unwise to claim that human cognitive plasticity is totally discontinuous from that of other species.

5.12. Brain

Grading: Connectionism: Best
ACT-R: Worse

Classical connectionism, as advertised, presents a strong position on how the mind is implemented in the brain. Of course, there is the frequently expressed question of whether the brain that classical connectionism assumes happens to correspond to the human brain. Assumptions of equipotentiality and the backprop algorithm are frequent targets for such criticisms, and many nonclassical connectionist approaches take these problems as starting points for their efforts.

There is a partial theory about how ACT-R is instantiated in the brain. ACT-RN has established the neural plausibility of the ACT-R computations, and we have indicated rough neural correlates for the architectural components. Recently completed neural imaging studies (Anderson et al. 2003; Fincham et al. 2002; Sohn et al. 2000) have confirmed the mapping of ACT-R processes onto specific brain regions (e.g., goal manipulations onto the dorsolateral prefrontal cortex). There is also an ACT-R model of frontal patient deficits (Kimberg & Farah 1993). However, there is not the systematic development that is characteristic of classical connectionism. While we are optimistic that further effort will improve ACT-R's performance on this criteria, it is not there yet.

6. Conclusion

Probably others will question the grading and argue that certain criteria need to be re-ranked for one or both of the theoretical positions. Many of the arguments will be legitimate complaints, and we are likely to respond by either defending the grading, or conceding an adjustment in it. However, the main point of this target article is that the theories should be evaluated on all 12 criteria, and the grades point to where the theories need more work.

Speaking for ACT-R, where will an attempt to improve lead? In the case of some areas like language and development, it appears that improving the score simply comes down to adopting the connectionist strategy of applying ACT-R in depth to more empirical targets of opportunity. We could be surprised, but so far these applications have not fundamentally impacted the architecture. The efforts to extend ACT-R to account for dynamic behavior through perception and action yielded a quite different outcome. At first, ACT-R/PM was just an importation, largely from EPIC (Meyer & Kieras 1997) to provide input and output to ACT-R's cognitive engine. However, it became clear that ACT-R's cognitive components (the retrieval and goal buffers in Fig. 1) should be redesigned to be more like the sensory and motor buffers. This led to a system that more successfully met the dynamic behavior criterion and has much future promise in this regard. Thus, incorporating the perceptual and motor modules fundamentally changed the architecture. We suspect that similar fundamental changes will occur as ACT-R is extended to deal further with the brain criterion.

Where would attention to these criteria take classical connectionism? First, we should acknowledge that it is not clear that classical connectionists will pay attention to these criteria or even acknowledge that the criteria are reasonable. However, if they were to try to achieve the criteria, we suspect that it would move connectionism to a concern with more complex tasks and symbolic processing. We would not be surprised if it took them in a direction of a theory more like ACT-R, even as ACT-R has moved in a direction that is more compatible with connectionism. Indeed, many attempts have been made recently to integrate connectionist and symbolic mechanisms into hybrid systems (Sun 1994; 2002). More generally, if researchers of all theoretical persuasions did try to pursue a broad range of criteria, we believe that distinctions among theoretical positions would dissolve and psychology will finally provide "the kind of encompassing of its subject matter – the behavior of man – that we all posit as a characteristic of a mature science" (Newell 1973, p. 288).

NOTE

1. The complete list of published ACT-R models between 1997 and 2002 is available from the ACT-R home page at: **http://act-r.psy.cmu.edu/**

ACKNOWLEDGMENTS

Preparation of this manuscript was supported by ONR Grant N00014–96–1-C491. We would like to thank Gary Marcus and Alex Petrov for their comments on this manuscript. We would also like to thank Jay McClelland and David Plaut for many relevant and helpful discussions, although we note they explicitly chose to absent themselves from any participation that could be taken as adopting a stance on anything in this paper.

