

In the first half 20th century behaviorism largely failed as an approach to understanding cognitive functions. What was the root cause of this failure? Are any of the concepts used in cognitive neuroscience today in similar jeopardy?

Behaviorism doesn't assist our understanding of cognitive functions for the most part because it doesn't purport to. Methodological behaviorism is an attempt to reconceptualize psychology as the science of behavior, rather than a science of mental states or cognitive functions. As such, methodological behaviorists maintain that we should either eliminate terms that refer to mental states from the vocabulary of a science of psychology, or that we should operationalize them so that they refer to observable behavior - or, more plausibly, to dispositions to behave. Dispositions are thought to be reputable scientific notions because they can be operationalized into observable behaviors. To say that a substance is soluble, for example, is just to say that in certain sorts of circumstances (such as being placed in water) certain sorts of observations (dissolving) would be found. Similarly, to say that a person is in a certain kind of mental state (e.g., pain) is just to say that if they were in certain kinds of circumstances then they would behave in a certain kind of way.

There are observations that the behaviorist paradigm seems incapable of explaining, however. Stimuli, response, and reinforcement history are unable to account for such phenomena as: filial imprinting, critical learning thresholds, fixed action patterns, species specific learning (e.g., easy to teach a hen to peck a key and a rat to push a lever but the converse is not the case), and species specific behaviors such as language (human infants will acquire syntax but not the infants of other species). Advances in ethology, genetics, and neurophysiology thus provided alternative paradigms for theorists and researchers who were interested in explaining those phenomena. Advances in computer science, linguistics, and cognitive psychology also provided alternative paradigms for researchers and theorists who were interested in explaining phenomena that didn't seem to be explainable from within the behaviorist paradigm. Examples of such phenomena include: a computer being programmed to perform an 'intelligent task' despite the lack of learning and reinforcement history, the seemingly rule governed combinatorial structure shared by syntax, logical thought, and computation, and limitations on what could be learned or remembered.

The first movement away from behaviorism thus came from the idea that there must be a categorical basis to dispositions. If we return to solubility it seems that there is something about the internal nature of substance in virtue of which it has its dispositional profile. To learn about the nature of bonds between atoms would seem to further our knowledge by explaining why the substance has the disposition to dissolve. Analogously, it was argued that behavioral dispositions must be realized in neurophysiology and that getting clearer on the categorical basis of dispositions was as legitimate a branch of psychology as getting clearer on the laws governing dispositional profiles. Thus it was argued that solubility should be identified with the categorical chemical bonds and mental states should be identified with the categorical neurophysiological underpinnings. The second movement away from behaviorism came from the notion that an adequate dispositional analysis required a more complex structure than that posited by behaviorists. While behaviorists limited themselves for the most part to explaining relatively simple stimulus - response (S-R) pairs, complex behavioral dispositions were best explained by positing intervening layers that interacted with other intervening layers (e.g., S - R/S - R/S - R). Getting clearer on the computational or functional profile of mental processing thus seemed to be as legitimate a part of psychology as the laws governing stimulus - response interactions. Thus it was argued that mental states should be defined according to their informational or computational structure in a network rather than with dispositions to behave.

Despite popular opinion that the behaviorist paradigm is dead and buried, behaviorism continues to be a profitable paradigm in psychology. This is evidenced by the number of journals devoted to it, the number of researchers and theorists who identify with it, and the considerable research grants that continue to be given to behaviorists investigating such things as: the management and treatment of autism, addiction, and intellectual handicap, the discriminations that animals are able to make, how hard animals are prepared to work for various commodities (e.g., access to scratching litter). We have seen how there are limitations to the observations that the paradigm was suitable to explain, however. The development of alternative paradigms resulted in theorists and researchers reorienting their attention

towards explaining phenomena that were anomalous for the behaviorist framework. Thus what once was the predominant paradigm for psychology came to be but one of many.

Cognitive neuropsychologists often claim that cognitive neuropsychology provides a paradigm that successfully integrates the relatively high level functional, informational, or representational mechanisms<sup>1</sup> of cognitive psychology and the relatively low level neurophysiological mechanisms of persons. One proceeds by ‘popping the hood’ (looking inside the box at the neurophysiology) which is thought to provide evidence to support or disconfirm different cognitive psychological theories of the mechanisms involved in mental processing. It is important to note that in order for neurophysiological mechanisms to be relevant for supporting or falsifying theories about cognitive psychological mechanisms they must (in the first instance) be found to be correlated, however. Insofar as computationally or representationally defined mechanisms fail to correlate<sup>2</sup> with physiologically defined mechanisms there is no integrated science of cognitive neuropsychology.

The point of the software / hardware distinction isn’t to deny that software (informational or computational) mechanisms must be realized or implemented on some (neurophysiological or silicone) hardware mechanisms or other. Rather, the point of the distinction is to show that what counts as a kind of mechanism at the software level (e.g., a ‘save function’ or a ‘pain’) may be massively multiply realized at the hardware level (e.g., neurological differences within an individual over time, between different individuals, between different species, between persons and computers...). Such massive multiple realizability is thought to undermine the plausibility of the lower level science attempting to identify mental mechanisms with kinds of hardware mechanisms (to provide the categorical basis of the suitably enriched behavioral disposition). Insofar as the kinds of mechanisms posited by cognitive psychology fail to be correlated with the kinds of mechanisms posited by neurophysiology it is hard to see how findings pertaining to one can be used to either support or falsify hypotheses pertaining to the other. In the absence of one-one type mappings neurophysiology seems as irrelevant for explaining cognitive mechanisms and functions as particle physics is irrelevant for explaining laws in economics such as Fisher’s law. Notice that this is not to say that Fisher’s law could be instantiated without there being some distribution of physical particles or other. It is simply to deny that the language (and ontology of mechanisms) of particle physics is relevant for capturing the counter-factuals (enriched behavioral dispositions) that are supported from the language (and ontology of mechanisms) in economics.

The failure of correlation suggests two broadly different ways that the future of psychology could go (with there being two options within each broadly different approach): 1a: We should revise the mechanisms, entities, and processes posited by cognitive psychology such that they do correlate with neurophysiological states (e.g., by stipulating that they need to be operationalized into talk of kinds of neurophysiological mechanisms). 1b: Talk of cognitive psychological mechanisms should be eliminated from a science of psychology. 2a: We should revise the mechanisms, entities, and processes posited by neurophysiologists so as to ensure correlation (e.g., by stipulating that kinds of neurophysiological states need to be operationalized so that they are functionally or computationally defined). 2b: Talk of neurophysiological mechanisms should be eliminated from a science of psychology. Critics maintain that cognitive neuroscience recommends that we reconceptualize the science of the mind/brain as one in which mental mechanisms, processes, and structures, are either operationalized into talk of neurological states in a way that they are not in the science of cognitive psychology - or that they simply be eliminated from a science of psychology. While cognitive neuroscientists and neurophysiologists often tend to recommend the elimination of what doesn’t seem profitable for the level of analysis that they are interested in the fate of behaviorism seems to warn that redefining the subject matter won’t give lasting success as theorists become interested in answering questions (e.g., about functionally defined computational states) for which the paradigm is ill suited.

---

<sup>1</sup>‘Mechanism’ should be understood as a shorthand for ‘state’, ‘process’, ‘entity’ etc.

<sup>2</sup>One way to defend the box factory / solubility analogy would be to maintain that there are correlations to be had after all. I’d be interested in seeing them.