Much of psychological science is concerned with understanding the underlying processes that drive particular behavioral responses. When such responses occur in only a few hundred milliseconds, as they often do, gaining insight into the cognitive processes that culminate in a given response has proven difficult. The most common solution to this problem is still in use today, a century-and-a-half after Donders (1868/1969) first measured the human reaction time (RT) to infer a dissociation between hypothetically distinct processes. Since then, the study of mental chronometry has unquestionably advanced, and using RTs or RT distributions to make inferences about the time course of cognitive processes is a gold standard in the field. Measuring neural activity with high temporal resolution (e.g., electroencephalography) or patterns of visual attention (eye tracking) has also led to unprecedented insights into the temporal evolution of behavioral responses. However, although such techniques shed light onto the cognitive and neural processes accompanying a given response, more direct measures of the real-time evolution of the response itself—and of potential activation of alternative responses—have been lacking.

Over the past decade, the measurement of hand trajectories en route to choices on a screen has opened up new avenues of investigation into the dynamics of a wide range of cognitive processes. Often obtained via computer-mouse movements, hand tracking in choice tasks—and mouse tracking more specifically—is now a popular method across many areas of the field, proving to be a temporally fine-grained measure by which participants’ tentative commitments to various choice alternatives can be tracked continuously over hundreds of milliseconds. Moreover, now that software specialized for running and analyzing mouse-tracking experiments is freely available (Freeman & Ambady, 2010; Kieslich & Henninger, 2017), researchers need only a computer and a mouse to use the methodology, making its availability on par with the common RT.

Neurophysiological research in both monkeys and humans supports the use of hand movement as a valid index of evolving decisions. Specifically, activity in neuronal populations of the premotor cortex is strongly linked to hand movement, and these neuronal populations are stimulated by the decision process in a dynamic fashion. For instance, single-cell recordings revealed that during tasks in which monkeys must use a hand to select one of two response options, directionally tuned cells in the premotor cortex initially
fire for both response options simultaneously. However, as the decision-making process evolves, neuronal activity for the selected option gradually increases, whereas that for the unselected option is inhibited (Cisek & Kalaska, 2005). Such findings suggest that ongoing updates of a decision process are made immediately available to the premotor cortex, which continuously guides response-directed hand movement as a decision unfolds (Cisek & Kalaska, 2010; also see Freeman, Ambady, Midgley, & Holcomb, 2011).

In the most popular of mouse-tracking tasks, participants begin a trial by clicking a button at the bottom-center of the screen, after which they are presented with a stimulus. They then move the cursor to response alternatives in either top corner of the screen. Response alternatives may be presented before or along with stimulus onset; in some cases the alternatives are the stimuli themselves (Fig. 1a). The original and most common use of this paradigm is to measure the extent to which participants’ mouse trajectory exhibits a conspicuous attraction toward responses considered temporarily but not explicitly selected. To provide a few examples, participants’ mouse trajectories may veer toward a “female” response as a result of a male face’s feminine features (Freeman, Ambady, Rule, & Johnson, 2008); toward a “candy” response because of a spoken word’s overlapping phoneme (e.g., “candle”; Spivey, Grosjean, & Knoblich, 2005); toward an “angry” response because of stereotypes linking Black faces to hostility (Hehman, Ingbretsen, & Freeman, 2014; Stolier & Freeman, 2016); or toward an image of a cupcake before selecting a banana because of an inability to resist unhealthy food (Stillman, Medvedev, & Ferguson, 2017).

A Hidden Attraction

The early days of mouse-tracking research focused on such parallel-attraction effects to advance various dynamic models of language (Dale, Kehoe, & Spivey, 2007; Farmer, Cargill, Hindy, Dale, & Spivey, 2007; Spivey et al., 2005), social cognition (Freeman & Ambady, 2009, 2011a; Freeman et al., 2008; Freeman, Pauker, Apfelbaum, & Ambady, 2010; Wojnowicz, Ferguson, Dale, & Spivey, 2009), visual attention (Song & Nakayama, 2006, 2008), and decision making (McKinstry, Dale, & Spivey, 2008), often opposing dual-systems or stage-based models. Researchers realized that the continuous nature of hand movement, as opposed to discrete RTs or ballistic eye movements, was able to provide evidence for continuous cognitive dynamics in a way previously not possible, in turn helping to rule out alternative models (for reviews, see Freeman, Dale, & Farmer, 2011b; Song & Nakayama, 2009; Spivey & Dale, 2006).

