Against Simplicity and Cognitive Individualism

by

Nathaniel T. Wilcox*

Abstract

Neuroeconomics illustrates our deepening descent into the details of individual cognition. This descent is guided by the implicit assumption that "individual human" is the important "agent" of neoclassical economics. I argue here that this assumption is neither obviously correct, nor of primary importance to human economies. In particular I suggest that the main genius of the human species lies with its ability to distribute cognition across individuals, and to incrementally accumulate physical and social cognitive artifacts that largely obviate the innate biological limitations of individuals. If this is largely why our economies grow, then we should be much more interested in distributed cognition in human groups, and correspondingly less interested in individual cognition. We should also be much more interested in the cultural accumulation of cognitive artifacts: computational devices and media, social structures and economic institutions.

*Nathaniel T. Wilcox is Professor of Economics at the University of Houston. His research interests are economic, political and cognitive science, experimental methods and applied econometrics.

Experimental and behavioral economics and, by association, neuroeconomics overwhelmingly involve this kind of experiment: *Simple* choices for *individuals* to make, perhaps in the context of a *simple* social interaction such as a two-person game. Importantly, each subject's access to other subjects is tightly circumscribed. In individual choice experiments, such access is simply forbidden for statistical reasons, and in game experiments it is closely regulated for both game-theoretic and statistical reasons. The hunch seems to be (1) that individual minds are the centrally important information-processing units of economic science, and (2) that simple decision problems and simple games usefully characterize the actual decision problems and actual games that govern the important economic phenomena, such as growth and inequality. For example, McCabe (2008, this volume) singles out trust and reciprocity as enablers of economic growth, and perhaps they are. We are invited to view experiments about simple trust games between two individuals (and what happens within the individual skulls involved in such games) as crucially relevant to economic growth.

Let me illustrate an alternative hunch. I rather suspect that economic growth is overwhelmingly due to the cultural accumulation of *external* cognitive artifacts, including *social distributions of information processing*, which implement *external* information representations and algorithms. This accumulation increasingly exploits the performance strengths of individual cognition (e.g. hand-eye coordination, fast pattern recognition) and increasingly avoids its weaknesses (e.g. algebra, emotional interference). It increasingly *distributes information processing across group members* in highly structured ways, aided by external cognitive artifacts, and *lessens the importance of individual brains* to economically important information-processing tasks. I also suspect this accumulation is *not* mostly the product of *individual* learning, but is rather mostly the product of *social* learning processes that operate across contemporaneous social space and (especially) across generations of social groups confronting related tasks. This story of economic growth is of the *increasing marginality of individual brains* and the *growing centrality of socially distributed cognition*. I think this is a plausible alternative hunch about cognitive science's primary relevance to economic growth.

I will call the viewpoint I am criticizing "cognitive individualism." It powerfully shapes and constrains theory, research questions and empirical inquiry. Importantly for

this volume, it unsurprisingly draws researchers and resources ever deeper into the study of individual cognition. In an intellectual milieu guided by cognitive individualism, the emergence of neuroeconomics (and the strong support it gets from many distinguished scholars) was inevitable. To be clear, neuroeconomics is not my target here; it is just one particularly vivid manifestation of my true target.

I suspect that many decision and discovery processes that govern the most important economic phenomena involve a high degree of social and/or computational complexity. This is notoriously true of many dynamic decision problems and their solution algorithms: I suspect that a substantial part of inequality is due to differential success (across dynastic families) at intergenerational innovation and transmission of good heuristics for dealing with life-cycle decisions. Related considerations apply to technological progress and hence growth. I simply don't believe I have this nifty laptop or this swell alphabet *merely* (or even *mostly*) because of individual cognitive solutions to simple tasks or games. Rich, external cognitive artifacts like these were accumulated by increment and innovation across years and indeed millennia. I suspect the information-processing units responsible for these increments and innovations were usually *groups* of people and *rarely* individual brains (notwithstanding our seeming need to render scientific and technological history in terms of individual heroes). That is, I suspect our accumulation of cognitive artifacts (that transform our tasks and shape distributed cognition) is overwhelmingly the product of *social learning processes*. And so I suspect that most of the experimental and theoretical work on individual learning, with or without brain scans, has limited relevance to this accumulated cognition, and hence limited relevance to economic growth. Or at least, so goes an alternative hunch.

