Response to Zuriff, G. E. (1985) *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press. Published in *Behavioral and Brain Sciences* 1986, 9, 712–714

# Average behaviorism is unedifying

It is a monumental compilation that Zuriff has delivered unto us. Never before has so densely representative a sampling of views by so many behaviorists over so broad a range of our indigenous philosophy-of-psychology concerns been collated in one document. By all rights, this should be the definitive sourcebook for exhibiting the vision, audacity, and zeal of what behaviorism once was and could yet again become.

It grieves me, therefore, to observe that the ambitious experiment of Zuriff's compositing technique is at best only modestly successful. For inasmuch as the behaviorist literature has never been an exception to Sturgeon's Law,[1] pooling abstract position summaries across multiple sources, even when these are selected for similarity, can only degrade peaks of insight and choice delicacies of conception in a regression to mediocrity. Despite the scattering of Zuriff's own astute comments, the prevailing result here is a mush wherein classic slogans retain their rote verbal familiarity even while their meanings, too, remain as elusive/incoherent as was unhappily their norm. I shall work through one example that has more importance than Zuriff's review allows one to appreciate.

Following a suggestion by Feigl in the 1945 APA Symposium on Operationism, a recurrent behaviorist argument for interposing an "intervening variable" Z between $m$ independent data variables $X_1, \ldots, X_m$ and $r$ response variables $Y_1, \ldots, Y_n$ has been that when this can be done tidily it reduces the $m \times n$ pairwise input/output relations $\{X_i \rightarrow Y_j\}(i = 1, \ldots, m;\ j = 1, \ldots n)$ to $m + n$ relations $\{X_i \rightarrow Z\}$ and $\{Z \rightarrow Y_j\}$. On pp. 64-66, Zuriff recapitulates this line of reasoning with a minor twist of his own (namely, inferring one of the supposed empirical laws from three others) without pointing out (a) that this bivariate-lawfulness model is inchoate, and (b) that under a more adequate conception of simultaneous multivariate dependencies, the intervening variable is seen to have a very different ontological status from what Zuriff, speaking for the past norm, assigns to it. (Zuriff himself recognizes fragments of the more sophisticated story when on p. 66 he cites disposition-manifesting "circumstances"; but by clinging to the received model's inappropriate formalisms he garbles the account's proper punch line.)

Letting $m = n = 2$, suppose that $X_1, X_2, Y_1, Y_2$ are certain measures of water deprivation, dietary-salt concentration, lever-pressing rate, and ad-libitum water

[1] Ace science-fiction writer Ted Sturgeon's legendary retort to the accusation that 90% of science fiction is crap was "Hell, 90% of everything is crap." I have never encountered an observation on the Human Condition that generalizes so robustly.

intake, respectively, as suggested by Zuriff. How the first two of these affect the last two in living organisms is severely conditional on many other causal antecedents, some of which (e.g., past training, species, and maturation) we can control by manipulation or selection, whereas others remain unknown. But even with all relevant background factors held constant, there is no tight dependency of lever pressing or water intake on water deprivation alone and another on dietary salt alone. Rather, there is one determination of lever pressing, and another of water intake, *jointly* by water deprivation, dietary salt and the relevant background factors; and the environment in which outputs on $Y_1$ and $Y_2$ are observed—call this variable "stimulus setting" $S$—also makes an enormous output difference. Thus the absence of the lever or of water necessitates zero lever pressing or water consumption, respectively; whereas the presence of both the lever and water will elicit extensive interference between these two response variables (strongly conditional on how that water is distributed) in blatant contradiction with Zuriff s equation (11) on p. 68. The proper empirical model, idealized as customary with linear residuals, is

$$Y_1 = \phi_1(X_1, X_2, S, B) + e_1, \quad Y_2 = \phi_2(X_1, X_2, S, B) + e_2, \tag{1a,b}$$

where $B$ comprises all relevant background factors we have managed to identify and $e_1$ and $e_2$ are residuals reflecting additional unknown behavior sources. It is important to be clear that the laws schematized in (1) are able to govern a common domain comprising all subjects of some broad kind (say, all living mammals) at all times in any environment, for example, regardless of whether levers or water dispensers are present, because they include stimulus setting as input variance.[2] How much we can learn about functions $\phi_1$ and $\phi_2$ in (1) empirically when in practice we can sample them under just a few of the copiously diversified alternatives on $S$ needn't concern us here.

