find persuasive—not to say mandatory—about the constructivist alternative is certain ontogenetic possibilities that it appears to offer. Specifically, if mental structures can be viewed as assembled from primitive elements, then perhaps mechanisms of *learning* can be shown to be responsible for effecting their construction. Here, then, is a real convergence between the motivations of classical associations and those which actuate its computational reincarnation: Both doc-trines find in constructivist analyses of mental structures the promise of an Empiricist (i.e., non-Nativist) theory of cognitive development.

But not, I think, with equal plausibility. We have seen that computational associationists are free to dispense with previously accepted constraints upon the sorts of mental structures that associationism can acknowledge; in principle, any computational mechanism can be reconstructed with the apparatus they have available. Arguably, however, it was only in light of his insistence upon an absolute *minimum* of virtual architecture that the classical associationist's Empiricism was remotely plausible.

The basic point about association was, surely, that it offered a mechanism for bringing about co-occurrence relations among *mental* events which mirror the corresponding relations among *environmental* ones. The feature of experience to which the formation of associations was supposed to be most sensitive was thus relative

frequencies of spatiotemporal contiguities among stimuli (Ideas be-come associated in virtue of spatiotemporal propinquities among the things that they are Ideas *of;* responses get conditioned in virtue of spatiotemporal propinquities between discriminative and reinforcing stimuli; and so forth). Correspondingly, the typical products of association are chains of Ideas (mutatis mutandis, response chains), these being the psychological counterparts of causal chains of environmental events. Not to put too fine a point on it, association was a mechanism for producing sequential redundancies in the mind (or in behavior) which mirror sequential redundancies in the world. This notion of mental structures, and of the environmental structures presumed to cause them, is no doubt depressingly crude; but at least one can imagine such associative chains being constructed from their elementary links under the influence of environmental regularities of the sorts that organisms actually do encounter. To that extent the classical associationist's ontogenetic theories fit together with his account of the structure of mature cognitive competence.

What the computational associationist offers instead is the possibility of mental structures of arbitrary complexity; he thus has a sort of guaranty that his associationism will never force him to accept an unduly impoverished notion of mental organization. But he pays a price: traditional associationist accounts of ontogeny can no longer be relied upon. There is simply no reason at all to believe that the ontogeny of the elaborate psychological organization that computational associationism contemplates can be explained by appeal to learning principles which do what principles of associative learning did—viz., create mental copies of environmental redundancies. In particular, the constructibility in *logical principle* of arbitrarily complicated processes from elementary ones doesn't *begin* to imply that such processes are constructible *in ontogeny* by the operation of any learning mechanism of a kind that associationists would be prepared to live with. This is a point about which I suspect that many contemporary psychologists are profoundly confused.

In short, as the operative notion of mental structure gets richer, it becomes increasingly difficult to imagine identifying the ontogeny of such structures with the registration of environmental regularities. Hence the main course of recent Cartesian theorizing, with its reiterated emphasis upon 'poverty of the stimulus' arguments: There would seem not to be enough ambient information available to account for the functional architecture that minds are found to have. You can, no doubt, make a language parser, or a visual scene recognizer, or a 'General Problem Solver' out of the sort of psychological elements that computational associationists acknowledge; this follows just from the assumption that parsers and scene recognizers and the rest are species of computers. What does *not* follow is that there is some way of constructing such systems from the information given *in experience.* But this consideration under- mines the main motivation for viewing mental structures as assembled in the first place—viz., that what is first exhibited as *assembled* can then be exhibited as *learned-indeed,* as learned *by association.* To put the point in a nutshell, the crucial difference between classical and computational associationism is simply that the latter is utterly lacking in any theory of learning. (There is, once again, a budget of heavy ironies to contemplate. After all, the historical point of associationism was largely to make Empiricism respectable. It was to do this precisely by providing a theory of learning which would show how mental structure could be accounted for without nativistic postulation. There was a guy in Greek mythology who got so hungry that eventually he ate himself; modern associationism may be said to have attained much the same condition.)

My present purposes being largely expository, I don't propose to pursue this line of argument; it is, in any event, familiar from Chomsky's work. Suffice it that insofar as environmentalist biases provide a main motivation for the computational associationist's constructivism, it is perhaps best seen as a failed attempt at reconciling faculty psychology with Empiricism. Conversely, latter-day nativists typically view constructivism in psychology with deep misgivings; if mental architecture is innately specified and if the ontogeny of cognition is primarily the unfolding of a genetic program, why should one expect that mental structures will prove to be assembled? The idea that they are hardwired—i.e., that the grain of their physical architecture quite closely parallels the grain of their virtual architecture-seems at least equally plausible.