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Newell's list

Joseph Agassi

Department of Philosophy, Tel-Aviv University, Tel-Aviv 69978, Israel.
agass@post.tau.ac.il **http://www.tau.ac.il/~agass/**

Abstract: Newell wanted a theory of cognition to abide by some explicit criteria, here called the Newell Test. The test differs from the Turing Test because it is explicit. The Newell Test will include the Turing Test if its characterization of cognition is complete. It is not. Its use here is open-ended: A system that does not pass it well invites improvement.

Alan Newell asserted that an adequate theory of a functioning system of human cognition should abide by some explicit criteria, and he offered a list of such criteria. The list includes characteristics such as flexible, adaptive behavior; possession of a vast knowledge base; and the ability to integrate knowledge, use a natural language, and learn. The target article authors say that, although this list is not complete, it certainly is "enough to avoid theoretical myopia" (sect. 1, para. 2). Hardly: Myopia is the outcome of the claim for knowledge of natural languages and learning sufficient to per-

https://doi.org/10.1017/S0140525X03220136 Published online by Cambridge University Press

mit decision as to whether a given theory of cognition captures them adequately. We just do not know that much as yet.

The authors say that the criteria deserve "greater scientific prominence." They therefore try to "evaluate theories by how well they do in meeting" the criteria (sect. 1, para. 4). This may be premature. Whether it is, depends on the merit of Newell's idea more than on its applications. So, it requires examination. What the authors call the Newell Test is a test not of Newell's idea but of the theories that should agree with it – provided that it is valid. Is it? How are we to judge this? (The claim that a characterization of "a functioning system of human cognition" is possible, is significant for those who plan to realize it. The claim and the plan pertain to studies of artificial intelligence that are no longer popular. See Agassi 1988, reprinted in Agassi 2003.)

Anderson & Lebiere (A&L) apply Newell's Test to two new ideas that are controversial, so the application cannot be such a test. Hence, their work is begging the question: Some test of it is required to show that it deserves a "greater scientific prominence."

"Newell is calling us to consider all the criteria and not pick and choose the ones to consider" (sect. 2.8). This remark renders the whole venture too questionable. The authors make it a propos discussion of the criterion of consciousness.

> Newell acknowledged the importance of consciousness to a full account of human cognition, although he felt compelled to remark that "it is not evident what functional role self-awareness plays in the total scheme of mind." We too have tended to regard consciousness as epiphenomenal . . . (sect. 2.8)

This is very shaky. Whether consciousness is or is not epiphenomenal is a red herring: It is an empirical fact that in many cases cognitive conduct differs depending on whether it is accompanied with consciousness or not, and the question may arise, should a system emulating human consciousness reflect this fact? Importantly, Turing's celebrated idea of the Turing Test is designed to avoid this question altogether.

The authors examine two sets, classical connectionism and ACT-R. Classical connectionism is a computerized version of behaviorism. (For the claim that connectionism is unscientific, see Agassi 1988, reprinted in Agassi 2003.) ACT-R is "a theory of cognition," "a subsymbolic activation-based memory" able "to interact with a symbolic system of production rules"; the "R" in ACT-R "denotes rational analysis" (sect. 4.1, first para.). The two sets, then, are artificial intelligence or expert-systems programs. The authors report a claim that classical connectionism passes the Turing Test. Presumably they disagree. The same holds for ACT-R. "ACT-R, except for its subsymbolic limitations, is Turing equivalent, as are most production systems" and "(proof for an early version of ACT appears in Anderson 1976)" (sect. 5.1, para. 2). This is a bit cryptic; I will explain the difference between the Turing and the Newell Tests in the following paragraph.

