

Moving hand reveals dynamics of thought

James S. Magnuson*

Department of Psychology, University of Connecticut, Storrs, CT 06269-1020; and Haskins Laboratories, 300 George Street, New Haven, CT 06511

Perception, cognition, and action occur over time. An organism must continuously and rapidly integrate sensory data with prior knowledge and potential actions at multiple timescales. This makes time of central importance in cognitive science. Crucial questions hinge on time, from the class of systems that may underlie cognition to debates about constraints on the functional organization of the brain and theoretical disputes in specialized domains. Often, the fine-grained time-course predictions that distinguish theories exceed the temporal resolution of available measures. In this issue of PNAS, Spivey *et al.* (1) introduce a method (“mouse tracking”) that provides a continuous measure of underlying perception and cognition in online language processing, promising badly needed leverage for addressing theoretical impasses, narrow and broad. I will describe examples of theoretical debates that hinge on time course, the difficulties in assessing time, and how the strengths and limitations of the new method complement current techniques for estimating time course. I conclude with a discussion of the potential of the new method to extend the tools and implications of dynamical systems theory (DST) to higher-level cognition.

The Time-Course Quandary

Precious little of perception and cognition can be observed directly and instead must be inferred from relationships between input and behavior, often from single postperceptual responses. A typical word-recognition paradigm is lexical decision. A subject sees or hears words and nonwords and hits keys indicating whether the stimulus was a word. The more frequently a word is used, the more quickly subjects respond yes; however, the more words that sound similar to it, the slower the response. So reaction times tell us a complex process of activation and competition underlies word recognition but indicate little beyond the approximate number of activated words and some of their characteristics. Theories make conflicting predictions about precisely which words are activated, and how strongly each competes over time as a word is heard (2–4). The temporal resolution needed to test these conflicting predictions exceeds by far that of methods like lexical decision.

A similar stalemate occurred in sentence processing. Theories agree that as each word is recognized, its grammatical

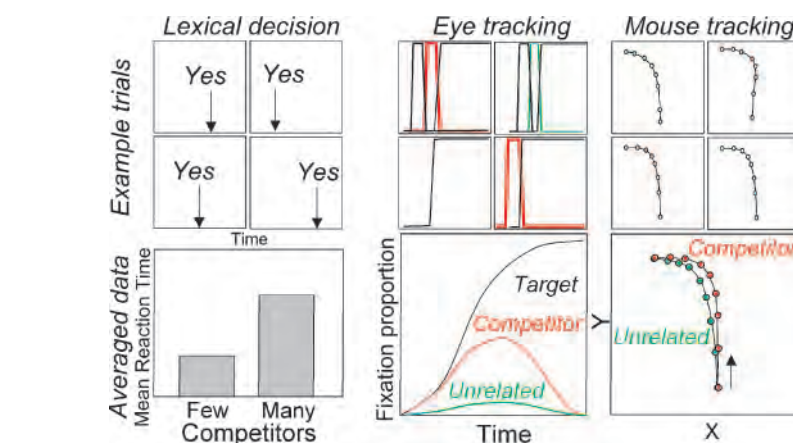


Fig. 1. Lexical decision yields single postperceptual response times that provide condition means and coarse clues to underlying processing. Eye-tracking trials provide a few discrete data points during processing; averaging over many trials yields time-course estimates of processing. Each mouse-tracking trial provides continuous time-course data (points correspond to normalized time steps), providing the most direct measure of time course and making it amenable to dynamical systems theory (DST) analysis.

category constrains the assembly of syntactic structures that in turn constrain semantics. The central theoretical debate is between syntax-first (5) and constraint-based (6, 7) theories. Syntax-first parsers initially construct the simplest syntactic structure without consideration of semantics and revise it later if syntax and semantics cannot be integrated. In constraint-based theories, semantics continuously constrain syntax. Distinguishing the theories requires testing how quickly semantics of specific words influence syntactic parsing. Given the temporal resolution of conventional techniques, cases where the semantics of individual words appear to constrain syntactic parsing (supporting constraint-based views) can be accommodated by extending syntax-first to generate multiple possible parses in parallel, ordered by simplicity, with semantic evaluation lagging a few hundred milliseconds behind (8). Resolving this debate requires measures of precisely how early semantics influences syntax.