For instance, in attitudes research, proponents of dual-systems models have long argued that people automatically activate an implicit attitude (e.g., Black people = bad) that may be subsequently corrected by an explicit attitude (e.g., Black people = good) if the attitudes conflict (Devine, 1989). In contrast, dynamic models propose that both attitudes are simultaneously activated and self-organize into a coherent evaluation (Wojnowicz et al., 2009). Or, in language research, stage-based models posit that one syntactic structure may be initially activated during sentence processing, but in ambiguous cases, this structure may be reanalyzed and replaced by a new syntactic structure if the first turns out to be inappropriate (van Gompel, Pickering, & Traxler, 2001). In contrast, dynamic-constraint-based models propose that multiple syntactic structures compete over time to stabilize on a given interpretation, without any subsequent reanalysis mechanism (Farmer et al., 2007). Using mouse tracking, researchers in these cases were able to provide evidence for a temporally continuous attraction toward two responses in parallel (i.e., both “like” and “dislike” when evaluating Black people or both syntactic interpretations when processing ambiguous sentences), thereby supporting dynamic models. Specifically, at each moment during the decision process, mouse trajectories always reflected some dynamically weighted coactivation of both implicit and explicit attitudes or both syntactic structures, which provided important challenges for dual-systems or stage-based models.

Several other early mouse-tracking studies adopted a similar logic, including studies in the domains of spoken-word recognition (Spivey et al., 2005), social categorization (Freeman et al., 2008), and decision making (McKinstry et al., 2008). Taken together, the early mouse-tracking research focused on the temporal continuity of trajectory-attraction effects to make claims about the continuous nature of underlying cognitive processes. Since then, more than 100 studies have exploited mouse tracking to index such attraction effects, but they have adopted a theoretical plurality that is no longer squarely focused on continuous dynamics. The surge of mouse-tracking research, now far broader than a debate between dynamic and dual systems, has shown that the technique can be leveraged across a wide range of domains to measure covert activations of responses that do not manifest in explicit decisions, including self-control (Sullivan, Hutcherson, Harris, & Rangel, 2015), emotion (Mattek et al., 2016), memory (Papesh & Goldinger, 2012), group processes (Lazerus, Ingbretsen, Stolier, Freeman, & Cikara, 2016), ambivalence (Schneider & Schwarz, 2017), intertemporal choice (Dshemuchadse, Scherbaum, & Goschke, 2013), theory of mind (van der Wel, Sebanz, & Knoblich, 2017).
Mouse Tracking

2014), self-esteem (Leitner, Hehman, Deegan, & Jones, 2014), moral cognition (Koop, 2013), subliminal perception (Xiao & Yamauchi, 2017), embodiment (Lepora & Pezzulo, 2015), and deception (Duran, Dale, & McNamara, 2010), among countless others. Mouse tracking has therefore become a powerful measure of multiple response activation with wide applicability across psychological science.

Microstructure of Decisions

From their beginnings, mouse-tracking studies sought to rule out dual-systems or stage-based models by demonstrating the continuity of trajectory-attraction effects, advancing the claim of a coactivation of competing processes that together coalesce into a stable response. Evidence in support of such alternative models would

Fig. 1. Mouse-tracking paradigms. In a depiction of a standard two-choice mouse-tracking paradigm (a), on each trial, participants click a start button at the bottom center (left), which reveals a stimulus (center). Participants then move the cursor and click on one of the responses in the top corners (right). There are many variants, including multichoice paradigms (e.g., four choices), sequences of stimuli (e.g., priming), or responses serving as stimuli themselves (as in panel b). Mouse tracking reveals decision microstructure. In conditions of conflict, dynamic models tend to predict simultaneously active processes (e.g., impulses toward eating unhealthy food vs. long-term goal of eating healthy food) that continuously self-organize into an explicit response. This leads to parallel attraction effects with a unimodal distribution. Dual-systems models tend to predict that a System 1 process occurs automatically (e.g., automatic impulse) on certain trials, and then a System 2 process intervenes (e.g., controlled goal). This leads to two subpopulations of trials: no-attraction trials (b, top) and extreme midflight correction trials (b, bottom), creating a bimodal distribution. An example of mouse tracking used as a time-course methodology is given in (c), adapted from Sullivan, Hutcherson, Harris, and Rangel (2015). The strength of the relationship (regression coefficients) between trajectories’ angle of movement and the relative tastiness and healthfulness of one food option over another is plotted as a function of time, separately for participants with low and high self-control ability. Vertical lines indicate onset of significant effects. Healthfulness was processed as early as tastiness for participants with high self-control; for those with low self-control, healthfulness was processed considerably later.
instead be reflected by discrete midflight corrections (e.g., automatic impulse toward unhealthy food vs. controlled correction in favor of healthy food; Stillman et al., 2017), such that an initial movement straight to one response is followed by a discrete corrective movement straight to the opposing response (Freeman et al., 2008, Study 3). Most realistic models of this kind assume that such stage-based corrections are probabilistic to some degree and do not necessarily take place on every trial; however, they generally assume that responses are being drawn from two subpopulations under conflict: some trials in which an “inappropriate” impulse must be squashed (e.g., “grab the cupcake—no, grab the banana!”) and other trials in which it is never activated in the first place (e.g., “grab the banana!”). Accordingly, it is the bimodal nature of trajectories’ response distribution that is often crucial in establishing a claim of dual-processing stages during mouse tracking (Fig. 1b; Freeman & Dale, 2013).