If the details of functions $\phi_1$ and $\phi_2$ are suitably cooperative, we may now find that equations (1a) and (1b) cry for explanation in terms of a hypothesized intervening variable $Z$—heuristically call this "thirst" without, however, presuming anything about its relation to mentalistic appetitive experience—such that lever pressing and water intake are each due jointly to thirst, stimulus setting, and background factors $B$, but not to water deprivation or dietary salt through any causal route unmediated by thirst, under certain laws

$$Y_1 = \psi_1(Z, S, B) + e_1', \quad Y_2 = \psi_2(Z, X, B) + e_2'; \tag{2}$$

while thirst, in turn, is caused jointly by water deprivation, dietary salt, and background factors $B$ with indifference to stimulus setting, in accord with some law

[2]In contrast, when the import of Zuriff's "defining-experiment" rider on his p. 64f. equations is explicated, it can be seen that his empirical laws (1), (4), (6), and (9) have narrow disjoint domains that preclude any one of these being entailed by the others.

$$Z = \psi_0(X_1, X_2, B) + e_0. \tag{3}$$

Space restrictions prevent my detailing how we achieve inference to (2,3) in practice, except to hint that it involves our being able to predict from what subjects do on levers in press-permissive environments to the water intake of those same subjects, under unchanged values of $X_1$ and $X_2$, in $S$-settings that allow unrestricted drinking. (A procedure we do not follow here is to write down equations (1a) and (1b) and observe that their right-hand sides both embed the right-hand side of (3). Not only is that impractical, but (2) and (3) do not even reproduce (1) exactly unless the functions in (1)–(3) are linear.) We can infer (2) and (3), however, at least for selected values of $S$ and $B$, albeit, as in statistical reasoning, the inference is fallibly ampliative.

The crucial point here is that "intervening variables" in cases like this are not *invented*, as explicitly defined abstractions from data measures, to simplify empirical equations. Rather, we *discover* them by a logic of explanatory induction (see Rozeboom, 1972 that with imperfect reliability but often overwhelming conviction discloses to us the hidden sources of intercorrelated observed phenomena. When interpreting real data, working behaviorists have made such inductions intuitively, with neither supervision by a received metatheory of their logical forms nor sufficient expertise in technical philosophy-of-science to explicate that on their own. Indeed, not until rather late in the behaviorist game did its more thoughtful partisans begin to recognize that their cherished "operationally" defined concepts were no different in kind from more conspicuously theoretical terms given meaning/referents by their nomological-network roles (see Rozeboom, 1984, and additional references cited there). Meanwhile, the mid-century bivariate-lawfulness account of intervening variables so misleadingly travesties the logic of theoretic discovery that to endorse this without significant upgrading is to portray seminal issues at the cutting edge of advanced epistemic engineering as empty symbol bashing.[3]

Technicalities of multivariate lawfulness aside, Zuriff's regression-to-mediocrity emphasizing of typical past slogans on intervening variables and operational definitions regrettably reinforces the tediously repeated slander by hostile outsiders that behaviorism was dedicated to positivistic rejection of the inner organism. It cannot be denied that a few influential behaviorists, notably Skinner (e.g., 1950), the early Spence under Bergmann's tutelage (e.g., Bergmann & Spence, 1941), and H. Kendler (e.g., 1952, ardently proscribed conjectures about internal mediators. And it is also true that MacCorquodale and Meehl (1948)), in their brilliantly definitive paper on this matter, unwisely used the label "intervening variable" to

