

# Why creative intelligence is hard to find

Daniel Dennett

Center for Cognitive Studies, Tufts University, Medford, Mass. 02155

The concluding question and answer of Whiten & Byrne's (W&B's) valuable survey is worth repeating, if only to save W&B from a likely misreading: "But can anecdotes ever be more than a jumping off point for more systematic work? We propose that the answer must be 'no' . . . ." I have been dismayed to learn that my own limited defense (Dennett 1983) of the tactic of provoking "anecdotes" – generating single instances of otherwise highly improbable behaviors under controlled circumstances – has been misinterpreted by some enthusiasts as giving them license to replace tedious experimentation with the gathering of anecdotes. But as W&B stress, anecdotes are a prelude, not a substitute, for systematic observation and controlled experiments.

They describe in outline the further courses of experimental work that would shed light on the phenomena of tactical deception, and suggest that the postponed question of whether these behaviors arise as a result of "creative intelligence" may be settled by the outcome of such research. This research should certainly be pursued, and if properly conducted its results are bound to shed light on these issues, but it is worth noting in advance that no matter how clean the data are, and indeed no matter how *uniform* they are, there is a systematic instability in the phenomenon of creatively intelligent tactical deception (if it exists!) that will tend to frustrate efforts of interpretation.

To see this, consider the range of possibilities available in the generic case, stripped to its essentials. Suppose AGENT intelligently creates a deceptive tactic that is devastatingly effective on a first trial against TARGET. Will it tend to be repeated in similar circumstances? Yes; *ex hypothesi* it was intelligently created rather than a result of blind luck or sheer coincidence, so AGENT can be supposed to recognize and appreciate the effect achieved. But then there are two possible outcomes to such repetition: Either it will provoke countermeasures from TARGET (who is no dummy, and can be fooled once or twice, but will eventually catch on), or it won't (TARGET is not so smart after all). If it doesn't, then the exploitative behavior will become (and be seen to be) stereotyped, and ipso facto will be interpretable not as a sign of creative intelligence but as a useful habit whose very cleverness is diminished by our lower regard for the creative intelligence of the TARGET. If, on the other hand, TARGET attempts countermeasures, then either they will work or they won't. If they don't work, TARGET will be seen once again to be an unworthy opponent, and the deceptive behavior a mere good habit. If the countermeasures tend to work, then either AGENT notices that they do, and thereupon revises his schemes, or not. If not, then AGENT's intelligence will be put into question, whereas if AGENT does come up with a suitable revision then he

will tend *not* to repeat the behavior with which we began, but rather some relatively novel successor behavior.

In other words, if both AGENT and TARGET are capable of creative intelligence, then what must ensue is either an escalating arms race of ploy and counterploy or a semistable equilibrium in which the frequency of deceptive tactics is close to "chance" (a tactic will work against a wily TARGET only if infrequently used, something a wily AGENT will understand). Escalations, however, are bound to be short-lived phenomena, punctuating relatively static periods. (If AGENTS and TARGETS are so smart, one might ask, why haven't they already discovered – and exhausted – the opportunities for escalation?) So the conditions one would predict in any community in which there is genuine creative intelligence are conditions systematically difficult to distinguish from "mere chance" fluctuations from a norm of trustworthiness. Any regularly repeated exploitative behaviors are in themselves grounds for diminishing our esteem for the creative intelligence of either AGENT or TARGET or both.

Our estimation of AGENT's cleverness is in part a function of our estimation of TARGET's cleverness. The more stupid TARGET appears, the less we will be impressed by AGENT's success. The smarter TARGET is with countermeasures, the less success AGENT will have. If AGENT persists in spite of failure, AGENT's intelligence is rendered suspect, while if AGENT largely abandons the tactic, we will not have any clear way of determining whether AGENT's initial success was dumb luck, unrecognized by AGENT, or a tactic wisely perceived by AGENT to have outlived its usefulness.

This can be made to appear paradoxical – a proof that there couldn't be any such thing as genuine creative intelligence. Once or twice is dumb luck; many times is boring habit; in between there is no stable rate that counts clearly and unequivocally as creative intelligence. If we set the threshold in this fashion we can guarantee that nothing could count as genuine creative intelligence. Using just the same move, evolutionists could prove that no history of natural selection *could* count as an unproblematic instance of adaptation; the first innovation counts as luck, while its mere preservation unenhanced counts as uncreative.

Clearly, we must adjust our presuppositions if we want to make use of such concepts as creative intelligence or adaptation. The stripped-down paradigmatic case exposes the evidential problem by leaving out all the idiosyncratic but telling details that are apt to convince us (one way or the other) in particular cases. If we see AGENT adjusting his ploys to the particularities of TARGET's counterploys, or even just reserving his ploys for those occasions in which he can detect evidence that TARGET can be caught off guard, our conviction that AGENT's infrequent attempts at deception are intelligently guided will be enhanced. But this seeing is itself a product of interpretation, and can only be supported by longitudinal observation of variegated, not repeated, behavior. There could be no surefire signs of creative intelligence observable in isolated, individual bits of behavior.

What we will always have to rely on to persuade us of the intelligence of the members of a species is the *continued* variation, enhancement, and adjustment of ploys in the face of counterploys. So it is not repeated behaviors but changing behaviors that are the sign of creative intelligence, and this can be observed only in the long term. Moreover, the long-term history of any genuinely intelligent agent will show this proclivity to make novel and appropriate advances mixed with a smattering of false starts, unlucky breaks and "bad ideas." So we should set aside the illusory hope of finding "conclusive" empirical evidence of creative intelligence, evidence that can withstand all skeptical attempts at a "demoting" reinterpretation. When we recognize that the concept of intelligence is not a neutral or entirely objective one, but rather an evaluative one, this should not surprise us.