Indeed, such systematic flip-flopping of mouse-trajectories has now been taken as evidence supporting dual-systems or stage-based accounts of several aspects of language processing (Barca & Pezzulo, 2015; Dale & Duran, 2011; Tomlinson, Bailey, & Bott, 2013) and ambivalence (Schneider & Schwarz, 2017). However, it need not be an either–or question. For example, social categorization has been found to exhibit dynamic effects (e.g., for a masculine female face, parallel attraction to “male” before selecting “female”) but also dual-systems-like effects (e.g., for a masculine female face, initial attraction to “male,” followed by abrupt correction toward “female”), even within the same task (Freeman, 2014). Indeed, in some cases, models that take a formal dynamic-systems approach (e.g., for social categorization; Freeman & Ambady, 2011a) may even predict a trajectory pattern that appears in its stage-like sequence to be what one would expect from dual systems but instead reflects a rapid “phase transition” within a single dynamic system (Spivey, Anderson, & Dale, 2009). Thus, the important question may not be which pattern is observed for a given cognitive process, but rather under what conditions these different patterns manifest.

Of course, dynamic and dual-systems models are only two accounts, albeit popular ones, of cognitive processing. In the context of mouse tracking, another way to conceive of dual-systems models is that they predict two movement components, each of which inhabits its own spatiotemporal dynamics (e.g., early movement to top left and late movement to top-right; but see Spivey, Dale, Knoblich, & Grosjean, 2010). This logic can be broadened, however, to more complex models that predict the tandem operation of more than two systems or processes. For example, the Quad Model is a popular model of implicit social cognition that posits the existence of four distinct processes (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005); in certain tasks, one may expect four movement components that inhabit different parts of the spatiotemporal sequence, with four factors biasing the decision process at different times. Recently, researchers have taken several approaches to characterizing such trajectory components, including changes in trajectory direction or acceleration/deceleration (e.g., Dale & Duran, 2011; Dale, Roche, Snyder, & McCall, 2008), dimensionality-reduction approaches (Hehman, Stolier, & Freeman, 2015), and entropy analyses that identify high-speed movements and motor “breaks” (Calcagnì, Lombardi, & Sulpizio, 2017).

Moreover, these and other velocity and acceleration analyses may be used to measure additional characteristics of a decision process, such as instability. For instance, individuals with less interracial exposure were shown to exhibit more unstable dynamics and abrupt race-categorization shifts when categorizing racially ambiguous faces, an effect predicted by dynamic computational models (Freeman, Pauker, & Sanchez, 2016).

Such recent work shows that mouse tracking has the ability to uncover a microstructure of real-time decisions, revealing dissociable dynamics and processing components that can inform theory. Further research is certainly needed to link trajectory components to the specific theoretical processes under study, but at this early stage, it is clear that even when explicit responses or RTs may be similar, mouse tracking can qualitatively distinguish between vastly different dynamics driving those responses or RTs.