Tanenhaus *et al.* (9) provided a dramatic step forward when they estimated time course by tracking eye movements as subjects followed spoken instructions to perform visually guided movements. Fixations lagged ≈ 200 ms behind disambiguating acoustic information, which is remarkably close, because saccades require 150 ms to plan and launch (10). Their most striking result supported constraint-based theories. An ambiguous instruction that would require a syntax-first parser to re-

ject the simplest structure (e.g., “put the apple on the towel in the box”) was paired with two visual contexts. The helpful context had two apples: one on a towel, one not. The unhelpful context had just one apple: on a towel. Both cases had an empty towel. The unhelpful context was confusing; subjects had trouble discarding their initial interpretation that they should move the apple to the empty towel, indicated by multiple fixations to the empty towel. The helpful context made the instruction as easy as an unambiguous control (“put the apple that is on the towel in the box”), with no looks to the empty towel, suggesting visual context was constraining the earliest moments of syntactic parsing.

This technique has been used to address questions in speech perception, word recognition, and sentence processing (11). This is not the end of the story, though; eye movements provide time-course estimates, but individual trials provide just a few discrete points, since a person can fixate only one location at a time. Time-course estimates are computed by averaging over many trials and are argued to reflect central tendencies of the underlying system that generated the behavior. Thus, one cannot rule out the possibility that apparently continuous estimates of

See companion article on page 10393.

*E-mail: james.magnuson@uconn.edu.

© 2005 by The National Academy of Sciences of the USA

activation based on fixations result from underlyingly discontinuous processes. One potential contribution of mouse tracking will be to resolve this concern about time course. Although individual mouse-tracking trials are averaged to reduce noise, each trial provides continuous time-course data (see Fig. 1). Spivey *et al.* (1) test the homogeneity of their tracking data to test the likelihood that it is not continuous and show that the time course of lexical activation emerges gradually in continuous behavior trial by trial. This suggests an immediate use of mouse tracking: resolving debates where eye tracking suggests immediate and continuous integration of bottom-up, top-down, and intermodality information, to test whether trial-by-trial behavior is consistent with aggregate time course.

The implications of these cases extend beyond psycholinguistics to the modularity debate. Fodor (12) argued that veridical perception depends on protecting sensory data from prior knowledge and other modalities, lest the senses lead us to hallucinate. Fodor argued for strict modularity of input systems, to ensure low-level percepts accurately reflect the physical world. The same sort of data are needed to test modularity; we must observe precisely how early top-down information influences perception, and so mouse tracking stands to play a pivotal role in this debate as well.

However, although mouse tracking has some advantages over eye tracking, future work will have to grapple with at least four limitations. First, mouse movements are largely under conscious control, whereas people are unaware of saccades unless they explicitly monitor them. Second, the physical arrangement of stimuli is crucial, limiting the naturalness of the tasks that can be used. Adding a third object directly above the starting point, or arraying items in a circle around the starting point, would have a large impact on what could be detected from mouse movements. Movement dynamics (which depend on distance from starting point, interobject distance, etc.) must not interact with language-processing dynamics, or the interaction must be quantified (e.g., by adding motor constraints explicitly to the

model). Third, effects of ambiguity persist longer in mouse than eye tracking. Fixation proportions typically diverge reliably from competitors within 200–300 ms of point of disambiguation. Mouse trajectories diverge reliably from cohort competitors over 500 ms after word offset. On the other hand, between-condition differences in trajectories emerge remarkably early. This suggests early portions of the trajectories likely reflect online processing, but care must be taken in interpreting latter

Converging evidence from eye and mouse tracking may prove key to resolving theoretical debates.

portions, which may necessarily transpire after word recognition is (largely) complete. Fourth, it must be demonstrated that trajectories do not depend on displaying competitors; we should see influences of factors like word frequency and undisplayed competitors even when targets are displayed with unrelated distractors, as is true with eye tracking (13).

Despite these limitations, converging evidence from eye and mouse tracking may prove key to resolving a range of theoretical debates. Further, the concerns with mouse tracking are minor in comparison to two more significant advantages it provides. First, eye trackers cost tens of thousands of dollars, but mouse tracking makes the study of time course possible for anyone with a computer. Second, although comparing time-course data from eye and mouse tracking is an important potential use, more importantly, this method opens up higher cognition to DST.