To put some formal structure on this alternative hunch, I begin by (tentatively) viewing the relationship between neoclassical economics and cognitive science from the perspective of Marr (1982). Marr divides understanding of an information processor into three levels. At the top level, we have a rational model of a processor—a description of *what* a processor ought to compute, and an argument as to *why* it should compute that. Marr calls this the "computational theory" or CT of the processor. Addition is a CT of a cash register; a Nash equilibrium strategy is a CT of a game player (given certain well-known assumptions about other players); and so on. The next level of understanding is

"algorithm and representation" or A&R. Here we find representations of inputs and outputs, and the algorithms that convert the former into the latter, such that the output approximates the CT's specified output in some satisfactory way—robustly, efficiently, accurately or with some tradeoff amongst these. Usually, a great variety of A&Rs can approximate the specified output of any CT. We might (with some abuse) call the A&R "software." Finally there is the bottom level of "physical implementation" or PI: What's the actual physical "machine" or "hardware" in which the A&R will "run?" Vacuum tubes, silicon, neurons or Tinker Toys?—Again, many PIs can "run" any given A&R.

From a Marrian perspective, neoclassical economics is CT without any specification of A&R or PI: It is the hunch that the outputs specified by decision and game theory—the neoclassical CTs—are, to an empirically fruitful approximation, the observed behavior of economic processors under almost all conditions.[1] Marr thought that CT was centrally important for understanding processors, and some philosophers such as Dennett (1987) seem sympathetic to that view (with caveats). This is why the title of Gul and Pesendorfer's (2008) spirited defense of neoclassical economics, "The Case for Mindless Economics," makes me wince. That title just invites distracting name-calling like "behaviorist" and "Skinnerian" for no good reason. With a title like "The Case for Algorithm-, Representation- and Machine-Free Economics," Gul and Pesendorfer would pretend to know and care about cognitive science in a respectable Marrian sense.

Marr's three-level view puts neuroeconomics in danger of appearing a sideshow about PI (hardware) which, in Marr's view, is frequently a distraction from the really important issues that largely revolve around an A&R's performance relative to CT's specifications. Philosophically, the notion of "multiple realizability" of the mental in the physical also suggests that hardware is a distraction: The important essence of the mental may not be usefully reduced to any particular machine (Fodor 1975; Putnam 1988). McCabe's quick and early attempt to dismiss the software/hardware question attests to

[1] Marr considered situations where "constraints" (within a CT's "why argument") define a *unique* desired mapping between inputs and outputs. Marr's "constraints" conceptually correspond to decision-theoretic axioms: Both show that a certain desired mapping between input and output *exists* (while saying nothing about how that might be computed) and has certain properties. But decision theoretic axioms are typically weaker: They do not define a *unique* mapping (e.g. risk attitude and/or belief diversity within the structure of subjective expected utility or SEU). We might therefore wish to distinguish between strong CTs (like Marr's example of stereopsis) and weaker CTs (such as SEU). The typically weaker form of economic CTs is a nontrivial complication of a Marrian view of neoclassical theory.

the danger. Neuroeconomics answers: We accept Marr's distinctions and views *in general, but* we suspect that the individual human nervous system PI is *in fact* an especially strong constraint on the kinds of A&R that will "run" in it. Some neuroscientists go even further, arguing to dispense with CT and dissolve the distinction between A&R and PI in their special case: They say that while it is a cogent distinction *in general*, it is practically useless in the case of biological nervous systems. This line of thinking has its own philosophical heroes (e.g. Churchland 1981), and usually goes along with some stories about evolution (McCabe hints at these without development).

Not all neuroeconomists want to dispense with CT: For instance, Glimcher (2003) seems quite sympathetic with Marr's view that rigorous cognitive science cannot proceed without a relevant CT. Philosophers such as Fodor (2000) have attacked *both* the notion that biological hardware is crucially special, *and* the relevance of adaptationist stories concerning that hardware's structure. I find evolutionary talk inherently seductive, but I try to keep Gould and Lewontin (1979) in mind when considering *ex post* just-so stories about "adaptations." Yet in spite of these difficulties, I'll grant neuroeconomists the specialness of the biological hardware. It matters little to my subsequent argument.

Since introducing Marr above, I've said nothing explicit about *which* processor is the centrally important processor of economic science—what we call an *agent* at the level of economic CTs. Fudenberg and Levine's (2006) "dual self" model of impatience and small-stakes risk aversion is just a recent example of theoretical fission of individuals into multiple noncooperative agents, a move endorsed by some philosophers (Ross 2005). Fissions of the agency of firms (into managers versus owners) and households (into spouse versus spouse) similarly create new noncooperative games between new smaller agents—an older and well-established move, as Camerer (2008) points out. What all of these moves have in common is *fission* of agency. Might it also be useful to explore some *fusions* of agency?

