Graded motor responses in the time course of categorizing atypical exemplars

RICK DALE, CAITLIN KEHOE, AND MICHAEL J. SPIVEY
Cornell University, Ithaca, New York

The time course of categorization was investigated in four experiments, which revealed graded competitive effects in a categorization task. Participants clicked one of two categories (e.g., *mammal* or *fish*) in response to atypical or typical exemplars (e.g., *whale* or *cat*) in the form of words (Experiments 1 and 2) or pictures (Experiments 3 and 4). Streaming *x*, *y* coordinates of mouse movement trajectories were recorded. Normalized mean trajectories revealed a graded competitive process: Atypical exemplars produced trajectories with greater curvature toward the competing category than did typical exemplars. The experiments contribute to recent examination of the time course of categorization and carry implications for theories of representation in cognitive science.

The past few decades have seen a transition from classical set theoretic accounts of categorization, wherein cognitive processes and categories are seen as discretely bounded with unique membership, to the development of nuanced theories of fuzzy categories and their interrelationships. For example, Reed (1972), Rosch (1973, 1975), and Rips, Shoben, and Smith (1973) challenged the classical conception of category structure early on, revealing that more graded semantic space effects are readily observable and that these cannot be easily accounted for in terms of classical conceptions (e.g., Bruner, Goodnow, & Austin, 1956; Collins & Quillian, 1969). The emerging probabilistic prototype framework has also faced challenges from alternative accounts. It has often been argued that an exemplar-based theory of categories accounts for a wider set of data (Medin & Ross, 1989; Medin & Schaffer, 1978; Nosofsky, 1988, 1992; for a recent account, see Storms, 2005; for an opposing perspective, see Smith, 2002). Aspects of these similarity-based accounts, both prototype and exemplar, have also been suggested to figure into a theory-based theory of category structure in our cognitive system (Medin, 1989; Murphy, 2002; Murphy & Medin, 1985; see, e.g., Burnett, Medin, Ross, & Blok, 2005; Heit, 1994; Lynch, Coley, & Medin, 2000). Despite this branching into multiple theoretical directions, experimental methodologies had remained largely unchanged. The upshot, until recently, has been that the *time course of processing* in categorization had been underexplored, in comparison with the development of theories regarding the static representation of category knowledge.

There is a growing body of research devoted to this question. The time course of categorization phenomena has begun to receive some attention, from perceptual categorization (Ashby, Boynton, & Lee, 1994; Lamberts, 1995, 2000; Nosofsky & Palmeri, 1997) to categorical

perception of speech (McMurray & Spivey, 2000; McMurray, Tanenhaus, Aslin, & Spivey, 2003) and lexical processing (e.g., Cree, McRae, & McNorgan, 1999). Most of these studies have made use of reaction time measures. For example, Lamberts (2000) introduced an information accumulation account of speeded classification of objects (see also Ashby et al., 1994; Nosofsky & Palmeri, 1997). The goal of these investigations of the time course of categorization has been to supplement research that has generally placed more emphasis on the outcome of the process, rather than on its continuous temporal dynamics (Nosofsky & Palmeri, 1997). Decision models of categorization, such as Lamberts's (2000, 2002) and Nosofsky and Palmeri's, uncover the time course of the process leading up to the buttonpress response.

It is possible, however, that the motor output of the decision process itself exhibits graded effects over time, providing further information about the continuous nature of the process. For example, McMurray et al. (2003) used eye movement data to investigate the graded temporal dynamics of speech sound classification over the course of several hundred milliseconds. Similarly, Nederhouser and Spivey (2004) used eye movement patterns to show competition between alternative taxonomic classifications of atypical animals. When given a toy whale to categorize by dropping it into either a "fish" bucket or a "mammal" bucket, participants frequently fixated the "fish" bucket first, before then fixating the "mammal" bucket to guide hand movement. These kinds of eye movement data represent *microdecisions* that are tentative intermediate emissions from a categorization process that is extended in time (Richardson, Dale, & Spivey, 2007; see also Rehder & Hoffman, 2005).

Motor responses of this kind epitomize what has often been considered to be *output* from cognitive processes: The

---

R. Dale, radale@memphis.edu

outcome of a pipeline, from perceptual processes through the association cortex and into premotor regions, then collapses into individual manual and oculomotor choice behaviors. More recent work on manual and oculomotor movements has suggested that these manual processes are contiguous with cognitive processes (Gold & Shadlen, 2001; Shin & Rosenbaum, 2002). Eye movement patterns themselves offer an unusually early glimpse into partially active microdecisions in motor output (cf. Gold & Shadlen, 2000; Magnuson, 2005). Unfortunately, the ballistic quality of most saccades prevents them from being able to exhibit effects of partially activated representations that are truly graded. On any given trial, participants either fixate the competing object or they do not. Eye movement data are usually not able to show continuous attraction effects within a trial (but see Doyle & Walker, 2001). In contrast, nonballistic arm movements regularly involve a curvature that reveals continuous attraction effects (Goodale, Pélisson, & Prablanc, 1986). In fact, continuous manual motor output has been studied for the very purpose of providing a variety of clues about graded underlying cognitive processes within individual trials. For example, Abrams and Balota (1991; see also Coles, Gratton, Bashore, Eriksen, & Donchin, 1985) used an arbitrary manual response task (pulling a handle to the right or the left) in a lexical decision and recognition memory test. The frequency of lexical items and the strength of recognition memory were strongly associated with a shorter time being spent pulling the handle to its limit (i.e., greater force and velocity of pull), independently of the latency of response onset. The authors recommended that anticipation components of response preparation and force and velocity parameters of motor execution should figure into our understanding of cognitive processing. Similarly, Spivey, Grosjean, and Knoblich (2005) used computer mouse trajectories to study spoken word recognition. On individual trials, participants heard instructions, such as "click the candle," and selected one of two objects in two corners of a computer monitor. Trials on which the two objects had similar-sounding names (e.g., a candle and a candy) revealed mouse movement trajectories that exhibited significant attraction toward the competing object, in comparison with control trials (e.g., where the objects were a candle and a spoon). Computer mouse trajectories thereby revealed a continuous dynamic partial activation of multiple competing representations during real-time spoken word recognition (see also Allopenna, Magnuson, & Tanenhaus, 1998, for related eye movement results).

The time course of categorization is a system that can be a suitable model for understanding the interaction among different processes as cognition unfolds (e.g., memory and attention; see Lamberts, 2000, for a discussion). For example, feature-sampling models of the time course of categorization contribute to an understanding not simply of categorization outcomes, but also of predictive or explanatory systems accounting for reaction time measures of categorization and how these measures reflect the perceptual information accumulation used in categorization (e.g., Lamberts, 2000, 2002). In the present study, four experiments revealed that categorization as a cognitive process flows into, rather than collapses onto, the effectors responsible for manual action.

The goal of the present study was to demonstrate, via continuous computer mouse trajectories, that graded representations of the kind discussed above are also involved in the real-time taxonomic categorization of animal names and animal pictures. It has been suggested that the temporal dynamics of matching a category exemplar to one of its potential categories exhibits a degree of *granularity* (Miller, 1982; Miller & Ulrich, 2003) that may raise problems for a discrete representational account of the categorization process. We present four experiments in which continuous manual motor output was tracked during lexical and perceptual categorization of atypical exemplars. This work adds a new methodology to the study of the time course of categorization and contributes to further specifying the processes by which the mind settles onto one categorical response versus another. These experiments illustrate how the gradual accumulation of evidence for a given category is not composed solely of that category's representation transitioning from zero activation to full activation. As in many complex dynamical systems, there is also competition from alternative category representations that are partially active at the same time. The trajectory of the categorization process reveals that multiple categories (attractor basins) are nearly visited as the system eventually settles into a unique outcome-based response, much like a dynamical system continuously traversing its high-dimensional semantic space over time.

EXPERIMENT 1

Typicality has become one of the most thoroughly studied aspects of categories and concepts since the transition from classical perspectives (Medin, 1989; Murphy, 2002). The empirical drive toward nonclassical accounts of categories has been guided by studies of membership and typicality judgments, in which corresponding reaction time measures have been used (e.g., McCloskey & Glucksberg, 1978; Rips et al., 1973; Rosch, 1975). Category members deemed more typical are recognized more quickly (Rips et al., 1973) and more consistently (McCloskey & Glucksberg, 1978), have many features in common (Rosch & Mervis, 1975), and can even result in facilitated language comprehension (Garrod & Sanford, 1977; see Murphy, 2002, chap. 2, for a review of these and other robust results). Recent research on categories and concepts has continued this emphasis (e.g., Burnett et al., 2005; Estes, 2003; Murphy & Ross, 2005; Op de Beeck & Wagemans, 2002; Smith, 2002; Verbeemen, Storms, & Verguts, 2003). For example, Burnett et al. revealed that typicality ratings of various fish species are guided by knowledge-level expectations regarding fish, rather than by a simple notion of category centrality (see also Lynch et al., 2000). Smith recently used expected patterns of typicality across category members as a focal point for theoretical debate, such as the effectiveness of exemplar-based perspectives. Typicality is thus central to our categorizing capacities but is also useful as a crucial measure for comparing opposing theories.