[3]If you have been indoctrinated by Popperian philosophy-of-science, you probably don't believe that any epistemically significant "logic of discovery" exists. But Popper was simply wrong in this; and cogent theorizing is something about which practicing scientists cannot afford to be romantically naïve.

distinguish logical abstractions on observables from covert factors hypothesized to explain data regularities. But both Hull and before him Tolman, who introduced the notion, were emphatically clear that *their* intervening variables were hypothesized causal mediators (see, e.g.,Hull, 1943; Tolman 1936). (Zuriff recognizes this, but buries the acknowledgment in a footnote (1985, Chapter 4, fn. 33 and 53) when its rightful place is in his text's foreground.) And although many mid-century behaviorists would have found congenial Zuriff's normative characterization of intervening variables as "summaries" of input/output correlations that are "generally conceived of as having no causal status (p. 207), that is mainly because no one had yet made clear how the ontology of an explanatory induction's conclusion transcends the data patterning which impels its inference.

The aim of all behavioristic approaches to psychological science, the essential unity behind the splendorous diversity of specifics so amply documented by Zuriff, is not to avert attention from covert sources of overt behavior (even Skinnerians grudgingly theorize) but to insist that *scientific* conclusions about these require tough-minded epistemic warrants—hard evidence, if you like—that free-spirited theory spinners and fantasizers in folk psychology consider unbearably spoilsport.[4] Behaviorism has no yen for empty organisms; rather, it recognizes the enormous gulf between our desire to comprehend the innerness of subjects and the modest reach of our commonsense ability to attain such knowledge; and it accepts responsibility for engineering reductions of this gap to whatever extent current technical epistemic competences make possible. It is precisely because our reach of credible understanding is *not* positivistically confined to observables, but has potentially unlimited scope if properly disciplined, that the behaviorist outlook is so important for continuing psychology; and this is why its repudiation by the current cognitive Zeitgeist is such a scientific disaster. (There is nothing wrong with targeting mentality for study; it's *how* this study is pursued that makes all the difference.) Despite his evident goodwill, Zuriff has done us a considerable disservice by exhibiting behaviorism mainly as a midden of past metatheoretical muddles. It would have been far more beneficial to make clear that our profession's need for the behaviorist *program*—never mind the polemical excesses and generally limited, though far from insignificant, achievements of its early implementation—is more urgent than ever.

[4] Savor this passage from Hull (1943, p. 23) "Driesch's entelechy fails as a logical construct or intervening variable *not because it is not directly observable* [my italics] . . . but because [its] general functional relationship[s to its observable causes and effects] are *both* left unspecified. This, of course, is but another way of saying that the entelechy and all similar constructs are essentially metaphysical in nature. As such they have no place in science. *Science has no use for unverifiable hypotheses* [Hull's italics]." Hull's understanding of what it takes for a conjectured mediator to have strong empirical support is somewhat ingenuous; but his intuition correctly shouts at him that this must consist somehow in the mediator's having nonarbitrarily theorized connections to hard data that pin it down as the Lilliputians did Gulliver.

## References

Bergmann, G., & Spence, K. W. (1941). Operationism and theory in psychology. *Psychological Review*, *48*, 1–14.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century.

Kendler, H. H. (1952). "What is learned?"—a theoretical blind alley. *Psychological Review*, *59*, 269–277.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*, 95–107.

Rozeboom, W. W. (1972). Scientific inference: The myth and the reality. In R. S. Brown & D. J. Brenner (Eds.), *Science, psychology, and communication: Essays honoring William Stephenson.* New York: Teachers College Press.

Rozeboom, W. W. (1984). Dispositions do explain; or, picking up the pieces after Hurricane Walter. *Annals of Theoretical Psychology*, *1*, 205–223.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193–216–484.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.