

## Intention in Mechanisms and the Baconian Criticism: Is the Modern Cognitivist Reviving Aristotelian Excesses?

Joseph F. Rychlak

*Loyola University of Chicago*

The Baconian Criticism holds that it is unnecessary to use final-cause conceptions when an explanation in terms of the other Aristotelian causes is sufficient to the task at hand. It is argued that modern efforts by cognitive psychologists to explain intentionality in machine terminology falls prey to the Baconian Criticism. Cognitive theory is framed extraspectively and relies basically and thoroughly on material/efficient-causation. Introducing final-cause description to such machine processing is superfluous because it adds nothing to our basic understanding of what is taking place. Telosponsivity, on the other hand, is exclusively introspective in formulation and is not open to the Baconian Criticism because of its basic reliance on oppositionality in cognition. The telosponding person is always "taking a position" within a sea of opposite possibilities, which allows for the fact that behavior could have unfolded differently all circumstances remaining the same. This permits a truly teleological understanding of human behavior, one that is not reducible to machine processing.

Hibbard's (1993) thoughtful comments on my paper (Rychlak, 1993, this issue) prompt a reply in which I return to a point that I merely alluded to initially as the criticism that Sir Francis Bacon made of Aristotle's use of the four causes. Bacon's comprehensive and sometimes quite personal attack on Aristotle has been well documented (see Farrington, 1949). But the specific point I am referring to has great relevance for our discussion of behavior as framed now by the meaning of the four Aristotelian causes. I find it helpful to keep these four causal meanings central to discussions of the sort we are now engaged in. Bacon (1952) divided natural philosophy into *physic* and *metaphysic*, stipulating that the former realm deals with what is inherent and transitory in matter whereas the latter deals with what is abstracted and fixed (p. 45). Description in the physical realm is limited to "being and moving,"

whereas in the metaphysical realm we have issues of “reason, understanding, and platform [i.e., grounds]” under consideration (*ibid.*, p. 43). Since natural philosophy is also to be thought of as dealing in causes Bacon added that: “The one part, which is phisic, inquireth and handleth the material and efficient causes; and the other, which is metaphysic, handleth the formal and final causes” (*ibid.*, p. 43).

How did this distinction detract from Aristotelian philosophy? Well, Aristotle never suggested that we limit the use of the four causal meanings like this. The more causes we could use to explain anything the better. From his perspective, “Nature belongs to the class of causes which act for the sake of something” (Aristotle, *Physics*; 1952a, p. 275). This means that a description of physical events could well include a final-cause conceptualization (recall that the definition of final cause is “that [reason, etc.], for the sake of which” events occur or objects exist). It therefore was plausible to Aristotle that plants grow leaves for the sake of shading the fruit on their branches (*ibid.*, p. 276).

Bacon found such final-cause description in science totally unacceptable. Farrington (1949) notes that Bacon considered Aristotelianism to have succumbed to the “sin” of intellectual pride in presuming to say “why” leaves are on the branches of trees (p. 148). And even if one wished to speculate on such matters, the realm of such investigation is not the physical [or “scientific”] but the metaphysical–theological. Thus it is that we find Bacon (1952) opining: “For to say that . . . the leaves of trees are for protecting of the fruit [on their branches] . . . is well inquired and collected in metaphysic, but in phisic [it is] impertinent . . . [T]he search of the physical causes hath been neglected and passed in silence” (p. 45). By “physical causes” Bacon of course means material and efficient causation.

The upshot here is that according to Bacon’s criticism, final-cause description in the physical realm (a) adds nothing relevant to its understanding, and (b) overlooks its relevant physical causes. The Baconian Criticism was to become extremely influential in the rise of Newtonian science, where the goal was to describe “how” natural events or products came about and refrain from speculating on “why” they occurred or existed, which is totally irrelevant to an understanding of natural objects and events in any case. There is also a deity teleology under attack here by the Newtonians, prompted by the repressive measures of the Inquisition against such heroes of science as Galileo. To speak of ends, reasons, or intentions in physical description could only mean that somewhere in the account a deity was lurking to direct things, including the expectation that science would defend Biblical teachings like the geocentric structure of the universe.