Table 1
Atypical and Typical Animals/Words in the Experiments,
With the Response Options Given to the Participants (in Parentheses)

Atypical	Typical	
Eel (*fish*; reptile; **mammal**)	Hawk (*bird*; reptile)	Cat (*mammal*; reptile)
Whale (*mammal*; fish; **bird**)	Dog (*mammal*; insect)	Sparrow (*bird*; mammal)
Sea lion (*mammal*; fish; **reptile**)	Horse (*mammal*; bird)	Goldfish (*fish*; amphibian)
Penguin (*bird*; fish; **mammal**)	Shark (*fish*; mammal)	Salmon (*fish*; mammal)
Butterfly (*insect*; bird; **reptile**)	Alligator (*reptile*; mammal)	Rattlesnake (*reptile*; amphibian)
Bat (*mammal*; bird; **reptile**)	Rabbit (*mammal*; reptile)	Lion (*mammal*; fish)
	Chameleon (*reptile*; insect)	

Note—The correct categories are given in italics. Noncompeting labels used as options on atypical trials in Experiments 2 and 4 are given in boldface.

In the following four experiments, we utilized an animal stimulus set (words and pictures) whose elements differ in typicality. We used a small stimulus set of well-known animals that are either highly typical members of a category (*dog* as *mammal*) or widely regarded as atypical (*penguin* as *bird*). In this first experiment, we recorded continuous manual responses in a categorization task. Participants performed a simple judgment: After seeing an animal word, the participants clicked on one of two category labels, indicating to which category the animal belonged. Some trials involved animals of an atypical nature—such as a whale, which has several properties that suggest another category label (*fish*), potentially causing competition with the correct label (*mammal*). Just as in Spivey et al. (2005), echoes of a competitive categorization process should be reflected in continuous mouse movements. In other words, when *whale* was categorized as *mammal*, the participants' mouse trajectories should gravitate toward the competing category (*fish*) more so than when a typical exemplar, such as *cat*, was categorized.

Method

Participants. Forty-one undergraduate participants signed up for this experiment for extra credit in their psychology classes. All the participants in this and the subsequent experiments used their right hand to perform the task.

Materials. Word stimuli were presented using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The participants were faced with the task of choosing the appropriate category for each animal word, presented in text on the monitor. These categories included pairings of the following: *mammal*, *reptile*, *bird*, *fish*, *amphibian*, and *insect*.

There were two within-subjects conditions. On control trials, the animal words were typical category members (e.g., *cat* as *mammal*). On the experimental trials, they were atypical category members (e.g., *penguin* as *bird*), and both the correct category name and the featurally similar competing category name were presented (*bird* and *fish*, respectively). Although we assume that typicality of category membership is a graded parameter, the two-condition design of these experiments required that we keep within-group variation among stimulus items to a minimum (particularly in the atypical animal condition). Therefore, only 6 highly atypical animals were used in the atypical condition, with other, merely moderately atypical animals being excluded from the design. Since typical category members are easier to generate, 13 highly typical animals were used for the control condition (see Table 1).

Procedure. The participants were presented with two different animal category names, randomly assigned to one of the upper corners of a computer screen. After a 2,000-msec moment in which to view the category options, the text "Click Here" appeared in the bottom center of the screen. The participants were instructed to click first on that text and to wait for an animal word to appear in its place, then to click on the upper (left or right) category name that was appropriate for that animal. The participants were provided with 3 practice trials before beginning the 19 target trials. All the trials were presented in random order.

It was predicted that in the experimental trials, mouse movement trajectories would show evidence of competition between the categories. This competition would be revealed by an analysis of mouse movement trajectory divergence: Atypical animal trials should have movement trajectories that reveal a slight bias toward the competing category (e.g., with *whale*, a slight attraction toward the category *fish*), when compared with control trials. Data for testing this prediction were collected by recording *x*- and *y*-coordinates of mouse movement trajectories. Due to occasional skipped samples, PsyScope's sampling rate averaged approximately 42 Hz. As a result, each trial collected about 40–80 mouse position data points.

Data analysis. Numerous analyses were conducted on these rich trajectory data. First, to enable averaging of full trajectories from multiple trials, all trajectories were normalized to 101 time steps and were translated to begin at an *x*, *y* coordinate of (0, 0). These *time-normalized* trajectories could be compared between typical and atypical conditions. Divergence of the two averaged trajectories was established by significant differences between the *x*-coordinates.

This first analysis provided information regarding the overall shape of the trajectories for both trial types. In addition to this time-normalized analysis, a *space-normalized* analysis was run in which the beginning and end coordinates of each trial were normalized to (0, 0) and (1, 1), respectively. Real-time information was retained by computing *x*, *y* coordinates as they traveled from 0 to 1 in time bins of 0–500, 500–1,000, and 1,000–1,500 msec. This provided a window onto the movement in real time from start to finish of a trial and included enough data in each time bin to permit an additional statistical test of the difference between trial conditions by subjecting these bins to a repeated measures ANOVA.

Finally, a number of properties of the trajectories were computed and compared between conditions: mouse movement initiation time, movement duration, total categorization response time, distance traveled in pixels, and direction in degrees of the first pair of mouse movement samples. This last measure, initial movement direction in degrees, offers insight into early stages of the trial. For example, it is possible that the participants strategically moved the mouse cursor along the vertical on all the trials before turning in the direction of the target category. Angle information on initial movement will reveal whether this was the case. Such a strategy would cause that initial movement from (0, 0) not to be significantly different from 0º off the vertical in both the typical and the atypical conditions.

These final measures were supplemented by two further analyses: (1) tests of bimodality in the distribution of trajectory curvatures and (2) a time series analysis to explore the complexity of the resultant trajectories. As has been discussed in Spivey et al. (2005), any pattern of competition seen in atypical trials might, in fact, have been the result of an averaged bimodal distribution. If half of the correct

trials involved movement straight toward the target and the other half involved movement straight toward the competitor, followed by a corrective movement toward the target, the average of all the trials could appear as a graded curvature toward the competitor—when, in fact, it would be better described as the result of a number of trials with rather discrete errors that were corrected midflight. To approach this problem, each trial's area (in pixels) was computed between the actual trajectory and a straight line connecting the start and the end points. A distributional analysis of this area should show bimodality if the averaged trajectory's apparent graded curvature was actually due to occasional discrete errors that got sharply corrected partway through the movement.[1] Finally, sample entropy (Richman & Moorman, 2000) was used to discern whether atypical trajectories were more *complex* than typical trajectories (see Feldman & Crutchfield, 1998, for a discussion of complexity measures and their advantages and problems). If the competing category in the atypical trials was acting as a substantial second attractor on the dynamics of manual output, this stress should be evident as less smoothness in the trajectories. Atypical trials should show more complexity, because two attractor forces were acting on them. Sample entropy gives a larger value the more complex or irregular a time series. Further detail on this measure will be presented below.

Results

The participants categorized atypical exemplars with 90% accuracy and typical exemplars with 95% accuracy, a significant difference ($p < .05$). Only correct trials will be used in the following analyses. All incorrect trials were removed from analysis in this and the subsequent experiments.

Time-normalized analysis. In this analysis, t tests were conducted to compare the difference between the x-coordinate values for typical and atypical trials to zero, at each of 101 interpolated time steps throughout the trajectories (see Figure 1A). Rightward and leftward responses were pooled to maximize statistical power. The relevant null hypothesis was that the difference between the atypical and the typical trajectories' x-coordinate at any given time step (out of 101) should be 0. As a more conservative test, in this and the subsequent experiments, a reliable divergence was defined as a mini-

mum of 8 consecutive time slices in sequence, in which the normalized trajectory differences were significant at a criterion of $p < .05$. This criterion was established by performing a bootstrap of 10,000 simulated experiments of the same mean and standard deviation (see the Appendix). Significant divergence between trajectories is thus observed when there is a substantial sequence of consecutive significant t tests between atypical and typical x-coordinates. In this experiment, the trajectories exhibited significant differences in x-coordinates for 38 consecutive time slices, from the 47th to the 85th time step ($ps < .05$).

