

How Predictable Are We?

CARL BROWNSON

THIS paper will address the question of how it is, exactly, that we are able to predict the behaviors of our fellow human beings. More particularly, it will address an answer given to this question by Daniel Dennett. According to Dennett, we can predict the behavior of our fellow human beings quite effectively using the following mechanism:

First you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning will in many—but not all—instances yield a decision about what the agent ought to do; that is what you predict the agent *will* do.¹

Dennett calls this strategy the Intentional Stance. I will argue that he grounds the Intentional Stance on assumptions of human optimality, assumptions that according to current thought in evolutionary biology, cannot be defended. Dennett's strategy would be more plausible were he to reestablish his strategy based on empathy rather than optimality.

So, what gives us the “many *but not all*” clause in the above statement about this predictive mechanism? Dennett says, “The objective fact is that the intentional strategy *works as well as it does*, which is not perfectly. No

Carl Brownson is senior studying philosophy at the University of Tennessee. He will pursue a Ph.D. at the City University of New York this fall.

one is perfectly rational, perfectly unforgetful, all-observant, or invulnerable to fatigue, malfunction, or design imperfection. This leads inevitably to circumstances beyond the power of the Intentional Strategy to describe.”² If this strategy works in the manner Dennett claims it does, it would be of interest to find out how often and in what capacity those imperfections in rationality, design, memory, et al., come into play. The usefulness of this predictive mechanism would be proportional in some way to the avoidance of these problems. If this predictive mechanism fails to escape these problems often enough, the usefulness of this strategy will quickly become marginal. Dennett’s strategy requires individuals to achieve relative optimality in the above-mentioned categories (rationality, et al.) for his predictive strategy to be of use in predicting human behavior. So why does Dennett think the intentional strategy works “as well as it does”?

The first answer to [this] question of why the intentional strategy works is that evolution has designed human beings to be rational, to believe what they ought to believe and want what they ought to want. The fact that we are products of a long and demanding evolutionary process *guarantees* that using the intentional strategy on us is a safe bet.³

This evolutionary answer, however, illustrates a critical weakness in Dennett’s answer, a weakness that he does not seem to acknowledge. This weakness, according to contemporary sentiment in evolutionary biology, undermines the relative effectiveness of Dennett’s predictive strategy. According to current trends in evolutionary biology, the assumption that evolution creates optimal systems is not a secure one.

I. Evolutionary Biology and the Assumption of Optimality

Certain strands of this contemporary thought can be seen in the work of Steven Gould and Richard Lewontin. Gould and Lewontin relay an anecdote given by a colleague of Herbert Spencer’s on the origin of fingerprints:

¹ Dennett 17.

HOW PREDICTABLE

It was obvious, he said, that the delicate mouths of the sudorific glands required the protection given to them by the ridges on either side of them, and there from he elaborated a consistent and ingenious hypothesis at great length. I replied that his arguments were beautiful and deserved to be true, but it happened that the mouths of the ducts did not run in the valleys between the crests, but along the crests of the ridges themselves.⁴

This anecdote illustrates that although an adaptive story can be told to explain why a particular feature of an organism exists as it does, it does not follow that the story told is necessarily true. Recently, much has been done to dispel the idea that evolution can be explained simply by telling stories about an all-powerful process of adaptation through natural selection, through which all traits are given their form. We know that other factors *do* come into play in evolutionary processes, some of which do not concern natural selection at all.

The dismissal of the idea that all evolutionary processes can be explained through the fabrication of stories about an all-powerful selective process leads us to a second, related point, a point upon which much energy has already been spent. Evolutionary processes do *not* necessarily work to optimize particular traits. This is the case for three reasons:

- 1) Evolutionary processes are not always powerful enough.
- 2) Evolutionary processes may not work on a particular trait at all.
- 3) Even when evolutionary processes *do* work on a particular trait, they may do so through some mechanism unrelated to natural selection (which is just a particular mechanism of evolution). These other mechanisms will provide no guarantee of optimality.

In response to these claims, Dennett adopts and defends the view generally called adaptationism. Dennett defends this stance in his critical essay in *The Intentional Stance*, "Intentional Systems in Cognitive Ethology: The Panglossian Paradigm Defended." Dennett's defense of adaptationism rests on the concept that adaptive explanations *are* necessary for particular traits, because they cannot, in non-adaptive terms, answer important questions regarding the purpose of various traits and characteristics. His defense will be examined here.