The Turing Test was meant to render the mind-body problem empirically decidable. Were there a computer program that could fool an expert, Turing suggested, then it would be empirically indistinguishable from humans, and so the attribution to humans of a metaphysical soul would be redundant. (For a critique of this test, see Agassi 1992.) Because Newell's criteria depict human characteristics, any interlocutor who can pass the Turing Test should certainly possess them, because the inability to exhibit any human characteristic the like of which Newell mentions would expose the impostor. And yet, the Turing Test is implicit and Newell's Test is explicit. This permits finding a partial success in passing the Newell Test. But, to be an explicit version of the Turing Test, the Newell Test must refer to a complete list of characteristics. We do not have this, and the Turing Test may be preferred just because it leaves this task to the experts who wish to test the humanity of their enigmatic interlocutor. Consequently, a Turing Test can never be decisive: Both expert and programmer can improve on prior situations and thus deem failure a merely temporary setback. True, the Turing Test is generally deemed possibly decisive, and, being a thought-experiment, actually decisive. Some writers, notably Daniel Dennett, claim that only exorbitant costs prevent the construction of a machine that will pass the Turing Test. That machine, then, should certainly be able to pass the Newell Test with flying colors. It is a pity that A&L do not refer to this claim and expose it as a sham. If they are anywhere close to being right, they should be able to do so with ease. (See Agassi & Laor 2000, pp. 556–62.)

The interesting aspect of the target article is that it is open-ended: Whenever the system A&L advocate, which is ACT-R, does not pass the examination as well as they wish, they recommend trying an improvement, leading to a retest. They should observe that such a move may be two-pronged. They refer to the improvement of the ability of a program to abide by the theory of flexibility; adaptive behavior; and the ability to integrate knowledge, use a natural language, and learn. They should not ignore the need to improve on these theories. When they refer to natural languages or to learning, they view the connectionist idea of them as more satisfactory than that of ACT-R, because it is more complete. Yet, whatever completeness is exactly, it is not enough: We seek explanations, and so to accept axiomatically what we want to understand is not good enough. We still do not know what a natural language is and how we learn; and we do not begin to understand these. Let me end with an insight of David Marr that should not be forgotten. Emulation is helpful for the understanding but is no substitute for it; sometimes, the very success of emulation, Marr observed, renders it less useful than a problematic one. We want understanding, not mere emulation. (See Marr 1982, p. 253 and Agassi 1987, p. 410.)

Think globally, ask functionally

Erik M. Altmann

Department of Psychology, Michigan State University, East Lansing, MI 48824. **ema@msu.edu** **http://www.msu.edu/~ema**

Abstract: The notion of functionality is appropriately central to the Newell Test but is also critical at a lower level, in development of cognitive sub-theories. I illustrate, on one hand, how far this principle is from general acceptance among verbal theoreticians, and, on the other hand, how simulation models (here implemented within ACT-R) seem to drive the functional question automatically.

Anderson & Newell (A&L) have been carrying integrative cognitive theory, in shifts, for the past 30 years or so (if one goes back to Newell 1973). We are fortunate that Anderson is young; formulating dichotomous questions – seeing the trees but not the forest – may be the modal tenure procedure in psychology departments today, but perhaps in another generation it will be acceptable not to conduct new experiments at all but simply to integrate old data into increasingly complete computational models.

In the meantime, how can we avoid theoretical myopia in our daily research? Applying the Newell Test is well and good once a decade or so, with that many years' interim progress available to assess it. In terms of the next chunk of publishable research, however, it's useful to have more immediate guidance.

Central to the Newell Test is the idea of functionality: A theory has to explain how the cognitive system accomplishes some particular function. Among the Newell Test criteria, this function is high level, related in some relatively direct way to the fitness of the organism. However, as one develops micro-theories within a larger theory, functionality is still a relevant question; one can ask, for each process within a model, what it is for. Its outputs could, for example, be necessary inputs for another process in a chain that leads ultimately to accomplishing the task at hand. Or, one could ask whether each behavioral measure reflects a distinct process at all; perhaps it reflects a side effect of some other functionally necessary process. In both cases, it is difficult if not impossible to address the functional question without a precise representation of the processes one is talking about. In practice, this implies a computational simulation.

https://doi.org/10.1017/S0140525X03220136 Published online by Cambridge University Press