As an additional statistical test, we computed pooled bins from these time-normalized trajectories and conducted a 2 (typical or atypical) × 3 (bins of 1–33, 34–67, or 68–101 steps) repeated measures ANOVA. This test revealed a strong effect of trial type [$F(1,40) = 21.8$, $MS_e = 1,824.8$, $p < .001$], a main effect of bin [$F(2,39) = 861.1$, $MS_e = 1,916.5$, $p < .001$], and a significant interaction [$F(2,39) = 6.2$, $MS_e = 783.7$, $p < .01$]. To reveal what portions of the trajectory were exhibiting this divergence, we conducted planned comparisons between trial types within each bin. These showed a significant difference between trial types in the second and third bins ($ps < .001$). Thus, by the second and final third of the time-normalized trajectories, atypical categorization exhibited a significant divergence in x-coordinate from typical categorization.

Space-normalized analysis. Figure 1B shows a graph of the trajectories, from leftward and rightward movements, in separate conditions in terms of time bins and normalized pixel coordinates. Whereas the previous analysis preserved raw space and normalized time into 101 bins, this analysis preserved real time and normalized spatial coordinates of the mouse from (0, 0) to (1, 1) and pooled these values into three time bins: 0–500, 500–1,000, and 1,000–1,500 msec. Once again, absolute left and right values of the x-coordinate were pooled for statistical comparison. A similar 2 (trial type) × 3 (time bin) repeated measures ANOVA was used. This re-
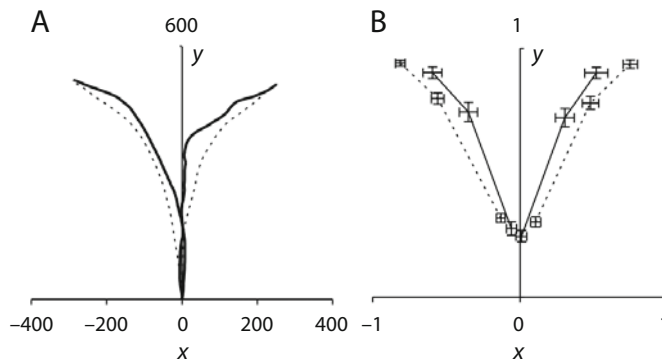

Figure 1. (A) Mean time-normalized mouse movement trajectories in Experiment 1, separated for left- and rightward responses to the correct target. Atypical trials (solid lines) indicate an attraction toward the competing category on the opposite side of the computer screen in the 101 time steps. (B) Mean mouse movement trajectories in space-normalized analyses, with x, y coordinates traveling from (0, 0) to (−1/1, 1) in time bins of 0–500, 500–1,000, and 1,000–1,500 msec. Atypical trials (solid lines) reveal slower movement toward the target.

vealed strong effects of trial type [$F(1,37) = 52.3$, $MS_e = 0.031$, $p < .001$] and time bin [$F(2,36) = 168.0$, $MS_e = 0.041$, $p < .001$] and a significant interaction between time bin and trial type [$F(2,36) = 5.2$, $MS_e = 0.030$, $p < .05$].[2] Once again, planned comparisons were run between types at each time bin to detect what portion of the trajectories had significant divergence. The difference between the trial types was significant at every time bin ($ps < .01$).

Additional measures. The movement duration for atypical trials was 1,573 msec, as compared with 1,338 msec for typical trials [$t(40) = 3.8$, $p < .001$]. Total categorization time (from animal word onset to final mouse click) for atypical animals was greater than that for typical trials [1,997 and 1,807 msec, respectively; $t(40) = 3.0$, $p < .01$]. Total distance traveled also differed significantly [atypical, 894 pixels; typical, 831 pixels; $t(40) = 3.0$, $p < .01$]. Curiously, movement initiation latency showed an opposite pattern but was only marginally significant [atypical, 424 msec; typical, 470 msec; $t(40) = -1.8$, $p = .09$]. An analysis of the initial direction of the mouse trajectory, from position (0, 0), extracted a measure in degrees from the vertical (from the start-click event). The typical trials exhibited a significant positive angle (toward the target) in a one-sample t test (null hypothesis that degrees are different from 0), with a value of 6.0º [$t(40) = 2.6$, $p < .05$]. Atypical trials, in fact, exhibited a slightly negative angle toward the competitor (-0.99º), but this was not significant.

Item-based repeated measures ANOVAs were also run, using these same measures. Both movement duration and distance were marginally significant in the expected direction [$Fs(1,17) = 3.2$ and 2.0, $MS_es = 72,125.3$ and 7,685.4, $ps = .09$ and .08, respectively], with mean atypical trajectories that were longer in time and in distance to reach their target. Neither movement initiation latency nor total response time was significant. These results may be difficult to interpret, due to the size of the stimulus set. However, the marginal significance suggests that the effects in the expected direction might be significant with a larger set, whereas the movement initiation latency would not be likely to be so [$F(1,17) = 0.8$, $p = .38$]. Initial angle of movement from the vertical, in fact, showed a marginal difference between conditions in an item-based analysis [$F(1,17) = 3.4$, $MS_e = 54.3$, $p = .08$], with the typical trials again showing the only significant difference from 0º at 6.6º [$t(12) = 3.1$, $p < .01$].

Two additional analyses were used to reveal more characteristics of the trajectories. First, as in Spivey et al. (2005), we explored the nature of the distribution of curvature across all the trajectories. The pattern of divergence (from a straight line) on atypical trials could, in principle, have simply been a reflection of an averaged bimodal distribution. In contrast, there was no theoretical reason to expect such bimodality on the typical trials. The trial data used were the area between the actual trajectory and a straight line from (0, 0) to the final click (see Spivey et al., 2005). All trials across participants were used in order to have sufficient numbers for the statistical test. All area values were converted into z-scores and were subjected to distributional analyses. A Kolmogorov–Smirnov test for a
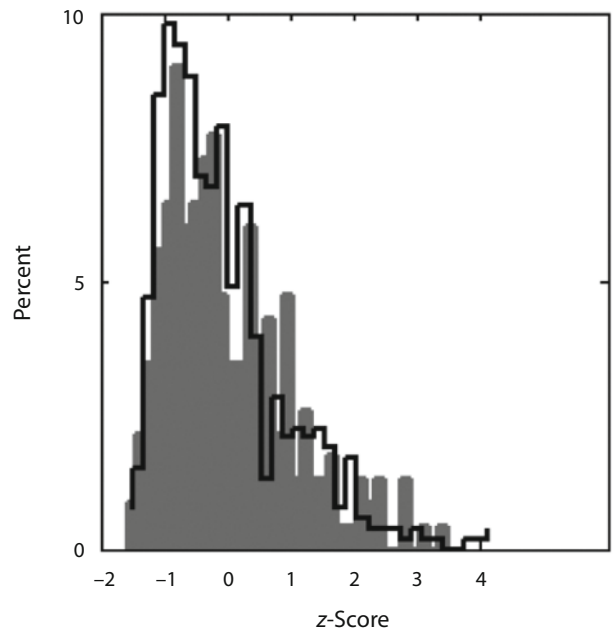

Figure 2. Percentage distribution histogram of z-scores for area in pixels on atypical (solid gray distribution) and typical (black outlined distribution) trials, superimposed. Distribution analysis reveals no significant difference between these distributions.

difference in distribution revealed no difference between typical and atypical trajectories ($\chi^2 = 1.9$, $p = .79$; see Figure 2). In addition, computed bimodality coefficients for both the typical and the atypical trials were in the unimodal range of $<.555$ (see Spivey et al., 2005). Although the typical (.536) and the atypical (.535) trials were close to this cutoff, at which concerns about bimodality may arise, the values were, in fact, almost identical. These results indicate that even if bimodality were a concern, it would not be exclusive to the atypical trials.

Finally, we explored the complexity of the trajectories, using a time series analysis giving an entropy-based value.[3] Sample entropy (Richman & Moorman, 2000) is computed by comparing windows of size m from a time series. A given window is said to be similar to another if their distance is less than some value r, known as the tolerance. This is given by the following equation, in which two windows (w_1 and w_2) of size m have a distance equal to the maximal difference between paired elements composing them, with $w_1(i)$ representing the ith element of window w_1:

$$d(w_1, w_2) = \max |w_1(i+k) - w_2(j+k)|, \text{ with } k = 0 \ldots m-1.$$

Sample entropy is then computed by taking the difference between averaged natural logarithms of counted similarities at size m and $m+1$. If, when window size is increased, the average number of similarities is similar, this value will be low. Conversely, with a higher difference between counts in m and $m+1$, it is less likely that windows similar to each other with size m will again be similar at $m+1$, indicating more irregularity in the time series. For this analysis, we used the time series of a trial's normalized
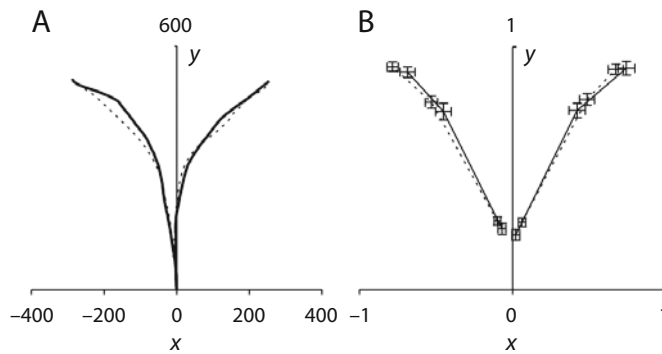
Figure 3. (A) Mean time-normalized mouse movement trajectories in Experiment 2, separated for left- and rightward responses to the correct target. Atypical trials (solid lines) reveal a diminished attraction toward the competing category. (B) Space-normalized time bins show no significant differences between trial types (atypical trials represented by solid lines).

x-coordinate fluctuations. This time series represents the extent to which horizontal movement is fluctuating toward one category label or back to the other. A range of window sizes was used (2–10), along with a tolerance of .2 multiplied by the standard deviation of all the *x*-coordinate fluctuations $(x_{t+1} - x_t)$ in normalized trajectories. Across all values for *m*, mean sample entropy was numerically greater (indicating higher complexity) for atypical trajectories than for typical trajectories. These differences were statistically significant or marginally significant when *m* had the values 3 through 6.