Teleology of all sorts — including the strictly human variety that I support — received a mortal blow in the rise of Newtonian science. To give just one

important example in psychology, when Titchener (1898) drew the distinction between a functional and a structural psychology he relied on the Baconian Criticism as follows: "There is . . . the danger that, if function is studied before structure has been fully elucidated, the student may fall into that acceptance of teleological explanation which is fatal to scientific advance" (p. 453). Titchener believed that functionalist psychologists like Brentano and James too quickly settled for accounts of behavior relying on intentionality, and thereby overlooked underlying structural components falling within what Bacon would consider the physical realm. In other words, the functionalists settled for final causation when material and efficient causation was called for. To employ intentionality in the early years of psychology was to theorize in the vein of Voltaire's (1930) satirical Pangloss, the "metaphysico-theologo-cosmo-nigologist," who taught his young charge, *Candide*, that: "everything is made for an end . . . noses were made to wear spectacles; and so we have spectacles" (p. 14).

There is obviously merit in the Baconian Criticism, particularly as it applies to explanations in the bio-physical realm, which always take what I have called an extraspective or third-person perspective (Rychlak, 1981, pp. 27-34). This theoretical slant invites explanation in a mechanistic (i.e., exclusively material- and efficient-cause) fashion. Much to my satisfaction, Hibbard (1993) recognizes the importance of framing psychological explanations in an introspective or first-person fashion, which means we would be trying to understand things from the person's unique, "internal" conceptualization of what is transpiring experientially rather than in terms of what is supposedly shaping him or her "from without." I think the widely cited Jones and Nisbett (1971) study of actors versus observers supports this introspective/ extraspective distinction. This study examined what actors and observers believed were the relevant causes of the behavior carried out by the actors in various situations. The actors claimed that it was the logically relevant factors (*formal/final causation*) in the circumstances facing them that caused them to behave as they did, whereas the observers of these same actors were prone to attribute the cause of the latter's behavior to presumed habits (*efficient causes*) or personality traits (*material causes*) that directed the actor no matter how the logic of the situation was framed introspectively.

My concept of the telosponse is an exclusively introspective formulation, relying upon the affirmation of meanings that are sequaciously extended into ongoing experience by such actors. Biological or socio-cultural "influences" must first be predicated (i.e., grasped either accurately or erroneously), following which the person behaves "for their sake" rather than "in response to" their unpredicated impact. The fact that I base telosponsivity on an intrinsic oppositionality in human cognition is what distinguishes my approach from the "system teleologies" that Hibbard (1993) is supporting. The reason affir-

mation is called for in a telosponding organism is because the person always confronts experiential alternatives by way of oppositional possibilities — possibilities that need not exist because they are not “in” some informational input, not patterned meaningfully “in” some unpredicated, independent reality! The person can reason from what is the case (“in” reality) to what is not the case (“in” imagination, etc.) and frame an alternative that is not a matter of “settling into” solutions based upon weighting preformed (i.e., already predicated) inputs according to an unchallengeable algorithmic rule (to refer to an example Hibbard gives us).

Though he sees the necessity of introspective formulations, the systems that Hibbard presents as examples of intentional processes are totally extra-spective formulations which do not “take a position” on anything hence cannot question their executive algorithms, never draw an implicit implication to the opposite of the goal that they are “pursuing,” and are intrinsically incapable of negating what Hibbard refers to as their “acting rationally in a goal directed manner.” Indeed, by his own definition, an intentional system cannot pursue goals in an irrational fashion.

As I rely upon oppositionality in my formulations I am not exactly a “rationalist” in theoretical persuasion. This label is typically used to describe a theorist who places reason above all other factors in the explanation of behavior. And reason per se is said to involve a capacity to draw logical conclusions — which comes down to “correct” or “sound” conclusions. Though telosponsivity surely involves drawing logical conclusions, it is much more than this for it includes the pre-affirmation process of predication, of bringing to bear a precedent (*formal cause*) pattern “for the sake of which” (*final cause*) the meaning under affirmation is understood and then enacted — or, thanks to the generic oppositionality of cognition, *not* enacted! This process makes no claims about extending only correct or sound predications into sound conclusions. Its intentionality stems from the oppositionality at its base, enabling the position to be taken or — all things held constant — negated. If the person reasons from a sound (realistic, etc.) premise to an unsound (unrealistic, etc.), illogical — hence, “irrational”— premise, and then extends such irrational meanings into overt behavior, this does not in any way contradict or subvert the intentionality implicit in the telosponsive process. I do not agree with Hibbard that intentionality has to be either rational or goal directed — if we mean by “goal” some definite objective. People are frequently irrational and they sometimes intend a goal-less, even dangerously spontaneous course of action.