**A Matter of Time**

A critical advantage of mouse tracking is that it can sensitively expose millisecond-resolution timing information. Time-course analyses can provide powerful information about when specific factors are computed during an evolving decision or how specific processes temporally unfold. In one study, participants were asked on every trial to indicate their preference for one of two food options for which ratings of tastiness and healthfulness were also obtained. At each time point, the relationships between mouse trajectories’ angle of movement and the relative tastiness and healthfulness of one food option over another was examined. For people with high self-control ability, tastiness and healthfulness began correlating with mouse trajectories at the same time during the decision process; for people with low self-control ability, healthfulness began correlating with mouse trajectories considerably later in time than did tastiness (Fig. 1c). These results suggest
that food options’ tastiness has an early advantage over healthfulness in driving real-time preferences for people with a weak ability to control their impulses (Sullivan et al., 2015).

In cultural psychology, research has long suggested that people from collectivistic East Asian societies are more attuned to contextual associations than people from individualistic Western societies (Nisbett, Peng, Choi, & Norenzayan, 2001). In one study, American and native Chinese participants were presented with White and Asian faces embedded in scene environments; some environments were more stereotypically associated with White individuals, and some were more stereotypically associated with Asian individuals. While categorizing a face’s race, incongruent contexts led trajectories to veer toward the context-associated response, whereas congruent contexts led trajectories to approach the context-associate response more directly. Moreover, the onset and peak of these contextual effects occurred earlier for native Chinese participants than for American participants. Such results show that visual context exerts an earlier effect for individuals from collectivistic societies, suggesting they may have a greater preparedness to integrate contextual information into real-time perceptions (Freeman, Ma, Han, & Ambady, 2013). Additional studies have adopted a similar approach to explore how specific facial features drive gender, race, and age categorization with different temporal ordering (Freeman & Ambady, 2011b; Freeman et al., 2010). In the domain of subliminal perception, recent research showed that top-down attention both delays and prolongs the time course of subliminal semantic processing, revealing novel information about how attention interacts with nonconscious perceptual processes (Xiao & Yamauchi, 2017). Such work shows that mouse tracking is a powerful methodology able to dissociate the timing of different cognitive processes and, in some cases, link such timing to individual differences.

Compared with other time-sensitive measures, mouse tracking has distinct advantages and limitations. Eye tracking in choice tasks relies on discrete saccades (tracked as fast as they occur, about 3–4 times/s), whereas mouse tracking relies on continuous hand motion (tracked as fast as possible, typically about 70 times/s; Magnuson, 2005). Moreover, the eyes may only fixate on one response at a time, whereas the hand may inhabit in-between states among multiple responses. These qualities make mouse tracking uniquely suited to measure how a response evolves continuously over time, including any tentative attraction to other possible responses. That said, eye tracking may be more sensitive to preattentive processes before than after initiation of hand movement, and thus combining the two may be valuable (e.g., Quétard et al., 2016). Event-related potentials (ERPs), on the other hand, provide an index of when neural processing relevant to a decision process is modulated; mouse tracking provides a more direct measure of how multiple response alternatives accrue evidence to drive the decision over time. However, mouse-tracking timing information is most meaningfully interpreted in relative terms. For instance, finding that a given facial feature begins affecting mouse trajectories at 432 ms during age categorization but at 332 ms during gender categorization suggests that the facial feature starts playing a role in gender categorization 100 ms earlier (Freeman & Ambady, 2011b). However, placing meaning in 432 or 332 ms with respect to underlying cognitive processing is unwarranted. But how long it takes for a cognitive change to manifest in hand movement should be uniform throughout the decision process (Cisek & Kalaska, 2005, 2010); consequently, relative differences in mouse-tracking timing can powerfully reveal how much earlier or later different factors reign over a decision process with millisecond-level precision. With ERP, however, timing can be more meaningfully interpreted in absolute terms.

The Trajectory Forward

In short, mouse tracking has become a widely applicable measure of multiple response activations, capable of exposing component processes within real-time decisions and their time-course information. Indeed, compared with the gold standard of the RT, even the most straightforward mouse-tracking measures (e.g., deviation) are dissociable from RTs or general indecision. For instance, greater deviation effects predict stronger activation of conflict-monitoring regions even when RT is accounted for statistically (Stolier & Freeman, 2017), and there are numerous cases in which a deviation effect is observed without any RT effect (e.g., Stillman et al., 2017; Wojnowicz et al., 2009) or is uniquely predictive independent of RT (e.g., O’Hara, Carey, Kerwick, Crowley, & Dabrowski, 2016). Although delayed RTs may suggest parallel activation of multiple response options, there are numerous alternative explanations as well (e.g., slow evidence accumulation of a single response). Such dissociations become even clearer when considering three- or four-choice paradigms (e.g., Cloutier, Freeman, & Ambady, 2014; Tomlinson, Gotzner, & Bott, 2017) in which mouse tracking can reveal which specific response among multiple unselected alternatives simultaneously attracts participants’ decision trajectory; a delayed RT, on the other hand, may suggest that another response was activated in parallel but cannot distinguish which one it was. Moreover, as described earlier, mouse tracking can detect qualitatively distinct
decision microstructure and temporal dynamics that may be wildly different, even when two RTs are identical. A mouse-tracking methodology may therefore complement the traditional power of RTs and RT distributions to gain wholly new insights into a wide range of processes across the field.