Discussion

The results indicate that the process of categorization exhibited nonlinear time course effects in mouse movement trajectories. The trajectories revealed significant attraction toward the competing category name in the atypical animal condition, when compared with the typical animal condition. Interestingly, the movement initiation latency did not generate significant differences, although measures based on the manual motion itself revealed spatial divergence between atypical and typical trajectories and a difference in categorization time. We therefore conclude that the manual output from the categorization process is reflective of a cognitive temporal dynamics wherein the mapping of exemplar to category evolves nonlinearly over time (cf. Lamberts, 2000).

EXPERIMENT 2

The same animal names were used in this second experiment. However, in the atypical animal condition, the alternative (incorrect) response option was *not* the taxonomic class that shared similar properties with the label's referent. To explore whether the graded attraction effect in computer mouse trajectories is, in part, due to competition between strongly activated categories or whether the manual curvature is solely a result of uncertainty in the match between atypical animal words and their correct categorization, we altered the response options for those trials.

For example, whereas in the previous experimental trials, the categories *fish* and *mammal* were used for the atypical animal *whale*, in the present experiment, an alternative (incorrect) category was used that did not share features with the animal word (e.g., *bird* and *mammal* for *whale*). If the mouse trajectories for the atypical trials showed equal curvature toward the incorrect response option (as in Experiment 1), this would suggest that the curvatures we observed were due merely to a slow accumulation of evidence for the atypical animal's being categorized in its correct taxonomic class; if this curvature was reduced, it would suggest that the differences in Experiment 1 were due, in part, to a dynamic attraction effect exerted by the featurally similar competing category.

Method

Participants. Forty-one undergraduate participants in this experiment received extra credit for psychology classes.

Materials and Procedure. In this experiment, the same technique and same words were used as in the first, but the atypical animal condition no longer involved competing categories. Instead, the correct category was paired with a noncompeting one (e.g., *whale = mammal* or *bird*). The control trials were the same as those in Experiment 1 (see Table 1).

Results

The participants categorized atypical exemplars with 92% accuracy and typical exemplars with 98% accuracy ($p < .01$). Again, only correct trials were used in the analysis.

Time-normalized analysis. Normalized trajectories revealed a single sequence of eight *t* tests ($p\text{s} < .05$), from the 72nd to the 79th time step, showing atypical − typical *x*-coordinate differences that were reliably greater than zero (see Figure 3A). Although significant, this brief divergence between atypical and typical trajectories is a substantially diminished sequence, in comparison with the previous experiment's 38 consecutive time steps with significant differences. Again, as an additional statistical test, pooled bins from these time-normalized trajectories were used in a 2 (typical or atypical) × 3 (1–33, 34–67, or 68–101 time

steps) repeated measures ANOVA. There was no significant effect of trial type [$F(1,40) = 2.2, MS_e = 1,341.2, p = .14$], but there was a significant main effect of bin [$F(2,39) = 1,147.7, MS_e = 1,289.0, p < .001$] and a significant interaction [$F(2,39) = 4.0, MS_e = 617.4, p < .05$]. Although there was no main effect of trial type, comparisons did show a significant difference between the typical and the atypical conditions at the third time bin ($p < .05$).

Space-normalized analysis. There was a main effect of time bin [$F(2,35) = 340.1, MS_e = 0.035, p < .001$], but no significant effect of trial type or interaction between trial type and time bin (see Figure 3B).

Additional measures. No significant differences were found in movement initiation latency [416 msec atypical vs. 403 msec typical; $t(40) = 0.6, p = .6$]. However, there were significant effects of total response time [1,840 vs. 1,718 msec; $t(40) = 2.5, p < .05$], movement duration [1,424 vs. 1,315 msec; $t(40) = 3.8, p < .05$], and distance traveled [836 vs. 780 pixels; $t(40) = 2.3, p < .05$]. Initial movement angle showed no significant effects. Interestingly, both were positive toward the target category (3.1° vs. 2.25° on atypical and typical trials, respectively). In item-based repeated measures ANOVAs, only total response time and distance traveled were significant [$Fs(1,17) = 4.7$ and 9.9, $MS_es = 20,023.7$ and 1,990.5, respectively; $ps < .05$]. Neither movement initiation latency nor movement duration was significant. There were no significant results for initial angle.

As before, a Kolmogorov–Smirnov test did not show a significant difference in the distribution of trajectory curvatures between the typical and the atypical trials ($\chi^2 = 1.8, p = .80$). Bimodality coefficients were again similar and below .555 (.546 and .549). Finally, sample entropy analyses did not show any significant or marginally significant differences, using window sizes of 2–10.

The greatly reduced trajectory attraction effects during categorization with noncompetitive alternative categories suggest that the spatial attraction effects in Experiment 1 were not due merely to the inherent atypicality of the

Figure 4. Mean total divergence (atypical trial − typical trial *x*-coordinates) for Experiments 1 and 2. Experiment 1 (solid line) exhibits significantly more divergence.

label's referent. The featural similarity between the animal and the (incorrect) alternative category—present in Experiment 1 and absent in Experiment 2—played a substantial role in eliciting curvature in the computer mouse trajectory. When differential divergence between the average time-normalized trajectory for atypical animals and the average trajectory for typical animals is plotted over time for Experiment 1 and for Experiment 2 (Figure 4), the competitive alternative categories in Experiment 1 clearly show a much stronger attraction effect than do the noncompetitive alternative categories in Experiment 2. These trajectories differ significantly for 19 time slices, from the 57th to the 75th time steps ($ps < .05$). In addition, when atypical animal trials from these experiments are compared, using space-normalized time bins, there is a main effect of experiment [$F(1,75) = 5.8, MS_e = 0.096, p < .05$] and an interaction between experiment and bin [$F(2,74) = 3.4, MS_e = 0.039, p < .05$]. These differences suggest that Experiment 1's effects were strengthened by the presence of competitor labels.

Discussion

Interestingly, and perhaps surprisingly, noncompeting alternative category labels in Experiment 2 still induced some significant, albeit small, attraction. For example, when faced with possible responses of *mammal* and *bird* for the exemplar *whale*, the participants' trajectories in Experiment 2 showed some significant attraction toward the putatively noncompeting label *bird*. Why should this be? The answer may lie in the between-category similarity.

In addition to considering the match between an exemplar and a potential category, we must also consider the similarity between categories themselves (cf. Storms, 2005). Although one may initially assume that the noncompeting alternative categories in Experiment 2 should remove the patterns of competition seen in Experiment 1, there actually remains substantial between-category similarity among the correct and alternative category responses. For example, in addition to *whale*'s being a good match for *mammal* and a partial match for *fish*, it also has a moderate amount of fit to *bird*. Both whales and birds move, breathe, and eat. Both have eyes, skin, muscles, bones, lungs, a brain, a heart, and so on. Thus, even our "noncompeting" category labels in Experiment 2 still shared a number of similar features with the target stimulus, since they were all in the superordinate-level category of *animals*. Even such minor similarity appeared to be sufficient to produce some mild attraction effects in the mouse movement trajectories. On the basis of the combined results of these first two experiments, we suggest that mouse movement curvature reflects a competition process wherein partially active categories pull the state of the cognitive system toward their respective attractor basins. Dynamic properties of the resulting nonlinear mental trajectory are emitted in the continuous manual output of computer mouse movement. As in many classic categorization studies (e.g., Rips et al., 1973; Rosch, 1973), we were able to induce these effects with atypical category members presented in lexical form.
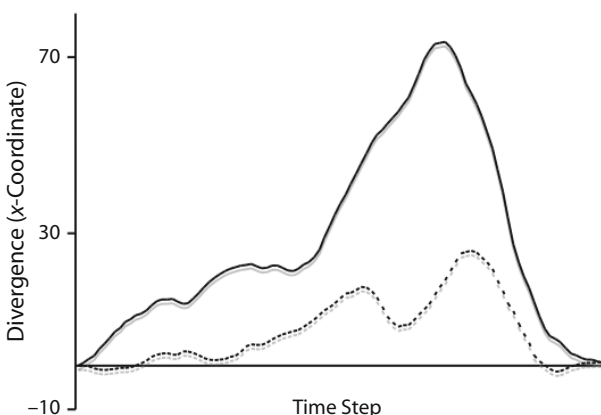
Indeed, the initial process of visual word recognition may also exhibit some attractor dynamics of its own (e.g., Hinton & Shallice, 1991; McRae, de Sa, & Seidenberg, 1997; Rueckl, 2002). Because initial processing in the trials of both Experiments 1 and 2 involved lexical recognition, the competitive process following them is attributable to the process of categorization. Some additional dynamics therefore take place following this word recognition phase, in the time course of mapping lexical processing onto a categorization response. These experiments reveal that even the manual output of this mapping reflects the partially active representations underlying it.