Various significant problems arise with adaptationism, particularly in regard to the features of the human mind. One major problem, as suggested above, is that there is empirical evidence that many of the features of the human mind are simply *not adaptations*. The ability to read is a prime example. Because there was never an episode in the history of humanity when the survival of the species depended on its members' ability to read, reading is not an adaptation. We have brains that developed for other purposes, and only in the last few thousand years were those brains applied to the interpretation of written symbols. This type of function (though they never explicitly mention reading) is what Gould and Lewontin refer to as a 'spandrel'. Spandrels are those features which are "secondary utilization of parts already present for reasons of architecture, development, or history."⁵ Nothing can be directly claimed in evolutionary terms about such features' optimality, since these features are not present solely due to their capacity to promote survival. They are essentially present accidentally.

Further, even those features that are historically likely to be adaptations through natural selection can quite often fail to be *optimal* features in their present utility. For example, humans generally crave fats and sugars. It is very likely that at some point in our history, these cravings were a positive adaptation, as those humans who stored fats and sugars were more likely to survive winters without central heating than those that did not. Therefore, these people would have reproduced and become a majority of the population. Now, however, cravings of (widely available) fats and sugars remain a feature of many humans, in a manner that is far from optimal. Those cravings are, in fact, a leading cause of early death in contemporary (Western) society.

As seen above, certain traits and behaviors may admit or reject adaptive explanations. Explanations concerning the nature of our brains often lead to the same conclusions when subsumed under the adaptationist theory. It seems almost sure that the development of our large brains themselves was an adaptation. We have *no* idea, though, how they became this

² Dennett 28.

³ Dennett 33, emphasis added.

⁴ Gould and Lewontin 88.

⁵ Gould and Lewontin 85. They are named 'spandrels' after a certain form of architectural "left-

HOW PREDICTABLE

large and complex. Similarly, it is likely that particular basic functions such as language and emotional processes were adaptations, but determining exactly how this came about is speculative work at best.⁶ Many other functions are even more controversial in origin, and if these likely “spandrels” are *not* adaptations, then speaking about them in terms of optimality seems impossible, at least from an evolutionary standpoint, and the evolutionary standpoint *is* the standpoint Dennett takes.

Dennett’s particular defense of adaptationism (or, as he calls it in the title of the relevant chapter in *The Intentional Stance*, “the Panglossian Paradigm”)⁷ rests on the idea that adaptationism can provide explanations regarding the features of organisms that a non-adaptationist approach cannot. As he puts it, “One has not yet answered the *why* question posed when one has abstemiously set out the long (and in fact largely inaccessible) history of mutation, predation, reproduction, selection—with no adaptationist gloss. Without the adaptationist gloss, we won’t *know why*.”⁸ This claim follows an earlier charge of hypocrisy leveled against Gould and Lewontin in which he suggests that they actually are espousing a “recommendation that we should all be more careful and pluralistic adaptationists.”⁹ Dennett implies that Gould and Lewontin will have to ultimately reject their “adaptationist view” if their positions are to remain coherent.

In response to Dennett’s claim, Gould and Lewontin refer to the set of unsubstantiated adaptive reasons (the “why” questions about traits) as “Telling Stories.” Telling stories is problematic for two reasons. First, the method is unfalsifiable: “Since the range of adaptive stories is as wide as our minds are fertile, new stories can always be postulated.”¹⁰ Second, stories are often accepted for less than adequate reasons. Plausibility, consistency, and coherence to a pre-established set of beliefs are often enough to merit the acceptance of an adaptive story. If a non-adaptationist approach cannot answer the “why” questions we want

over,” in which the primary features of a certain dome leave spaces, or “spandrels.” The particularly imaginative artwork on these spandrels might lead the viewer to believe that the spaces were created explicitly for the artwork. In fact, the spaces are first and foremost an architectural byproduct, regardless of their later use.

⁶ Sterelny 328–32.

⁷ Gould and Lewontin first call adaptationism the Panglossian Paradigm, after Voltaire’s Dr.

answered, then perhaps those questions are just out of our reach at the present time. A non-adaptationist evolutionary pluralist can have plenty of “why” answers, where there are actual good, adaptive, “why” answers; however, if there are no such answers to be found, it is prudent to accept that and avoid “story-telling”.

Lacking an all-powerful mechanism of adaptation through a natural selection that acts on every identifiable trait, we lack the optimality that Dennett’s predictive strategy needs. Based on adaptationism, Dennett’s strategy requires a mechanism that ensures the optimal development of various traits and features. Working sometimes, on some traits, natural selection will not provide such a mechanism. As seen above, the optimality of certain traits (reading, for example) may not be the result of natural selection. The trait(s) remaining as a result of natural selection is not necessarily optimal. Furthermore, numerous essential traits completely escape the influence of natural selection. Thus we see that as the adaptationistic ice becomes increasingly thin, Dennett’s predictive strategy begins to lose its support.