Coming down to my central point in this reply to Hibbard (1993), it is my charge that he as all modern cognitivists who are attempting to “account for” intentionality by way of systems theory, computer modeling, or information processing leaves himself open to the Baconian Criticism. Cognitive theo-

rists who seek to account for intentionality without altering the mechanism on which their views are positioned necessarily return to the excesses of Aristotle when they now seek to “tack on” final causation to what is sufficiently accounted for based on material and efficient causation. Paraphrasing Bacon’s “leaves shading fruit” example, we can ask of them: what does it add to our understanding of how a system, computer, or thermostat “works” to say that such a process has an intention? We understand and can explain this mechanical process completely without bringing in such irrelevancies. Telosponsivity, on the other hand, is a process involving the framing of a predication within a sea of opposite meanings and extending such meanings in formal-final cause fashion. Telosponsivity cannot be explained through use of well-understood, mechanical conceptions. Its description of intentionality cannot therefore be considered superfluous.

Hibbard has me criticizing a now passé hegemony of stimulus-response psychology. He apparently believes that we have moved on to a new era, in which cognitive psychology and teleological behaviorism are meeting the criticisms that I have been leveling at psychological explanations for over 30 years to date. My reaction to his opinion is that it does not matter what happened to behaviorism’s dominance in the field because the stimulus-response concept is not the historic culprit in any case. The culprit is the hegemony of efficient causation, which I contend has not changed one iota in the so-called cognitive revolution of recent years (see Rychlak, 1991). Today we have input-output instead of stimulus-response, and although the feedback conception does offer an advance in complexity, the fundamental cast of such theorizing continues to be efficiently causal, extraspective, and linear (i.e., lacking in oppositionality). The lyrics have changed but the mechanical melody lingers on.

Turning to specific examples of intentionality offered by Hibbard (1993), I believe that I have shown how Tolman’s “purposive behaviorism” was not the same type of purpose that people like William McDougall were referring to when they spoke of a person’s purpose (see Rychlak, 1988, pp. 150–154). Purpose is an introspective concept if ever there was one. Yet Tolman’s (1932/1967) concept of purpose was framed extraspectively. He held that a response that reflected docility (i.e., teachableness) in relation to some goal was purposeful, so that as we observe a rat’s growing facility across trials in reaching the goal box of a maze we are literally *seeing* purpose take place (ibid., p. 14). The problem here is that if some (stupid?) rat failed to improve its performance we would presumably not be witnessing an animal moved by purpose (no docility). Now we know that human beings work for goals that do not yet exist (like the rat’s goal box always exists at the end of a run)— in fact, cannot possibly exist (e.g., “perfection”) — and yet in their intentional efforts people foul up, again and again, showing no docility to speak of. Even

so, who would deny that they are behaving purposively, "for the sake of" such unachievable goals? I would like quickly to add that these goals need not exist in an independent reality, or "down the road of time," drawing the person forward by the "suction" of some kind of reverse efficient causation; the goals (ends, reasons, purposes, etc.) occur strictly and totally in the meaningful predications framed by the telosponding organism in the immediate present. I have shown in my writings that even certain lower animals exhibit such telosponsivity (Rychlak, 1992, Chapter 7).

Another example Hibbard (1993) gives us is Bandura's modeling theory. I find that a careful reading of Bandura's (1986) theory establishes that although he frequently sounds as though he is accepting teleology, when it comes down to specifics he is still a conventional mediational theorist. The problem is that he lacks a concept like telosponsivity on which to base behavioral processing. I commented on this once in print, in reaction to what he was saying about self-reflective thought (see Rychlak, 1979), only to have him lecture me on the benefits of mechanism as follows: "There is a difference between analyzing cognition as a contributing factor in the reciprocal determination of events and conceptualizing cognition as a psychic agent that orchestrates behavior. Understanding of how people exert some influence over their actions is more likely to be advanced by delineating and exploring the nature of self-regulatory mechanisms than by simply ascribing behavior to a psychic agent" (Bandura, 1979, p. 440).

This statement is actually a form of the Baconian Criticism, for Bandura obviously believes that to speak of a telosponding organism is to concoct psychic agents or homunculi instead of looking for the underlying material- and efficient-cause mechanisms. I, on the other hand, believe that the only reason we may require a homunculus in a theory is because we have already turned human beings into machines. Machines need decision makers, and so we have the little person from within "driving" the machine about. But is it not obvious that the homunculus has to be doing something like telosponding in order to make the decisions about what to believe, where to go, and what to do? Logical learning theory shortcuts this clumsy theoretical maneuver by making the human being out to be a telosponder *per se*. No little person from within need drive this organism about because the organism is driving itself from the outset.