EXPERIMENT 3

In the next two experiments, participants categorized *pictures* of our animal set, rather than lexical items. In numerous studies, the distinction between the processing of pictures and of words and the relationship between this processing and a proposed core conceptual representation underlying them have been considered. Early on, for example, Snodgrass (1984) argued that pictorial and verbal codes have separate functions and distinct neurophysiological realizations, rather than being mere surface forms for a central underlying concept representation (see also Paivio, 1991). More recently, numerous studies have explored the efficiency with which picture versus word stimuli are categorized, with most showing a small but significant advantage for pictorial stimuli (e.g., Job, Rumiati, & Lotto, 1992; Snodgrass & McCullough, 1986; Viswanathan & Childers, 2003). For example, Viswanathan and Childers demonstrated a processing advantage for pictorial stimuli in a task in which participants judged whether two stimuli were members of the same category. The authors suggested that visual stimuli induce simultaneous processing of category-relevant visual features *and* of their categories or concepts. Thus, the accumulation of featural match information may be faster and more robust with pictures than with lexical presentation alone, which relies only on conceptual information acquired through orthographic input.

This discussion suggests a few predictions for Experiments 3 and 4 with visual stimuli. First, categorization on typical animal picture trials will likely be quicker and more efficient than that on typical animal lexical trials. Previous research has suggested this processing advantage for visual exemplars. In addition, because processing of visual stimulus features occurs simultaneously with conceptual activation, there may, in fact, be more competition for atypical picture trials (Experiment 3) than there was for atypical word trials (Experiment 1). Visual features of atypical animals often provide some partial match to competing categories. For example, whale images offer visual information that greatly resembles fish. The atypical functional features of the concept *whale*, along with potentially misleading perceptual information, could increase the attraction exhibited in mouse movement trajectories (Experiment 3 in comparison with Experiment 1). Finally, this prediction regarding atypical competitive trials with pictures and the previous discussion of moderate

similarity even between noncompetitive categories and exemplars suggest that substantial competition should be seen in conditions with noncompetitive category alternatives when pictorial stimuli are used (Experiment 4 in comparison with Experiment 2). If misleading visual features cause more equivocal information to be partially active on atypical animal trials, even when the category label is supposedly noncompeting, it should take more time for the correct category to reach maximal activation.

Method

Participants. Forty-five Cornell undergraduates participated in this experiment in order to receive extra credit for their psychology classes.

Materials and Procedure. The picture stimuli were color images of approximately equal size, collected from the Internet. The images were chosen so as to be easily identifiable. For example, goldfish (and whale) images were from a side-angle view. All the atypical photos were whole-body images of the animals. None of the animal pictures was miscategorized persistently in Experiments 3 and 4. Besides employing pictures, rather than words, the same materials and procedures were used in the following two experiments as in the previous two.

Results

The accuracy rate was 96% for atypical animals and 99% for typical animals ($p = .078$). Once again, all incorrect trials were discarded before analysis.

Time-normalized analysis. As can be seen in Figure 5A, the atypical animal trials elicited substantially more curved computer mouse trajectories than did the typical animal trials. Collapsed across leftward and rightward trajectories, t tests revealed atypical − typical x-coordinate differences that were significantly greater than zero ($p < .05$) across 61 consecutive time steps, from the 26th to the 87th slices. This is considerably larger than the 38 from Experiment 1. In a repeated measures ANOVA with binned time steps, there were significant main effects of trial type [$F(1,44) = 34.5$, $MS_e = 2,012.5$, $p < .001$] and bin [$F(2,43) = 1,773.6$, $MS_e = 1,475.3$, $p < .001$] and a significant interaction [$F(2,43) = 14.7$, $MS_e = 1,007.6$, $p < .001$]. Planned comparisons show that, in all three normalized time bins, atypical trials diverged in the x-coordinate from typical trials ($ps < .05$), exhibiting attraction to the competing response category.

Space-normalized analysis. By normalizing the coordinates for travel from the origin $(0, 0)$ to $(1, 1)$, the three real-time bins revealed robust main effects of trial type [$F(1,43) = 45.6$, $MS_e = 0.056$, $p < .001$] and time bin [$F(2,42) = 289.1$, $MS_e = 0.028$, $p < .001$] and an interaction between trial type and time bin [$F(2,42) = 8.1$, $MS_e = 0.034$, $p = .001$]. Planned comparisons showed that all three time bins $(0–500, 500–1,000, \text{ and } 1,000–1,500 \text{ msec})$ contained significantly different mean x-coordinates for the atypical and the typical animal trials ($ps < .001$).

Additional measures. The difference between movement initiation latencies was significant and in the expected direction [440 msec for atypical animals vs. 383 msec for typical animals; $t(44) = 2.5$, $p < .05$]. There was also a significant difference between the two conditions in the total time taken to perform the categorization

[2,150 vs. 1,560 msec; $t(44) = 5.2$, $p < .001$]. Movement duration differed significantly [1,710 vs. 1,177 msec; $t(44) = 7.7$, $p < .001$], and total distance traveled differed significantly as well [1,017 vs. 830 pixels; $t(44) = 5.4$, $p < .001$]. Item analyses for all the measures revealed significant differences between typical and atypical stimuli ($ps < .01$). In contrast to Experiment 1, *every* dependent measure showed a strong reliable difference between the atypical and the typical conditions.

Analysis of initial direction in degrees shows a significant difference between the atypical and the typical trials [$t(44) = 2.1$, $p < .05$], with the atypical trials, in fact, showing a negative angle toward the competitor ($-2.98°$) and the typical trials showing a significant positive angle of $4.5°$ toward the correct category [$t(44) = 2.1$, $p < .05$]. The negative angle on the atypical trials was not significantly different from $0°$ [$t(44) = -0.94$, $p = .35$].

A distribution analysis with the Kolmogorov–Smirnov test again showed that curvatures for atypical and typical trajectories did not have different distributions ($\chi^2 = 0.94$, $p = .97$). Bimodality coefficients were similar and below .555 (.503 and .457 for atypical and typical, respectively). Sample entropy measures showed much stronger effects than those in Experiments 1 and 2. Again, using window sizes of 2–10 and 0.2 times the standard deviation of all x-coordinate fluctuations across trials as the tolerance, the atypical trials consistently showed reliably higher sample entropy for all the window sizes ($ps < .05$).

Discussion

The results indicate that the process of categorizing pictures of animals exhibited temporally dynamic spatial attraction effects in the mouse movement trajectories. In fact, the perceptual ambiguity of the atypical animals seemed to induce greater competition effects, in both trajectory and time measures, than did the lexical stimuli, as was predicted above. In Experiment 1 (with animal *words*), the spatial divergence between the atypical animal trajectories and the typical animal trajectories was significant for 38 consecutive normalized time slices, whereas in the present experiment (with animal *pictures*), the spatial divergence was significant for 61 consecutive normalized time slices. Moreover, the atypical/typical differences in total response time, movement duration, and movement distance were also greater in the present experiment than they were in Experiment 1.
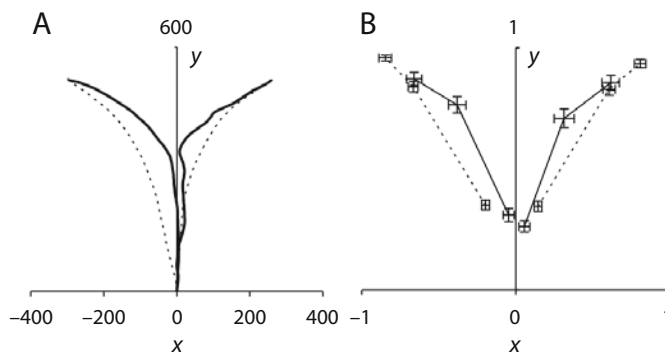
EXPERIMENT 4

Just as Experiment 1 (with animal words) had its control comparison in Experiment 2, using less competitive, alternative response options, Experiment 3 (with animal pictures) has its control comparison in the present experiment. In Experiment 4, the taxonomic class response options were the same as those in Experiment 2—for example, *mammal* and *bird* for the exemplar *whale*—but the exemplar was presented as a picture, instead of a word. Thus, the alternative (incorrect) category in the atypical animal condition here (e.g., *bird*) had less featural match to the exemplar (e.g., *whale*) than the one in Experiment 3 had (e.g., *fish*). However, the poor match between the salient visual properties of the picture and the correct category response (e.g., whales do not *look* much like mammals) may cause the correct category to be somewhat less competitive in the categorization process than it is when the exemplars are presented as words. With closer relative competitiveness of the two categories, the alternative (incorrect) category response may be able to exert a spatial attraction effect more substantial than that seen in Experiment 2.