Edwin Hutchins is a cognitive anthropologist known to some as the "father of modern cognitive ethnography," and a contemporary developer of theories of socially distributed cognition. Hutchins enormously influenced philosophers of mind such as Clark (1997). In his book *Cognition in the Wild*, Hutchins (1995) argues that the founding metaphor of cognitive psychology—that individual cognition is like a computer—was always flawed.

Hutchins instead suggests that it is socio-cultural symbol manipulation systems (cognition distributed across groups of humans, in interaction with external cognitive artifacts and both natural and artificial environments) that are *really* like a computer.

Hutchins (1995, p. 50) begins gently and craftily, pointing out that "Marr intended his framework to be applied to the cognitive processes that take place inside an individual, but there is no reason, in principle, to confine it to such a narrow conception of cognition." To elaborate "Hutchins' hunch," individual brains are not the machines that matter the most in social science. The important machines of social science are social groups with special artifacts, communicating with one another both now and across time and even across generational time. In these larger machines, individual brains are not the machines of interest: Instead, individuals (brain and body) are just *one type of cog* in larger socially distributed cognitive machines. We might think of the usual path of cognitive-cultural history (and economic growth) as the successive accumulation and modification of *external* cognitive artifacts: *external* representations of information, *external* physical devices and *external* patterns of information flow and aggregation. These are to some extent social A&R, but also *additional types of cogs* in larger socially distributed machines. Cognitive-cultural change may tend to (1) increasingly exploit strengths of human cogs and avoid their weaknesses, (2) make human cogs interchangeable and redundant where possible, and (3) increase the robustness of the bigger social machines to their errors, weaknesses, and arrivals and departures.

Now let us return to the neuroeconomists' attempted dismissal of the Marrian suspicion that hardware is a sideshow. From the viewpoint of Hutchins' hunch, I answer: While I sympathize with your desire to know a source of the relative strengths and weaknesses of *one part* (the brain) of the human cogs, this knowledge might not be very important for explaining (say) economic growth, or wealth, poverty and inequality. You are looking at just *one part* (brain, not body) of what may be a *relatively inconsequential cog*. Put differently, you may be mistaken about what the most important processor is, how flexible it is, and how biologically determined it is. In recounting the evolution of Western navigation artifacts, Hutchins (1995, pp. 155) notes that

> "…the existence of such a wide variety of specialized tools and techniques is evidence of a good deal of cultural elaboration directed toward avoiding algebraic reasoning and arithmetic…The kinds of cognitive tasks that people face in the wild cannot be inferred from the [CT of navigation] alone. The specific implementations of the task determine the kinds of cognitive processes that the performer will have to organize in order to do the task. The implementations are, in turn, part of a cultural process that tends to collect [external artifacts embodying specialized A&R] that permit tasks to be performed by means of simple cognitive processes."

To Hutchins, social hardware (in the form of both human cogs and external cognitive artifacts) does matter *in the short run*, but in the long run the hardware of distributed cognition isn't even given: It is *culturally evolved* to *sidestep* limits of human cogs and exploit their strengths. *Both* the software *and* hardware of distributed cognition are, as Hutchins says, "artificial through and through." Neural biology *isn't* social cognitive destiny.

From the viewpoint of Hutchins' hunch, doing rigorous Marrian cognitive science is *overwhelmingly easier in social machines than in brains*. I cannot put this better than Hutchins (1995, pp. 128-129):

> The basic computations of navigation could be characterized at the computational, representational/algorithmic, and implementational levels entirely in terms of *observable* representations. On this view of cognitive systems, communication among the actors is seen as a process *internal* to the cognitive system. Computational media, such as diagrams and charts, are seen as representations internal to the system, and the computations carried out upon them are more processes internal to the system. Because the cognitive activity is distributed across a social network, many of these internal processes and internal communications are *directly observable*. If a cognitive psychologist could get inside a human mind, he or she would want to look at the nature of the representations of knowledge, the nature and kind of communication among processes, and the organization of the information-processing apparatus. We might imagine, in such a fantasy, that at some level of detail underlying processes (the mechanics of synaptic junctions, for instance) would still be obscured. But if we could directly examine the transformations of knowledge representations we might not care about the layers that remain invisible. Any cognitive psychologist would be happy enough to be able to look directly at the content of the cognitive system. With systems of socially distributed cognition we *can* step inside the cognitive system, and while some underlying processes (inside people's heads) remain obscured, a great deal of the internal organization and operation of the system *is* directly observable. On this view, it might be possible to go quite far with a cognitive science that is neither mentalistic (remaining agnostic on the issue of representations "in the head") nor behavioristic (remaining committed

> to the analysis of information processing and the transformation of representations "inside the cognitive system") [all emphasis in original].