II. Suboptimal Behavior and the Intentional Stance

Dennett’s strategy is based on a third-person perspective. He assumes that “we can see more and better if we start here, now, than if we try some other track.”¹¹ From this perspective, obviously, it is going to be difficult to know exactly what people in the first-person perspective are going to do. Thus, he needs to have some mechanism to get from pure speculation about what people in that first-person perspective believe and desire to what they *actually* believe and desire. As a result, his method requires the predictor to attribute only those beliefs and desires that the agent “ought to have”; this is the only place where the predictor and the predicted have much chance of meeting on those beliefs and desires. According to Dennett, true believers are the only people with whom this meeting actually occurs. Surely people believe things suboptimally all the time. I will come back to this problem shortly.

Using this mechanism, however, the agents who possess the beliefs

HOW PREDICTABLE

they “ought to have” will obviously not meet. Dennett requires a common set of “oughts” between predictors and agents in a large number of situations for his intentional strategy to be of much use. This large common set is provided by a belief that we are optimal, and so will predict each other well enough and behave well enough of the time to *guarantee* a large number of meetings. As we have seen, however, the evolutionary mechanism on which he relies for that optimality will not necessarily provide it.

What, then, happens to Dennett’s strategy when people do behave suboptimally? Stephen Stich posed this problem to Dennett: in particular, the problem of people making mistakes. “If we ... trade up,” he says, “to the intentional-systems of notions of belief and desire ... then we simply would not be able to say all those things we need to say about ourselves and [others] when we deal with each other’s idiosyncrasies, shortcomings, and cognitive growth.”¹² Dennett’s system is a *predictive mechanism*, and so the response he should give to Stich would be a method through which *these* failures of optimality in rationality, memory, or cognitive growth could be predicted. Dennett claims, instead, that “such errors, as either *malfunctions* or the outcomes of *misdesign*, are unpredictable from the intentional stance.” He goes on to say that “there will inevitably be an instability or problematic point in the mere *description* of such lapses at the intentional system level—at the level at which it is the agent’s beliefs and desires that are attributed.”¹³ The determination of exactly what mistake is made in a case of suboptimal behavior is not easy to come by, and reasons, to name just a few, can include memory failure, irrationality, and inexperience. Thus, Dennett gives us an explanation of why such behavior was suboptimal, after the event, but does not answer how such behavior could have been predicted before the event (he thinks it could not have); and he does not give an answer, potentially as useful, as to *why* such behavior is unpredictable.

I think that the answer to the latter question can be approached using, in part, another idea that Stich proposes to counter Dennett’s intentional strategy. Dennett seems to think that there is some set of ideals that constitute optimally rational, fully in possession of memory, optimally experienced, etc., which intentional systems either approximate or

Pangloss, who was a caricature of Leibniz and his belief that this was “the best of all possible

do not approximate, and against which such intentional systems can be assessed in terms of that approximation. The question, “What ought this person do in this situation?” is asked relative to that set of ideals. However, Stich proposes that what we are actually asking in that circumstance is, “What would I do if I were in his place?” Stich claims that “the notion of idealized rationality plays *no role at all*,” and that “in ascribing content to belief states we measure others not against an idealized standard but against ourselves.”¹⁴ Dennett responds by claiming, “Measuring [someone] against ourselves is measuring against an idealized standard.”¹⁵

There are at least two reasons for dismissing this response. The first is that very little guarantee is given that the “ourselves” being measured against will not fall victim to the unpredictable mistakes that Dennett previously admitted (unavoidably) do exist. The second is that, without a powerful optimizing force (in the form of adaptationism), such mistakes may very well be rampant. However, if we admit this suboptimality into the form of our predictive strategy, we might come a step closer to being able to predict behavior. Since suboptimality is at least an occasionally present feature in the behavior of intentional systems, even if it is not rampant (which it seems, in fact, to be), possible explanations of these sub-optimal behaviors would seem to merit some attention.

III. Empathy and the Intentional Stance

By illustrating the weaknesses inherent in Dennett’s assumption of optimality—that is, the inability to assume optimality through natural selection—we can also see a possible solution to the problem. I propose that Dennett’s predictive strategy, and any predictive strategy for that matter, should be based on the assumption that individuals will not act optimally. As such, the predictive strategy would be directed from the principle of empathy rather than optimality. According to Stich, statements about what an intentional system ought to do actually take the form, “What would I do in that situation?” If we transform this question to say instead, “What would I do in that situation, if I were like him?” we

worlds.” This kind of optimality is necessary for Dennett.