Method

Participants. Thirty-nine undergraduate participants signed up for this experiment for extra credit in their psychology classes.

Materials and Procedure. In this experiment, the same technique and same pictures as those in Experiment 3 were used, but the experimental (atypical animal) condition involved less competitive categories. The same noncompeting response options as those in Experiment 2 (see Table 1) were employed here.

Figure 5. (A) Mean time-normalized mouse movement trajectories in Experiment 3, separated for left- and rightward responses to the correct target. Atypical picture trials (solid lines) indicate strong divergence toward the competing category. (B) Space-normalized time bins show a strong attraction and slower progress toward the target on atypical trials (solid lines).
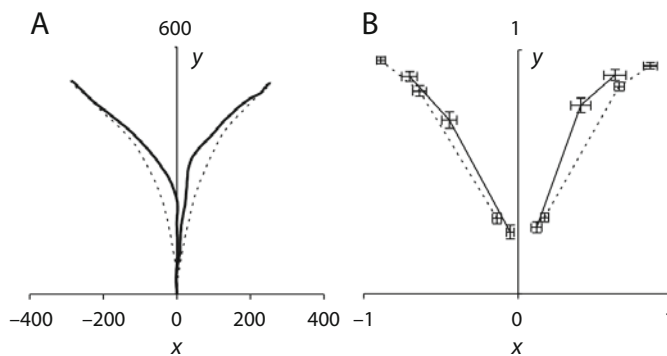
Figure 6. (A) Mean time-normalized mouse movement trajectories in Experiment 4. Atypical picture trials (solid lines) also reveal a significant attraction toward the competing category. **(B)** Space-normalized time bins again show a strong but diminished attraction and slower progress toward the target on atypical trials (solid lines).

Results

The accuracy rate was 94% for atypical animals and 97% for typical animals ($p = .58$). Once again, all incorrect trials were discarded before analysis.

Time-normalized analysis. An analysis of mean trajectories in combined left and right trajectories again showed differences in x-coordinates that were significantly greater than zero across 66 consecutive time steps, from the 28th to the 94th slices ($ps < .05$; see Figure 6A). A repeated measures ANOVA with binned time steps, as in the previous experiments, showed highly significant main effects of trial type [$F(1,38) = 17.0, MS_e = 1,538.6, p < .001$] and bin [$F(2,37) = 2,401.5, MS_e = 1,134.8, p < .001$] and an interaction with time step bin [$F(2,37) = 7.8, MS_e = 589.6, p < .001$]. Planned comparisons show significant differences between trial types in the second and third normalized time bins ($ps < .01$).

Space-normalized analysis. Figure 6B shows space-normalized trajectories in the three real-time bins. A repeated measures ANOVA showed significant effects of trial type [$F(1,36) = 28.7, MS_e = 0.047, p < .001$] and time bin [$F(2,35) = 739.0, MS_e = 0.016, p < .001$] and a significant interaction between trial type and bin [$F(2,35) = 9.1, MS_e = 0.017, p = .001$]. All three bins show a significant difference between trial types ($ps < .05$).

Additional measures. There was no significant difference in initiation latencies between the two conditions [392 msec atypical vs. 349 msec typical; $t(38) = 1.6, p = .12$]. As in Experiment 3, there was a significant difference between the two conditions in the total time taken to perform the categorization [1,813 vs. 1,461 msec; $t(38) = 5.2, p < .001$]. Both movement duration [1,421 vs. 1,112 msec; $t(38) = 4.8, p < .001$] and total distance traveled were significantly greater on the atypical trials [806 vs. 747 pixels; $t(38) = 3.3, p < .01$]. Initial angle from (0, 0) for the typical trials was at 4.9°, significantly greater than 0° [$t(38) = 2.1, p < .05$], whereas the atypical trials were again slightly negative ($-1.7°$) but not significantly so. Item-based analyses showed the same patterns of significance ($ps < .05$), although movement initiation

latency showed a marginal significance [$t(17) = 8.7, p = .07$]. Finally, a Kolmogorov–Smirnov test showed no significant difference between typical and atypical curvature distributions ($\chi^2 = 1.1, p = 1.0$). Interestingly, both bimodality coefficients were near but greater than .555 (.584 and .597 for atypical and typical, respectively). The value for the typical trials, however, was greater than that for the atypical trials. Thus, if each of these distributions was genuinely the result of two independent classes of motor movement trajectories (those that started out correct and those that were sharply corrected in midflight), this bimodality cannot be attributed solely to the atypicality of the whale, seal, penguin, and so forth; even the typical animal pictures elicited this pattern. Finally, just as in Experiment 2, sample entropy showed no significant differences between the atypical and the typical trials. Since Experiments 1 and 3 showed reliably greater sample entropy for atypical trials than for typical trials, this measure may be a critical indicator of the difference between competitive and noncompetitive category trials. When both category response options were substantially competitive (Experiments 1 and 3), the mouse movement trajectories exhibited a conspicuous complexity that may have been due to the nonlinear dynamics inherent in an attractor landscape that had multiple strong attractor basins.

Discussion

With visual images of exemplars, a noncompetitive alternative category still induced considerable curvature in the mouse movement trajectory. When divergence between the average trajectory for atypical animals and the average trajectory for typical animals is plotted over time for Experiments 3 and 4 together (see Figure 7), the competitive alternative categories in Experiment 3 show a numerically stronger attraction effect than do the noncompetitive alternative categories in Experiment 4. However, the two atypical animal trajectories in Experiments 3 and 4 do not exhibit a significant difference for more than eight consecutive normalized time steps. Nevertheless, when independent samples t tests of the outcome-based dependent measures for the experimental trials in Experi-
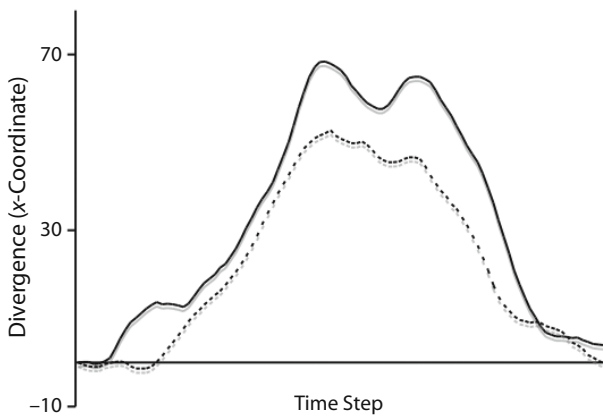
Figure 7. Mean total divergence (atypical trial − typical trial *x*-coordinates) for Experiments 3 and 4. Experiment 3 (solid line) exhibits greater, although not significant, divergence, for eight time steps (see the text for details).

ments 3 and 4 were done, total response time [$t(82) = 2.6$, $p < .05$], total time in motion [$t(82) = 2.5$, $p < .05$], and distance traveled [$t(82) = 4.8$, $p < .001$] all indicated that mouse movement was slower and more extensive when the incorrect alternative category was competitive (e.g., *fish* for the *whale* picture) than when the incorrect alternative category was noncompetitive (e.g., *bird* for the *whale* picture). Item-based comparisons of the atypical trials between Experiments 3 and 4 also showed a significant difference for distance in the expected direction [$t(5) = 6.2$, $p < .01$].

