

Jesús H. Aguilar and Andrei A. Buckareff

Causing Human Actions: New Perspectives on the Causal Theory of Action.

Cambridge, Mass.: MIT Press 2010.

vii + 327 pages

\$35.00 (paper ISBN 978-0-2625-1476-7)

In *Causing Human Actions: New Perspectives on the Causal Theory of Action*, Jesús H. Aguilar and Andrei A. Buckareff have drawn together a fine collection of essays that discuss, defend, and criticize the causal theory of action (CTA) in the context of, and applying it to, a wide range of issues in the philosophy of action, including the nature of intentional action, the causal role of the agent, reasons explanation, autonomy, omissions, developmental psychology, and more. The collection consists of 17 essays (including the introduction by Aguilar and Buckareff), 15 of which were previously unpublished (essays 8 and 9 by Carolina Sartorio and Randolph Clarke were previously published in *Noûs*). Some contributors defend CTA, while others emphatically reject it; some question its scope, while others contend that it is compatible with a wide range of theories in other areas of philosophy and psychology; some consider its historical progression and interact with the main objections against and arguments for CTA, while still others seek to raise new problems. In light of this, the collection will be of interest to well versed scholars in the philosophy of action, and appropriate for newcomers to this lively area of philosophy.

The number of essays in this contribution, the complexity of their main arguments, and the limited space afforded to me render it impossible to give the essays the attention they merit. I will reserve my discussion to offering a brief narrative of the collection and conclude with some critical remarks about Aguilar and Buckareff's introductory essay.

Aguilar and Buckareff define CTA as follows:

(CTA) Any behavioral event *A* of an agent *S* is an action if and only if *S*'s *A*-ing is caused in the right way and causally explained by some appropriate nonactional mental item(s) that mediate or constitute *S*'s reasons for *A*-ing. (1)

A few comments are in order. First, it is important to note that CTA offers a reductive picture of agency, in which an agent's doing something is reduced to mental items (such as desires and beliefs) doing something. Given its reductionism, CTA has appeared to be a promising strategy to naturalists for showing how desirable features of human agency are compatible with this framework. Second, as Aguilar and Buckareff note, it is better to think of CTA as a schema, and proponents of CTA as those theorists who accept the schema but might end up filling it out in different ways. The two clauses that require filling out are, 'caused in the right way' and 'some appropriate nonactional mental item(s).' The first clause is required to rule out cases of causal deviance, i.e., cases in which an event is caused by an "appropriate mental item" and yet intuitively, due to the manner of the causation, there is no action. Proponents of CTA disagree not only over how best to rule out cases of causal deviance, but also which mental items must and/or are always among the causal antecedent of action. Classically, Donald Davidson ('Actions, Reasons, and Causes', *Journal of Philosophy* 60 (1963): 685-700) conceived of the

causal antecedents as belief-desire pairs, though many have questioned whether this is sufficient. Finally, there is the question of scope: is CTA intended to capture all agential activity? Only some subset of agential activity, such as intentional action but not free action? I will return to this question in my critical remarks below.

Essays 2–6 explore the issues of reductionism and causal deviance surrounding CTA. In essay 2, Michael S. Moore argues that CTA, when properly understood, can account for intentional omissions, withstand worries that its reductionism actually bleeds into eliminativism, and solve the problem of causal deviance. Essays 3 and 4 are an exchange between Michael Smith and Jennifer Hornsby about the tenability of CTA. Smith focuses his article on responding to a series of criticisms leveled by Hornsby in her earlier work ('Agency and Actions', in *Agency and Action*, eds. H. Steward and J. Hyman, Cambridge: Cambridge University Press, 2004), focusing most of his attention on responding to her contention that CTA leaves out the crucial causal factor of "the agent's exercise of her capacity to do things" (52). Hornsby maintains that Smith's response does little to forestall her original worries about the inadequacies of CTA.