Neuroscience does not "look directly at the content" of any information representation in any brain as concrete, observable, rich and obvious as a ship's fix and projected bearing on a navigation chart, or a family's current and projected asset position as represented in its own external financial accounts and plans. As near as I can make out, they haven't yet agreed exactly where, or even how, the content of nouns like "ship" are physically represented in a brain (Martin and Caramazza 2003). Hutchins is, I think, still correct in spirit after thirteen years: What is still a fantastic voyage for neuroscience has *long* been feasible for social scientists examining distributed cognition systems. We can *now* (with the right ethnographic skills) literally step into these and *directly observe* just about every theoretically central entity of rigorous Marrian cognitive science.

For these reasons, I believe it may be fruitful to move away from studying what individuals do with simple decision problems and games, and toward studying distributed cognition in groups that confront complex problems and games. I now view my past work on social learning about hard economic problems (Ballinger, Palumbo and Wilcox 2003) as my initial steps in that direction. Ballinger and I continue to study the evolution of advice (about hard economic problems) across generations of subjects. Advice is, after all, a relatively simple cultural-cognitive artifact, and one whose evolution we can follow over workable experimental timescales. Schotter and Sopher (2003) share similar interests with us.

Though it is rare relative to all work, some experimental economists *have* been contrasting the behavior of individuals and groups. So far, though, these groups are overwhelmingly either unstructured groups or "symmetric committees" in which all group members have equal access to information, cognitive artifacts and each other, and all have equal sway in decision making (e.g. majority or unanimity rule). This is not distributed cognition in Hutchins' interesting sense (though such groups may be an interesting control treatment when studying distributed cognition, and such groups may have, on occasion, endogenously created their own truly distributed cognition). Most distributed cognition in real groups has asymmetric and hierarchical structural features

that divide tasks into subtasks, restrict data availability and information flows, and vest final decision making in only a subset of group members.[2]

Asymmetry and hierarchy are consequential features of distributed cognition. When viewed from the perspective of its place in distributed cognition, we discover that some shopworn normative failings of individual cognition may be social virtues in correctly structured groups. Hutchins (1995, p. 240) shows how "…some ways of organizing people around thinking tasks will lead to an exacerbation of the maladaptive aspects of [confirmation bias], whereas other forms of organization will actually make an adaptive virtue on the group level of what appears to be an individual vice." Hutchins' simulation studies of networks of individual interpreters suggest that it is just those social networks with *lots* of symmetric linkage of individuals (e.g. symmetric face-to-face committees) that would exacerbate confirmation bias. We might want to look closely at some real economic groups (e.g. extended families, parts of firms and/or public bureaucracies) to identify some typical structures of distributed economic cognition: Those might be good candidate structures for study. Symmetric committees are probably quite rare in distributed economic cognition—and with good reason, given Hutchins' arguments.

Told that I could select only one tool from a neighboring discipline, it would be ethnography, not medical imaging. A simple and false choice, I hear you cry: It excludes a large middle of more complex alternatives you actually have, and also ignores a rich division of cognitive labor across researchers and disciplines. Fair enough. But if that cry *was* yours, then have some sympathy for individuals in typical experiments: They too are confronted with simple and false choices that exclude many middles (and extremes) they'd actually face in the field; and they too are sealed off from advisors, artifacts and other distributed cognitive resources they'd have in the field. The *situated* character of individual cognition in the field is one reason Harrison and List (2004) advocate more field experiments. Understanding situated individual cognition is almost definitionally crucial to understanding distributed cognition systems. But Hutchins' hunch that the

[2] Cox and Hayne (2005) do have a small *ex post* asymmetry in groups that are bidding in common value auctions—group members with different signals of value; but all group members are still *ex ante* symmetric, and group members all seemed to immediately share their information with one another. I have in mind very different experiments, in which multiple types of information (and not simply multiple draws from a univariate distribution of a decision-relevant random variable) must be processed and then combined to reach a decision.

theoretical entities of classical cognitive psychology are better applied to distributed cognition than individual cognition is quite a bit beyond that.