Thus, although not quite as much as in Experiment 3, images of atypical animals in Experiment 4 *did* induce computer mouse trajectories that exhibited more spatial attraction toward an incorrect category label than did the trajectories for typical animals. When words, instead of images, were used with this noncompetitive arrangement in Experiment 2, a much diminished trajectory curvature was observed, as compared with the corresponding competing-label scenario in Experiment 1. One simple explanation for this strengthened attraction of the noncompeting labels in Experiment 4 (as compared with Experiment 3) may be that the divergence between trajectories was the result of a reduced featural match of the correct category with the visual properties of the atypical exemplar, rendering the alternative category's meager competitiveness relatively more substantial. When one then compares Experiment 4 with Experiment 3, the exceptionally long mouse movements (in time and space) on atypical animal trials in Experiment 3 can be seen as having been due to that same reduced competitiveness of the correct category, combined with a considerable competitiveness of the alternative category, resulting in a particularly lasting and laborious competition between the categories.

GENERAL DISCUSSION

The lexical stimuli in Experiment 1 revealed that categories in competition lead to a dynamic spatial attrac-

tion of the resultant mouse trajectories. This competition diminished but did not go away when the featurally similar alternative category was replaced by a less competitive alternative category in Experiment 2. Interestingly, this pattern changed when image-based stimuli were used in Experiments 3 and 4. Images of atypical exemplars produced robust mouse trajectory attraction toward both highly competitive alternative categories and less competitive categories alike. These results offer further insight into the nature of categorization. As was mentioned in the introduction, the time course of categorization has only recently been rigorously explored (e.g., Lamberts, 2000; Nosofsky & Palmeri, 1997). Like Abrams and Balota's (1991) results with lexical decision and recall memory, these results further suggest that manual responses themselves may be reflective of a continuous, dynamic process of categorization underlying participant performance in these experiments.

Importantly, these findings may extend the information accumulation theory of Lamberts (2000) and the exemplar random-walk model of Nosofsky and Palmeri (1997). These theories may, in fact, predict the results above. For example, Nosofsky and Palmeri's random-walk model involves a race among exemplars governed by their similarity to a test item. In our case, right versus left response boxes (competing categories) attract the continuous manual movement in accord with the extent to which they fit the exemplar presented. Depending on how its representations are mapped onto motor output, this race could result in a substantial pull toward highly competitive alternative categories (as in Experiments 1 and 3) and somewhat less of a pull with less competitive alternative categories. This basic finding that continuous, graded processing of the exemplars during categorization flows into the effectors can be seen as an extension of an iterative sequential-sampling, information accumulation perspective. Although the perspective of both Lamberts (2000) and Nosofsky and Palmeri is applied largely to the process leading up to a discrete categorization response, our results suggest that the response is itself a component of this continuous, probabilistic processing. A full synthesis may accompany future experiments that overcome some limitations of the present ones. For example, it must be acknowledged that in the present research, a small stimulus set was employed. Future experiments may further bridge the various levels of the categorization process (from perception to response preparation and production) through more extensive stimulus sets. Although the present experiments serve as an initial demonstration of these patterns, using a small set of commonly known animals, there remain numerous issues to explore. These might include the resolution of item-by-item typicality, frequency, and other variables and their concomitant effects on continuous output. Despite the present limitations, these experiments demonstrate a potentially fruitful avenue for eventually mapping out an uninterrupted explanatory landscape from perceptual input to motor output.

Classical perspectives on cognitive processing as organized into discrete serial stages would likely predict

that only noise would account for graded output patterns in such tasks as taxonomic categorizing—and that mean trajectories should, in fact, simply indicate the target response. Although the present results constitute a particularly strong recommendation against this purely serial perspective on the *process* of categorization, we would argue that they also have broader theoretical significance regarding representational issues in cognitive theories. Throughout the cognitive sciences, debate continues over the nature of conceptual representation. A central dichotomy in this debate has to do with the format of representation, and related processing-flow issues, that underlie cognition. On one side of this debate, representations are largely characterized as discrete and symbolic and undergo computational or algorithmic manipulation akin to a Turing machine (e.g., Dietrich & Markman, 2003; Fodor, 2000; Marcus, 2001; Pinker, 1997). On another side, there are a variety of proposals that see representational states as probabilistic, graded informational states that either undergo computational manipulation (Massaro, 1989, 1998) or are part and parcel of a system that blends content and process in its operation (Elman et al., 1996; Port & van Gelder, 1995; Rumelhart & McClelland, 1986; Spivey & Dale, 2004). This dichotomy only approximates the rich spectrum of opinion in cognitive science (cf. hybrid systems such as those by Sun, 1997; Young & Lewis, 1999). Nevertheless, debate tends to focus on the extent to which format of representation contributes most to our understanding of cognition.

A range of behavioral data are typically adduced to support one and challenge the other. A classic example of phenomena interpreted as supporting the existence of discrete-symbolic representations is the very act of categorization itself: "It follows . . . that if a system categorizes, then it has discrete representations" (Dietrich & Markman, 2003, p. 102). For example, early research on categorical perception of speech sounds not only fit perfectly into the traditional perspective on symbolic computation, but also led to proposals that there exist specialized processes of this sort for speech (see Massaro, 1998, for a review). Categorization is but one property of several behaviors that animals exhibit that have been proposed to support the centrality (or at least existence) of discrete-symbolic representations in cognition (for other properties, see Dietrich & Markman, 2003; Marcus, 2001). It has, however, been of central importance for the understanding of our cognitive architecture.

In relation to recent research (e.g., McMurray et al., 2003; Nederhouser & Spivey, 2004), lexical and perceptual categorization in the present experiments revealed graded response patterns in the participants' mouse trajectories (see also Spivey et al., 2005). As in Abrams and Balota (1991), it seems that the time course of categorization is reflected in the manual output from the process. In particular, given our results that competing category labels generate greater trajectory curvature, slower time measures, and longer trajectory excursions, this temporally sensitive deviation in motor output suggests that partially active states underlie the time course of categorization. The upshot for cognitive representation is subtle but important. If discrete representational states underlie

cognition at some point following perceptual processing, there must still exist some *granularity* of information, in Miller's (1982) sense, by the time this process is being converted into motor output. If this is granted, there is a limit on the discreteness imputed to the representational states underlying categorization. Discrete representational perspectives may, indeed, account for these kinds of results by allowing differing levels of granularity in representational discreteness. Indeed, by further pursuing the line of research in Miller (1982; see also Miller & Ulrich, 2003), these theoretical perspectives may be reconciled if a common ground between purely continuous and purely discrete representations can be found (cf. Dale & Spivey, 2005). This would involve specifying how refined the granularity is at various stages of processing.

The present experiments have demonstrated that echoes of continuous (nondiscrete) informational states can be observed in the dynamic properties of resultant motoric responses—granularity evident even in the manual output. Any postulated discrete representational states mediating sensors and effectors must carry at least some relevant information from early graded states in order for the motor output to mimic the continuity of the sensory input. In other words, although reaction time and accuracy measures may reveal information about the decision process during discrete, algorithmic processing, the graded manual output from the system observed here indicates that even when these discrete decision processes collapse onto the effectors, there remains some granularity.

AUTHOR NOTE

REFERENCES

ABRAMS, R., & BALOTA, D. (1991). Mental chronometry: Beyond reaction time. *Psychological Science*, **2**, 153-157.

ALLOPENNA, P. D., MAGNUSON, J. S., & TANENHAUS, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language*, **38**, 419-439.

ASHBY, F. G., BOYNTON, G., & LEE, W. W. (1994). Categorization response time with multidimensional stimuli. *Perception & Psychophysics*, **55**, 11-27.

BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. (1956). *A study of thinking*. New York: Wiley.

BURNETT, R. C., MEDIN, D. L., ROSS, N. O., & BLOK, S. V. (2005). Ideal is typical. *Canadian Journal of Experimental Psychology*, **59**, 3-10.

COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.

COLES, M. G. H., GRATTON, G., BASHORE, T. R., ERIKSEN, C. W., & DONCHIN, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 529-553.

COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **8**, 240-247.

CREE, G. S., MCRAE, K., & MCNORGAN, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, **23**, 371-414.

DALE, R., & SPIVEY, M. J. (2005). From apples and oranges to symbolic dynamics: A framework for conciliating notions of cognitive representations. *Journal of Experimental & Theoretical Artificial Intelligence*, **17**, 317-342.

DIETRICH, E., & MARKMAN, A. B. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind & Language*, **18**, 95-119.

DOYLE, M., & WALKER, R., (2001). Curved saccade trajectories: Voluntary and reflexive saccades curve away from irrelevant distractors. *Experimental Brain Research*, **139**, 333-344.

EFRON, B., & TIBSHIRANI, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

ELMAN, J. L., BATES, E. A., JOHNSON, M. H., KARMILOFF-SMITH, A., PARISI, D., & PLUNKETT, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

ESTES, Z. (2003). Domain differences in the structure of artifactual and natural categories. *Memory & Cognition*, **31**, 199-214.