Time and the Nature of Cognition

There is growing awareness in cognitive science of the analytical and conceptual insights to be gained from applying DST to cognition (14). DST not only offers

tools for analyzing systems that exhibit nonlinearities over time but also has conceptual implications for the nature of cognition itself. There is growing evidence that the dynamic qualities of cognition extend to coupled interactions between (what we typically describe as) distinct cognitive domains, as well as with noncognitive aspects of organisms and their environments (15). Such interactions are consistent with the view of the organism as a self-organizing system closely adapted to and coupled with its environment and imply that we cannot partition cognition from the organism as a whole.

But there are significant obstacles to exploring the analytical and conceptual implications of DST. Cognition emerges over time, but we get only glimpses of the continuous processes that underlie cognition from discrete external behaviors. Discrete measures coupled with tight experimental constraints can be assessed for dynamic properties (15–17), but continuous measures would allow more flexibility in tasks and far greater transparency in analysis. For this reason, considerably more progress has been achieved in understanding action and coordination dynamically (18, 19).

Many connectionist models have self-organizing and dynamical properties (20, 21), and to the degree that they predict behavior, one can argue the organism also has self-organizing, dynamical properties. However, without continuous behavioral data, this argument is compelling only to the point that behavior can be shown to depend crucially on dynamical properties and that it cannot be fit by models without those characteristics. If we were able to compare directly the continuous behavior of the organism and the model, rather than discrete states at disjoint instants, we would be on much firmer ground for drawing conclusions about the underlying nature of cognition. The continuous data provided by the mouse-tracking technique introduced by Spivey *et al.* (1) is an excellent start, with the potential to address not only specialized theoretical debates but also some of the biggest questions facing cognitive science.

This work is supported by National Institute on Deafness and Other Communication Disorders Grant R01 DC-005765.

1. Spivey, M. J., Grosjean, M. & Knoblich, G. (2005) *Proc. Natl. Acad. Sci. USA* **102**, 10393–10398.
2. Luce, P. A. & Pisoni, D. B. (1998) *Ear Hear.* **19**, 1–36.
3. Marslen-Wilson, W. (1993) in *Cognitive Models of Speech Processing*, eds. Altmann, G. T. M. & Shillcock, R. (Erlbaum, Hillsdale, NJ), pp. 187–210.
4. McClelland, J. L. & Elman, J. L. (1986) *Cognit. Psychol.* **18**, 1–86.
5. Frazier, L. & Clifton, C. (1996) *Construal* (MIT Press, Cambridge, MA).
6. MacDonald, M. S., Pearlmutter, N. J. & Seidenberg, M. S. (1994) *Psychol. Rev.* **89**, 483–506.
7. Trueswell, J. C. & Tanenhaus, M. K. (1994) in *Perspectives in Sentence Processing*, eds. Clifton, C., Frazier, L. &

- Rayner, K. (Erlbaum, Hillsdale, NJ), pp. 155–179.
8. Altmann, G. & Steedman, M. (1988) *Cognition* **30**, 191–238.
9. Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995) *Science* **268**, 1632–1634.
10. Saslow, M. G. (1967) *J. Opt. Soc. Am.* **57**, 1030–1033.
11. Trueswell, J. C. & Tanenhaus, M. K., eds. (2005) *Approaches to Studying World-Situated Language Use* (MIT Press, Cambridge, MA).
12. Fodor, J. A. (1983) *The Modularity of Mind* (MIT Press, Cambridge, MA).
13. Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N. & Dahan, D. (2003) *J. Exp. Psychol. Gen.* **132**, 202–227.
14. Beer, R. D. (2000) *Trends Cognit. Sci.* **4**, 91–99.

15. Van Orden, G. C., Holden, J. G. & Turvey, M. T. (2005) *J. Exp. Psychol. Gen.* **134**, 117–123.
16. Tuller, B., Case, P., Ding, M. & Kelso, J. A. S. (1994) *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 1–16.
17. Rueckl, J. G. (2002) *Ecol. Psychol.* **14**, 5–19.
18. Kelso, J. A. S. (1995) *Dynamic Patterns: The Self-Organization of Brain and Behavior* (MIT Press, Cambridge, MA).
19. Kugler, P. N. & Turvey, M. T. (1987) *Information, Natural Law, and the Self-Assembly of Rhythmic Movement* (Erlbaum, Hillsdale, NJ).
20. Elman, J. L. (1991) *Machine Learn.* **7**, 195–225.
21. Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. E. (1996) *Psychol. Rev.* **103**, 56–115.