I'm well aware that the ideas set out above will raise many hackles. To spare others the effort, I'll hurl the epithet "group selectionist" at myself, since I seem to advocate the fusion of individual persons into a larger agent. I'll also admit that *I'm personally* not all that interested in dealing with "crucial foundational problems" about groups. For instance, subjects seem to effortlessly overcome the pull of many bad equilibria when they can talk, observe and gossip the way they actually do in the real world (Duffy and Feltovich 2006). Yes, it *is* an interesting question why that is so: I hope for good answers and look forward to reading them with deep interest. No, the serious study of distributed economic cognition *does not need to*, and *should not*, wait for good answers to that question. So don't bug me with the predictable question: "But what about the (insert favorite game-theoretic puzzle) problem you are obviously ignoring in this group?" One person's—even one paper's—foundational problem is another's distracting detail.

More seriously and sympathetically: Economic theorists have already added important insights to a full understanding of distributed cognition, and I hope this continues. By now it is banal that economic institutions *are* important external cognitive artifacts that mediate *specific kinds* of distributed cognition at many scales: Philosophers of mind such as Clark (1997) are now getting this. At the smallest scales, *within* groups, smooth distributed cognition might very well be aided and abetted by specifically economic institutional gizmos that help solve problems of coordination, shirking, trust and so forth. Yes, of course—good thinking. But distributed cognition is not *merely* the solution of those economic problems: Solving them is no guarantee that distributed cognitive systems will increment and innovate their own hardware and software so as to grow an economy. More generally, distributed cognition systems, and economic mechanisms, are almost necessarily interdependent. I view neoclassical economic theory as a *necessary part* of a complete and rigorous cognitive science but, unlike Gul and Pesendorfer (2008), I think neoclassical economic theory just doesn't address certain important economic issues—genuinely cognitive ones below the level of computational theory. And I'm beginning to doubt that the most important kinds of economic cognition occur solely in machines that only reveal their secrets to medical devices. That is a

criticism of our general obsession with individual cognition, and neuroeconomics is just one manifestation of that.

Acknowledgments

In very different ways, Daniel Friedman, Glenn W. Harrison, John S. Hutchinson, Anne Jaap Jacobson, Tamar Kugler, David K. Phillips and my father John T. Wilcox have all helped me with this work. The usual disclaimer applies with special force: No opinion or mistake here should be ascribed to any of these people.

References

Ballinger, P., Palumbo, M. and Wilcox, N. (2003). Precautionary Saving and Social Learning Across Generations: An Experiment. *Economic Journal 113*, 920-947.

Camerer, C. (2008). The Case for Mindful Economics. In A. Caplin and A. Schotter, eds., *Foundations of Positive and Normative Economics*. New York: Oxford University Press.

Churchland, P. (1981). Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy 78*, 67-90.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

Cox, J. C. and Hayne, S. C. (2005). Barking Up The Right Tree: Are Small Groups Rational Agents? *Experimental Economics 9*, 209-222.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Duffy, J. and N. Feltovich. (2006). Words, Deeds, and Lies: Strategic Behavior in Games with Multiple Signals. *Review of Economic Studies 73*, 669-688.

Fodor, J. (1975). *The Language of Thought*. New York: Thomas Cromwell.

Fodor, J. (2000). *The Mind Doesn't Work That Way*. Cambridge, MA: MIT Press.

Fudenberg, D. and Levine, D. K. (2006). A Dual Self Model of Impulse Control. *American Economic Review 96*: 1449-1476.

Glimcher, P. (2003). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. Cambridge, MA: MIT Press.

Gould, S. J. and Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences 205*, 581-598.

Gul, F. and Pesendorfer, W. (2008). The Case for Mindless Economics. In A. Caplin and A. Schotter, eds., *Foundations of Positive and Normative Economics*. New York: Oxford University Press.

Harrison, G. and List, J. (2004). Field Experiments. *Journal of Economic Literature 42*, 1013-1059.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision*. Cambridge, MA: MIT Press.

Martin, A. and Caramazza, A. (2003). Neuropsychological and Neuroimaging Perspectives on Conceptual Knowledge: An Introduction. *Cognitive Neuropsychology 20*, 195–212.

McCabe, K. (2008). Neuroeconomics and the Economic Sciences. *Economics and Philosophy* [this volume, volume number and pages to be determined].

Putnam, H. (1988). *Representation and Reality*. Cambridge, MA: MIT Press.

Ross, D. (2005). *Economic Theory and Cognitive Science: Microexplanation*. Cambridge, MA: MIT Press.

Schotter, A. and Sopher, B. (2003). Social Learning and Coordination Conventions in intergenerational games: An experimental study. *Journal of Political Economy 111*, 498-529.