I do not find Newell and Simon's theorizing to be in the least bit teleological. To say, as Hibbard (1993, p. 378) does, that the mechanism's "maneuvering within a problem space" is "for the sake of" reaching a goal is simply to tack on telic phraseology where it is not required in order to account for the (extraspectively conceived) process taking place. No mechanism ever works "for the sake of" a goal, because it never takes a position (via affirmation) on this goal to begin with. The "goal" of a machine is a fait accompli, a "done

deal" from its processing inception, even if this "goal" is to create new (programming) goals. Simply because such mediational modelers as Newell and Simon (1972) refer to goals under processing does not mean that they are accounting for a teleological course of behavior. Logical learning theory locates its teleology in the *process* of telosponsivity, not in the *contents* of this process. Whatever meanings the telosponsive process may be carrying forward, whether concerning a sentimental recall of some past life event, the contemplation of a beautiful painting in the present, or the projection of some desired goal in the future, *all* such mental contents are cognized by the *same* teleological process.

On the other hand, mediational modelers find their teleology in the content being mechanically processed — so that one item being moved along by efficient causation is said to be a goal while another is not. Newell and Simon frankly admitted that they focused on the mediating "internal, symbolic mechanisms" (*ibid.*, p. 4) rather than on how people learn to behave differently, select alternative goals, and so on (*ibid.*, pp. 556; 866). Since they rest everything on an efficient-cause process I fail to see how Newell and Simon avoid the Baconian Criticism. Goals and intentions are superfluous attributions, assigned to contents being carried along by the efficient-cause (non-telic) process. It is as if the mediation modeler says "color this mediating content a goal" versus "color that mediating content a non-goal," when the underlying process is the same, non-intentional, efficiently caused mechanism for both goal and non-goal. Can we really explain teleology by reducing it to mechanism in this way?

Motion is vital to the mechanistic account, because this is the *sine qua non* of efficient causation. In telosponsivity, the precedent-sequacious extension of meaning occurs instantaneously—literally outside of time considerations. It is not the passage of antecedents to consequents over time that matters, but the logical ordering or patterning of a wider realm of meaning extending to a more focused, targeted realm of meaning. Hibbard (1993) speaks of the "intentional stance." The word "intention" devolves from Latin, meaning "to stretch out." I view sequacious meaning-extension as such a "stretching" from a wide-ranging, predicating context of meaning to a narrower, targeted realm of meaning — all taking place immediately, completely outside of time.

Hibbard (1993, p. 380) asks what I think of the teleological behaviorism advocated by Rachlin (1992). Well, unfortunately, not too much. I would like first of all to note that I have suggested in many of my writings that Skinnerian theory has capitalized on the telic nature of human beings without giving clear credit to the actual (i.e., final-) causation taking place in human behavioral processing (for a recent example, see Rychlak, 1992). In October of 1989 I had the honor of discussing this issue briefly with Dr.

Skinner. To my surprise, he was not bothered when I suggested that he may be propounding a teleology. It was clear to me, however, that he was construing teleology extraspectively — as an observable consequence (contingency) following an observable action (operant) in the style of Tolman. This is readily understood in non-telic fashion, of course, because the person has no intrinsic capacity to negate (through oppositionality) what is shaping his or her behavior over time — which to my way of thinking means there is no real “that, for the sake of which” occurring in the first place.

I become very uneasy when Skinner and the neo-Skinnerians like Rachlin begin accepting the label of being teleological theorists. What is my problem? Well, in line with what I said above concerning the lack of a true change in the efficient-cause style of theorizing by today’s cognitivists, it has been my experience that when mechanists (behaviorists, information processors, etc.) begin using telic language they change meanings to suit themselves. You cannot subsume the meaning of final causation by efficient causation and convey the same meaning that would be expressed if final causation were understood in its own right. To give an example from the teleological behaviorist, Rachlin (1992), when he defines the final cause for us he is anxious to include the concept of motion, as follows: “A final cause is a form of movement (or classification of movement) abstractly conceived; its effect is a particular movement” (pp. 1371–1372).

I cannot accept this interpretation of final causation based on my reading of Aristotle, whose entire discussion of the primary or unmoved mover, as well as the concept of universals, hinges upon the view that a final cause need not move (i.e., involve local motion) [see Aristotle, 1952a, *Physics*, Book VIII, Chapters 5–7; pp. 341–348]. Indeed, Aristotle (1952b) explicitly states that “the final cause may exist among unchangeable entitles” (*Metaphysics*, Book XII, Chapter 7; p. 602). Ends (e.g., goals, ideals, hopes) provide motives for motion, but they need not themselves move or change in any way over time.