FELDMAN, D. P., & CRUTCHFIELD, J. P. (1998). Measures of statistical complexity: Why? *Physics Letters A*, **238**, 244-252.

FODOR, J. A. (2000). *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press, Bradford Books.

GARROD, S., & SANFORD, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning & Verbal Behavior*, **16**, 77-90.

GOLD, J. I., & SHADLEN, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, **404**, 390-394.

GOLD, J. I., & SHADLEN, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, **5**, 10-16.

GOLDBERGER, A. L., AMARAL, L. A. N., GLASS, L., HAUSDORFF, J. M., IVANOV, P. C., MARK, R. G., ET AL. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, **101**, e215-e220.

GOODALE, M. A., PÉLISSON, D., & PRABLANC, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature*, **320**, 748-750.

HEIT, E. (1994). Models of the effects of prior knowledge on category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1264-1282.

HINTON, G. E., & SHALLICE, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, **98**, 74-95.

JOB, R., RUMIATI, R., & LOTTO, L. (1992). The picture superiority effect in categorization: Visual or semantic? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1019-1028.

LAKE, D. E., RICHMAN, J. S., GRIFFIN, M. P., & MOORMAN, J. R. (2002). Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology*, **283**, R789-R797.

LAMBERTS, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, **124**, 161-180.

LAMBERTS, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, **107**, 227-260.

LAMBERTS, K. (2002). Feature sampling in categorization and recognition of objects. *Quarterly Journal of Experimental Psychology*, **55A**, 141-154.

LYNCH, E. B., COLEY, J. D., & MEDIN, D. L. (2000). Tall is typical: Central tendency, ideal dimensions, and graded category structure among tree experts and novices. *Memory & Cognition*, **28**, 41-50.

MAGNUSON, J. S. (2005). Moving hand reveals dynamics of thought: Commentary on Spivey, Grosjean, & Knoblich (2005). *Proceedings of the National Academy of Sciences*, **102**, 9995-9996.

MARCUS, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press, Bradford Books.

MASSARO, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, **21**, 398-421.

MASSARO, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA: MIT Press, Bradford Books.

MCCLOSKEY, M. E., & GLUCKSBERG, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, **6**, 462-472.

MCMURRAY, B., & SPIVEY, M. (2000). The categorical perception of consonants: The interaction of learning and processing. *Proceedings of the Chicago Linguistics Society*, **34**, 205-220.

MCMURRAY, B., TANENHAUS, M. K., ASLIN, R. N., & SPIVEY, M. J. (2003). Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research*, **32**, 77-97.

MCRAE, K., DE SA, V. R., & SEIDENBERG, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, **126**, 99-130.

MEDIN, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, **44**, 1469-1481.

MEDIN, D. L., & ROSS, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 189-223). Hillsdale, NJ: Erlbaum.

MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.

MILLER, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 273-296.

MILLER, J., & ULRICH, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology*, **46**, 101-151.

MURPHY, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press, Bradford Books.

MURPHY, G. L., & MEDIN, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.

MURPHY, G. L., & ROSS, B. H. (2005). The two faces of typicality in category-based induction. *Cognition*, **95**, 175-200.

NEDERHOUSER, M., & SPIVEY, M. J. (2004). Eye-tracking and simulating the temporal dynamics of categorization. In K. Forbus, D. Gentner, & T. Reiger (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (p. 1612). Mahwah, NJ: Erlbaum.

NOSOFSKY, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 54-65.

NOSOFSKY, R. M. (1992). Exemplars, prototypes and similarity rules. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of W. K. Estes* (Vol. 1, pp. 149-168). Hillsdale, NJ: Erlbaum.

NOSOFSKY, R. M., & PALMERI, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, **104**, 266-300.

OP DE BEECK, H., & WAGEMANS, J. (2002). Visual object categorisation at distinct levels of abstraction: A new stimulus set. *Perception*, **30**, 1337-1361.

PAIVIO, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, **45**, 255-287.

PINKER, S. (1997). *How the mind works*. New York: Norton.

PORT, R. F., & VAN GELDER, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

REED, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, **3**, 382-407.

REHDER, B., & HOFFMAN, A. B. (2005). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 811-829.

RICHARDSON, D. C., DALE, R., & SPIVEY, M. J. (2007). Eye movements in language and cognition: A brief introduction. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 323-344). Amsterdam: John Benjamins.

RICHMAN, J. S., & MOORMAN, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology*, **278**, H2039-H2049.

RIPS, L. J., SHOBEN, E. J., & SMITH, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning & Verbal Behavior*, **12**, 1-20.

ROSCH, E. H. (1973). Natural categories. *Cognitive Psychology*, **4**, 328-350.

ROSCH, E. [H.] (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, **104**, 192-233.

ROSCH, E. [H.], & MERVIS, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573-605.

RUECKL, J. G. (2002). The dynamics of visual word recognition. *Ecological Psychology*, **14**, 5-19.

RUMELHART, D. E., & MCCLELLAND, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, Bradford Books.

SHIN, J. C., & ROSENBAUM, D. A. (2002). Reaching while calculating: Scheduling of cognitive and perceptual–motor processes. *Journal of Experimental Psychology: General*, **131**, 206-219.

SMITH, J. D. (2002). Exemplar theory's predicted typicality gradient can be tested and disconfirmed. *Psychological Science*, **13**, 437-442.

SNODGRASS, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning & Verbal Behavior*, **23**, 3-22.

SNODGRASS, J. G., & MCCULLOUGH, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 147-154.

SPIVEY, M. J., & DALE, R. (2004). The continuity of mind: Toward a dynamical account of cognition. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 45, pp. 85-142). Amsterdam: Elsevier.

SPIVEY, M. J., GROSJEAN, M., & KNOBLICH, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, **102**, 10393-10398.

STORMS, G. (2005). Exemplar models in the study of natural language concepts. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 45, pp. 1-39). Amsterdam: Elsevier.

SUN, R. (1997). Learning, action and consciousness: A hybrid approach toward modeling consciousness. *Neural Networks*, **10**, 1317-1331.

VERBEEMEN, T., STORMS, G., & VERGUTS, T. (2003). Determinants of speeded categorization in natural concepts. *Psychologica Belgica*, **43**, 139-151.

VISWANATHAN, M., & CHILDERS, T. L. (2003). An enquiry into the processing of categorization of pictures and words. *Perceptual & Motor Skills*, **96**, 267-287.

YOUNG, R. M., & LEWIS, R. L. (1999). The Soar cognitive architecture and human working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 224-256). Cambridge: Cambridge University Press.

NOTES

1. Although this bimodality test will rule out the possibility that, within a condition, half of the trajectories are roughly straight and half of the trials are substantially curved, it cannot rule out the possibility that most trajectories start out toward the incorrect alternative and then a strategy of decisively switching responses takes place at varying points in time. Such a strategy would, indeed, produce a unimodal distribution of trajectory curvatures; however, it is unclear why the start of such a process would always involve moving toward the incorrect alternative in the first place.

2. Degrees of freedom are different in space-normalized analyses, because some participants had too many empty cells in one or more time bins.

3. We used a MATLAB script from the PhysioNet package to compute sample entropy (Goldberger et al., 2000; Lake, Richman, Griffin, & Moorman, 2002).

APPENDIX

In order to establish a basis for our criterion used in the multiple *t* tests on time-normalized trajectories, we conducted simulations akin to the bootstrap method (Efron & Tibshirani, 1993). The mean and standard deviation of each of 101 time steps in atypical and typical trajectories in Experiment 1 were recorded. Next, 10,000 simulated "experiments" were computed, using the same mean and standard deviation. In other words, in each of these simulations, we constructed 41 model participants (*N* in Experiment 1) by constructing atypical and typical trajectories from these time step means and standard deviations. Within each trial type, each time step (out of 101) was sampled from a normal distribution with the mean and standard deviation from the time steps of the *actual* mean trajectories, thus preserving the nonindependence between time steps. We then ran *t* tests within each of the 10,000 simulations for each time step (101 time steps; *N* = 41). Of these 10,000, we recorded the frequency with which *sequences* of significant differences occurred. The simulations revealed significantly different sequences of 6, 7, and 8 time steps with percentages of 3%, 1%, and 0.5%, respectively (see Table A1). A sequence of eight consecutive significant *t* tests was thus produced by chance in these simulations with less than .01 probability and was selected as a conservative measure. This simulation therefore recommends a criterion of 8 consecutive significantly different time steps for discerning whether the interdependent time steps in our experiments have sequences that exceed what would occur by chance alone. This criterion was also used in Experiments 2–4.

Table A1
Frequencies (in Percentages) of Sequences Discovered in the
10,000 Simulated Experiments

Sequence Size	%	*p*
3	56	
4	22	
5	9	
6	3	<.05
7	1	
8	0.5	<.01
9	0.2	
10	0.001	<.001