HOW PREDICTABLE

⁸ Dennett 267, emphasis added.

⁹ Dennett 261.

¹⁰ Gould and Lewontin 79–81.

¹¹ Dennett 7.

¹² Stich 48.

¹³ Dennett 84.

¹⁴ Stich as quoted by Dennett 99.

¹⁵ Dennett 99.

might be able to drop assumptions of optimal behavior from the picture entirely. Dennett thinks that suboptimal behavior, although it cannot be predicted, can be described and that it usually is. However, those descriptions of the reasons for suboptimal behavior can be incorporated into the belief-desire attribution structure in appropriate ways. If I were to drive from New Orleans to Los Angeles in one day, and it was important that I made it for a certain meeting, and I was not as attentive as I usually am, etc., then I would fall asleep at the wheel. Therefore, I predict that someone in this situation will fall asleep at the wheel. There seems to be no reason, given the incorporation of this shared set of human sub-optimal characteristics, that sub-optimal behavior should be any less predictable than optimal behavior.

Dennett accepts an optimality-assuming version of an intentional strategy because it works, and it works because people usually behave in optimal ways. The practical reason, however, that the strategy works (when it does work) is not because people approximate optimal levels (as if we could even describe such levels in concrete terms). Rather, it works in those cases because the predicted behaviors *match* the behaviors performed by the predicted agents. Dennett's strategy works if the predicted people behave in optimal ways. However, even if people do not act optimally, then a predictive strategy may still work if the predictions match the behaviors of the predicted. Those predictions can be reached using a "What would I do in that situation, if I were like him?" question. The behaviors of intentional systems (without the label of optimality) can be predicted based on people's *actual* beliefs and desires. The burdensome label of "optimal" behavior can, in this case, be discarded.

Discarding the assumption of optimality would seem unavoidable; we are not optimal beings, as evolutionary theory and common sense might indicate. Furthermore, we seem to use the empathetic approach every day, in the same way that Dennett claims that we use the optimal formulation. The idea "I bet that guy's going to fall asleep on the road," or some related sub-optimal prediction, has undoubtedly taken shape at some point in everyone's mind. Taking these bad predictions with the good would require more work—attributing the likely flaws in belief and desire of particular people at particular times is another potentially complex step—but it seems to work as well as the optimal formulation. Every day, based on empathy, we predict the actions of others.

HOW PREDICTABLE

Some may continue to think that Dennett's strategy *needs* his optimality to function, as there would presumably be no starting point for predictions without it. I do not think this is the case, for the reasons stated above. Malfunctions and design imperfections are pervasive in human beings, and there seems to be a large question concerning what optimal people would even look like. Granting that we all share this set of maladies and shortcomings, some empathetic thinking should be able to give us the same starting block that Dennett seeks with his predictions of optimal behavior. The question "What would I do in that situation, if I were like him?" is useful, because to a large extent the predictor will be like the predicted, much more than either one of them *will* be like an optimal human. For this reason, the empathetic approach gives a better chance for the kind of matching between predicted behavior and actual behavior Dennett needs for his strategy to work than optimality does.

Optimality, then, is a major problem in Dennett's theory. Evolutionary theory would require an adaptationist reading to support optimality, and there are many reasons to reject an adaptationist version of evolutionary theory. Further, an adaptationist reading would require the glossing over of such primary features of our behaviors as reading, given that reading and many other behavioral features are not adaptations but spandrels. Yet many of these behaviors are central and deserve explanation in some terms, if not adaptationist ones. Most importantly, it seems as though intentional predictions can be and often are made accurately using a strategy that does not incorporate optimality at all. For these reasons, I think that Daniel Dennett's Intentional Stance would be stronger without the assumption of optimality.

CARL BROWNSON

References

- Dennett, Daniel. *The Intentional Stance*. Cambridge: MIT Press, 1987.
- Gould, Steven and Lewontin, Richard. "The Spandrels of San Marcos and the Panglossian Paradigm: A Critique of the Adaptionist Programme." *Proceedings of the Royal Society B*205 (1979): 581-98.
- Sterelny, Kim and Paul Griffiths. *Sex and Death*. Chicago: University of Chicago Press, 1999.
- Stich, Stephen. "Dennett on Intentional Systems," *Philosophical Topics* 12 (1981): 38-62.