Rachlin (1992) has now defined final causation to suit his behavioristic tradition of motility, lumping it together with efficient causation as the supposedly “dynamic” (ibid., p. 1371) descriptors of Aristotle’s schematic. He proceeds to build an extraspective characterization of behavior on this dubious foundation which is just as subject to the Baconian Criticism as the other examples that Hibbard (1993) advances. Actually, it is possible to see both an introspective and an extraspective formulation of final causation in the writings of Aristotle (1952c, *On the Parts of Animals*, Book I, Chapter 1; pp. 161–162). That is, he not only described leaves as purposively shading fruit, he spoke of the physician as taking the health of a patient as the “that [end, reason, goal], for the sake of which” therapeutic efforts were carried out. Note that this formulation can be understood introspectively, from the physician’s assumptive beginnings in setting about to diagnose a patient.



Note also that this "end" of health is not in motion; it is an ideal toward which therapeutic speculations and efforts can be aimed. And thanks to the oppositionality of the physician's capacity to reason dialectically, this end could also be negated! The efficient-cause efforts to cure which the physician carries out as local motions (locomotions) are purely instrumental behaviors, intended or aimed at this motionless ideal (which may or may not be realized). This is the kind of teleology I am trying to capture, and I contend that it is not possible to achieve unless we construe behavior according to a non-mechanistic process, a telosponsive process in which people are seen to affirm premises (encompassing predications) for the sake of which they behave *in all that they do*. As a descriptive label, teleological behaviorism is an oxymoron. Once we have a teleology we drop the strictly efficient-cause mechanisms of behaviorism. Put another way, we add to these mechanistic descriptions a richer account of behavior. Since the telosponder is shown to make a difference in what transpires behaviorally, over and above any mechanistic formulations of that behavior, the LLT advocate is not subject to the Baconian Criticism. I close by reaffirming the assertion of my original paper: to capture the essence of human behavior, we need a concept like telosponsivity in psychology.

### References

- Aristotle. (1952a). *Physics*. In R.M. Hutchins (Ed.), *Great books of the western world* (Volume 8, pp. 257–355). Chicago: Encyclopedia Britannica.
- Aristotle. (1952b). *Metaphysics*. In R.M. Hutchins (Ed.), *Great books of the western world* (Volume 8, pp. 495–626). Chicago: Encyclopedia Britannica.
- Aristotle (1952c). *On the parts of animals*. In R.M. Hutchins (Ed.), *Great books of the western world* (Volume 9, pp. 161–229). Chicago: Encyclopedia Britannica.
- Bacon, F. (1952). *Advancement of learning*. In R.M. Hutchins (Ed.), *Great books of the western world* (Volume 30, pp. 1–101). Chicago: Encyclopedia Britannica.
- Bandura, A. (1979). Self-referent mechanisms in social learning theory. [Reply to Rychlak]. *American Psychologist*, 34, 439–442.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Farrington, B. (1949). *Francis Bacon: Philosopher of industrial science*. New York: Henry Schuman.
- Hibbard, S. (1993). Behavior, adaptation, and intentionality: Comments on Rychlak, Leahey, and Jenkins. *The Journal of Mind and Behavior*, 14, 373–384.
- Jones, E.E., and Nisbett, R.E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior*. Morristown, New Jersey: General Learning Press.
- Newell, A., and Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Rachlin, H. (1992). Teleological behaviorism. *American Psychologist*, 47, 1371–1382.
- Rychlak, J.F. (1979). A nontelic teleology? [Comment on Bandura]. *American Psychologist*, 34, 435–438.
- Rychlak, J.F. (1981). *A philosophy of science for personality theory* (second edition). Malabar, Florida: Krieger.
- Rychlak, J.F. (1988). *The psychology of rigorous humanism* (second edition). New York: New York University Press.

- Rychlak, J.F. (1991). *Artificial intelligence and human reason: A teleological critique*. New York: Columbia University Press.
- Rychlak, J.F. (1992). A teleologist's reactions to "On private events and theoretical terms." *The Journal of Mind and Behavior*, 13, 347-357.
- Rychlak, J.F. (1993). Behavior as telosponsivity rather than responsivity. *The Journal of Mind and Behavior*, 14, 365-372.
- Titchener, E.B. (1898). The postulates of a structural psychology. *Philosophical Review*, 7, 449-465.
- Tolman, E.C. (1967). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts. (Originally published 1932)
- Voltaire, F.M.A. (1930). *Candide*. New York: J.J. Little & Ives.