Essays 5–6 are, in my opinion, the brightest part of this collection and focus their attention on solving the problem of causal deviance. John Bishop's essay argues that CTA offers a way for explaining how our ethical and naturalistic perspectives are compatible. Crucial to substantiating CTA, however, is solving the problem of causal deviance, and thus Bishop focuses his attention on responding to recent objections to CTA based on the problem of causal deviance. Aguilar contends that Bishop's original defense of CTA (*Natural Agency: An Essay on the Causal Theory of Action*, Cambridge: Cambridge University Press, 1989), while making progress towards solving the problem of causal deviance, fails to handle problems posed by prosthetic agents; Aguilar then attempts to shore up the shortcomings in Bishop's account. In essay 7 Rowland Stout calls into question the view of causation underlying CTA, arguing that we need a more dynamic theory to account for agency.

Essays 8–11 consist of an exchange between Sartorio and Clarke about whether CTA can account for intentional omissions, Sartorio arguing that it cannot and Clarke arguing that it can. In essays 12 and 13, David-Hillel Ruben and Alfred R. Mele turn to issues of reasons explanation. Ruben discusses the neglected question of the causal role of con-reasons (reasons against action) on CTA, and Mele argues that CTA has decided advantages over non-causal accounts of reasons explanation.

The remaining essays consider CTA as it relates to other areas of inquiry. In essay 14, Josef Perner and Johannes Roessler consider challenges to CTA stemming from developmental psychology, and construct a "teleological account" of children's conception of intentional explanation that both fits the experimental data and is consistent with CTA. Fred Adams in essay 15 considers worries that CTA is incompatible with embodied cognition and argues that these worries are misplaced. In essay 16, Alicia Juarrero argues that an appeal to dynamical systems allows for considerable headway in areas that have exasperated philosophers of action for over two millennia, and in the concluding essay, Thomas Nadelhoffer contends that the well-known "Knobe effect" raises substantive worries for CTA, as it suggests that part of the defining feature of action is the extra-mental fact of the action's moral valence.

Although this is hardly the place to engage these wonderful essays critically in the manner they deserve, I want to raise some worries about how Aguilar and Buckareff formulate CTA. My worries begin with a question: of what range of phenomena is CTA supposed to be an analysis? As Aguilar and Buckareff define the theory (but see Moore and Smith who follow suit), CTA is confined to accounting for *bodily* action, ignoring mental action such as decision, attention, and deliberation. There is also a question of how action is being understood. Is ‘action’ merely used to refer to intentional action, in the sense that an agent can perform an intentional action without acting freely? Or is ‘action’ used in a more technical manner: does it refer to the kind of agential activity that is always free? This lack of precision leads to a lack of clarity concerning how to evaluate CTA. First, consider Bishop’s essay in which he contends that CTA promises to reconcile the ethical perspective, in which we hold one another morally responsible, and the naturalistic perspective, in which human action is amenable to scientific explanation. Given Aguilar and Buckareff’s formulation of CTA, it is clear that CTA *cannot* affect this reconciliation, perhaps doubly so. After all, some of the most important forms of agency can occur without behavior (such as focusing our attention), yet CTA only offers an analysis of bodily action. Moreover, if CTA is limited to intentional action, and cannot account for free action, then it is doubtful that CTA can reconcile the ethical and naturalistic perspectives.

Issues also become blurred in Aguilar and Buckareff’s discussion of “the problem of the absent agent” (12). According to this problem, given that on CTA the agent’s causal role in action is exhausted by the causal role of mental states and events involving him, it seems that the agent is simply a passive bystander in the process of action, and thus CTA fails to accord the agent the required active role in action production. Aguilar and Buckareff note that J. David Velleman (‘What Happens When Someone Acts?’, *Mind* 101 [1992]: 461–481) has offered the “most extensive analysis of this problem” (12) and they discuss Velleman and Michel Bratman’s strategies for solving it (Velleman and Bratman, ‘Reflection, Planning, and Temporally Extended Agency’, *Philosophical Review* 109 [2000]: 35–61). However, Velleman’s objection has traction against CTA only if we understand the theory as attempting to analyze more than merely intentional (and especially bodily) action. Velleman never argues that CTA fails as an analysis of intentional action, only that it fails, in its present form, to account for “human action *par excellence*.” The same is true of Bratman.

This lack of precision in formulating CTA unfortunately pervades many of the otherwise excellent contributions in this book. I urge subsequent writers to define more clearly the aims and scope of CTA.

Christopher Franklin
Marymount University