As the last paragraph should suggest, neurological speculations are quite close to the surface here. Perhaps you can't tell from

outside whether a computational system is assembled or primitive, but you certainly ought to be able to tell from *inside*. The view of faculties as assembled comports with a view of the corresponding neurology as, at least initially, diffuse and equipotential; environmental tuition may effect local alterations in connectivity (for example), but it would be astonishing if it produced neural architecture and neural specificity on a large scale. By contrast, since the traditional faculty psychologist is a nativist down to his boots, he predicts a brain that is parsed into big, perhaps even macroscopic, neural structures. In this respect at least, the tradition that includes Gall runs through Wernicke and Broca (see Caplan, 1981).

This is, no doubt, all pretty loose-a matter less of demonstrative arguments than of elective affinities. Thus the constructivist may be interested in formalisms with the expressive power of universal computers, but I doubt that anybody actually thinks that the brain is really much like a Turing machine. Nor does the adjudication between virtual architecture and physical structure have to be made in the same way for every faculty; it is perfectly possible that operations that are primitive in one cognitive process may be assembled in another. For that matter, innately specified computational systems *could,* in logical principle, be put together from elementary operations; and learning could, in logical principle, result in elaborate and specific neural morphology. All we have is that neither of these contingencies seems very likely as a matter of fact. Let's leave it at this: the standard reason for stressing the distinction between virtual and physical architecture is to exhibit the actual organization of the mind as just one of the possibilities that could have been realized had the environment dictated an alternative arrangement of the computational elements. And a natural interpretation of neural hardwiring is that it packages into unanalyzed operations what may be quite powerful primitive computational capacities.

This looks like a good place for a little summary and prospectus.

Summary: In effect, what we have done so far is to suggest a number of questions that one can ask about a cognitive system in aid of locating it in relation to a general taxonomy of such systems. In particular:

1. Is it domain specific, or do its operations cross content domains? This is, of course, the question of vertical versus horizontal cognitive organization; Gall versus Plato.

2. Is the computational system innately specified, or is its structure formed by some sort of learning process?

3. Is the computational system 'assembled' (in the sense of having been put together from some stock of more elementary subprocesses) or does its virtual architecture map relatively directly onto its neural implementation?

4. Is it hardwired (in the sense of being associated with specific, localized, and elaborately structured neural systems) or is it implemented by relatively equipotential neural mechanisms?

5. Is it computationally autonomous (in Gall's sense), or does it share horizontal resources (of memory, attention, or whatever) with other cognitive systems?

Prospectus: I now propose to use this taxonomic apparatus to introduce the notion of a *cognitive module.* Two preliminary points, however. First, each of questions 1-5 is susceptible to a 'more or less' sort of answer. One would thus expect-what anyhow seems to be desirable-that the notion of modularity ought to admit of degrees. The notion of modularity that I have in mind certainly does. When I speak of a cognitive system as modular, I shall there-fore always mean "to some interesting extent." Second, I am not, in any strict sense, in the business of 'defining my terms'. I don't think that theoretical terms usually have definitions (for that matter, I don't think that nontheoretical terms usually do either). And, anyhow, the taxonomic apparatus just sketched is incomplete; what I take to be perhaps the most important aspect of modularity-something that I shall call "informational encapsulation"-has yet to appear. So what I propose to do instead of defining "modular" is to associate the notion with a pattern of answers to such questions as 1-5. Roughly, modular cognitive systems are domain specific, innately specified, hardwired, autonomous, and not assembled. Since modular systems are domain-specific computational mechanisms, it follows that they are species of vertical faculties.

I shall assume, hopefully, that this gives us a notion of modularity that is good enough to work with. The rest of this essay is devoted to doing the work. First, I want to try to refine the modularity concept by enriching the taxonomy. The goal is to suggest more properties that modular systems might have in common than the ones just mentioned, and also to try to see what it is that underlies the taxonomy: Why should there be modular systems? Why does

this cluster of properties tend to co-occur? Second, I want to say something about the extension of the concept; to propose a hypothesis about which cognitive systems are, in fact, modular. This second line of inquiry will provide the main structure of the discussion, the first emerging as opportunity provides targets. By the time I've finished, I shall have made the following suggestions:

(a) That the set of processors for which the modularity view currently seems most convincing is coextensive with a functionally definable subset of the cognitive systems.

(b) That there is some (more or less a priori) reason to believe that cognitive systems which do not belong to that functionally defined subset may be, in important respects, nonmodular (e.g., mediated by horizontal faculties). And finally,

(c) I shall make some depressed remarks along the following lines: though the putatively nonmodular processes include some of the ones that we would most like to know about (thought, for example, and the fixation of belief), our cognitive science has in fact made approximately no progress in studying these processes, and this may well be because of their nonmodularity. It may be that, from the point of view of practicable research strategy, it is only the modular cognitive systems that we have any serious hope of understanding. In which case, convincing arguments for non-modularity should be received with considerable gloom.