An important direction under way is mapping mouse-tracking data to neural representation. Recent research has synchronized in-scanner mouse tracking with multivariate functional MRI (fMRI) to map the covert activation of specific responses with underlying neural representations. For instance, on a given trial, the extent to which the mouse trajectory was simultaneously attracted to the opposite-category response (e.g., “male” for a masculine female face) predicted the extent to which the neural representational pattern in face-processing regions was similar to that opposite category (Fig. 2; Stolier & Freeman, 2017). Or, because of automatic stereotype-driven expectations, the extent to which a participant was attracted to the “angry” category, even for a Black face displaying no anger, predicted the extent to which face-processing regions’ neural representational pattern was similar to the “angry” category (Stolier & Freeman, 2016). Combining mouse tracking with fMRI decoding approaches has tremendous potential because this paradigm can identify which levels of neural representation are affected by specific changes in a decision trajectory. Combining mouse tracking with ERPs could provide unprecedented information about decision-related...
timing. However, an important challenge for future work aiming to synchronize two such high-resolution time series would be to provide sufficient timing precision. ERP artifacts, which could easily arise as a result of motor movement, would need to be minimized (see also Fischer & Hartmann, 2014). Transcranial magnetic stimulation (TMS) has also been usefully combined with mouse tracking, allowing a causal test of a brain region’s role in resolving competitive dynamics, such as in semantic categorization (Hindy, Hamilton, Houghtling, Coslett, & Thompson-Schill, 2009).

Mouse tracking has often been referred to as an implicit measure, but future research should better establish the implicit nature of specific effects. Although there is some theoretical dispute regarding what constitutes an implicit measure, investigations into whether mouse-tracking effects are resistant to social desirability and reflect nonconscious or introspectively inaccessible representations would be important, as well as what roles played by automaticity versus control or activation processes versus validation processes (e.g., Gawronski, LeBel, & Peters, 2007). But the question itself may be regarded as problematic, in that it is akin to asking whether the RT method in general is implicit. The answer, of course, is that it depends on how one uses it. An RT effect in the context of semantic or evaluative priming or an implicit association test may be referred to as implicit, yet an RT effect in a time-unconstrained discrimination task may not. Further work is needed to rigorously test the implicit nature of mouse-tracking effects in particular task contexts, including the roles of methodological factors.

Once upon a time, research on motor control was dubbed the “Cinderella of psychology” (Rosenbaum, 2005), because the broader field neglects it, it was argued, believing that motor processes have little to do with the cognitive processes of interest (but see, e.g., Wolpert & Landy, 2012). Researchers’ shift from using discrete responses to using continuous hand movement is not only shows that cognitive and motor processes are far more coextensive than was previously believed but also opens the door to new avenues of investigation into a wide range of cognitive processes. The famous Milner and Goodale (1995) finding—that a brain-lesioned patient could not report visual attributes of a bar in front of her, but when her hand reached for the bar, its trajectory clearly reflected knowledge of those attributes—tells us that the moving hand may reveal more than we think. This may be an extreme example, but it makes crystal clear that in hand movement lies novel—sometimes covert—information about cognition. Indeed, this notion is becoming increasingly clear as the mouse-tracking approach to psychological science grows and evolves. Further work is certainly needed to more deeply understand the link between specific hand-movement parameters and theoretical constructs, but if the past decade is any example, the trajectory looks to be on the rise.

**Recommended Reading**

Freeman, J. B., & Ambady, N. (2010). (See References). An in-depth description of freely available specialized software for conducting and analyzing mouse-tracking studies, including discussion of measures and methodological validation.


**Declaration of Conflicting Interests**

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

**Funding**

This work was supported in part by National Science Foundation (NSF) Division of Behavioral and Cognitive Sciences (BCS) Grant 1423708, NSF BCS Grant 1654731, and National Institute of Mental Health Grant R01-MH112640.

**References**


