Mouse Tracking Shows Attraction to Alternative Targets While Grounding Spatial Relations

Jonas Lins (jonas.lins@ini.rub.de)
Gregor Schöner (gregor.schoener@ini.rub.de)

Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract
Evidence that higher cognitive processes are coupled in a graded and time-continuous way to sensory-motor processes has been seen through interaction effects in reaction time tasks that probed potentially overlapping perceptual and motor representations (e.g., Bastian, Schönner, & Riehle, 2003; Cisek & Kalaska, 2005). That higher-level cognitive processes, such as understanding spatial concepts, interact with perceptual and motor processes has been seen through interaction effects in reaction time tasks that probed potentially overlapping perceptual and motor representations (e.g., Richardson, Spivey, Barsalou, & McRae, 2003). Mouse tracking and similar techniques have been a major tool to study and establish this link between cognitive and sensory-motor processes (for review see Freeman, Dale, & Farmer, 2011; Song & Nakayama, 2009).

In mouse tracking paradigms, participants are asked to solve a cognitive task that may engage high-level concepts. The response is usually ensured to begin early in relation to the decision process, analogously to the timed movement initiation paradigm (Ghez et al., 1997). The theoretical account (Erhagen & Schönner, 2002) linked this phenomenon to graded distributions of population activation in the motor and premotor cortex (Bastian, Schönner, & Riehle, 2003; Cisek & Kalaska, 2005). That higher-level cognitive processes, such as understanding spatial concepts, interact with perceptual and motor processes has been seen through interaction effects in reaction time tasks that probed potentially overlapping perceptual and motor representations (e.g., Richardson, Spivey, Barsalou, & McRae, 2003). Mouse tracking and similar techniques have been a major tool to study and establish this link between cognitive and sensory-motor processes (for review see Freeman, Dale, & Farmer, 2011; Song & Nakayama, 2009).

Introduction
Over the last two decades, a major theme in cognitive science has been that cognitive processing is graded in nature, unfolds continuously in time, and is coupled to perceptual and motor processes (Schöner, Spencer, & the DFT Research Group, 2015; Spivey, 2007). Movement preparation, for instance, was shown to occur in a graded and continuous form when the time interval between imperative stimulus and response was varied experimentally in the timed movement initiation paradigm (Ghez et al., 1997). The theoretical account (Erhagen & Schönner, 2002) linked this phenomenon to graded distributions of population activation in the motor and premotor cortex (Bastian, Schönner, & Riehle, 2003; Cisek & Kalaska, 2005). That higher-level cognitive processes, such as understanding spatial concepts, interact with perceptual and motor processes has been seen through interaction effects in reaction time tasks that probed potentially overlapping perceptual and motor representations (e.g., Richardson, Spivey, Barsalou, & McRae, 2003). Mouse tracking and similar techniques have been a major tool to study and establish this link between cognitive and sensory-motor processes (for review see Freeman, Dale, & Farmer, 2011; Song & Nakayama, 2009).

In mouse tracking paradigms, participants are asked to solve a cognitive task that may engage high-level concepts. The response is usually ensured to begin early in relation to the decision process, analogously to the timed movement initiation paradigm. Typically, a computer mouse must be moved toward one of two (sometimes a few) response locations. What varies over time is the certainty over the response, which is reflected by the movement deviating somewhat into intermediate directions. A limitation of previous mouse tracking research with respect to the interaction of cognitive and sensorimotor processes is that the motor responses are usually fixed and assigned symbolically to the solutions of the cognitive task (e.g., left button for “yes” and right button for “no”). When a small number of possible movement targets is known in advance, it is not the decision itself that specifies the metrics of the required movement.

We developed an experimental paradigm in which the spatial target of a mouse movement was directly specified by the cognitive task and unknown in advance. Participants read a target description invoking colors and spatial relations (e.g., “the yellow to the left of the green”) and then saw a complex layout of colored objects (e.g., Fig. 1c). A selection decision had to be made among multiple same-colored target items (“yellow”), clicking the one best matching the spatial term. Although some aspects of the visual arrays are tightly coupled and impossible to vary independently, we were able to unravel their effects by counter-balancing those aspects that posed potential confounds, separately for each comparison.

We looked for three signatures of interaction among the task’s cognitive and sensory-motor dimensions. First, the alternate targets (i.e., distractors) were predicted to metrically attract the trajectories, in analogy to the effect of alternate but incorrect choice alternatives in classic mouse tracking research. Second, while the reference item is never an alternate target, it likely engages attentional focus during processing (Franconeri, Scimeca, Roth, Helseth, & Kahn, 2012), which we predicted to also cause attraction. Third, based on previous evidence (Tower-Richardi, Brunye, Gagnon, Mahoney, & Taylor, 2012), we expected a bias into the direction described by the spatial term. In our recent neural process model of spatial language grounding (Richter, Lins, & Schöner, 2017), discrete amodal representations of target, reference, and spatial term guide activation in continuous perceptual representations. In the model, target and reference must become active sequentially, because overlapping substrates are engaged to spatially index the corresponding visual items. We thus expected some temporal displacement in the biases toward these items. The spatial term, in contrast, impacts another substrate and is active early and in parallel to target and reference. We thus expected the spatial term effect to bias movement metrics globally and early on.

Methods
Participants Twelve participants (5 female, 7 male, mean age 27.4 years ± 3.8 s.d., one left-handed) were recruited by notices at the local department, receiving EUR 10 for participation. All were naïve to the experimental hypotheses, native German speakers, and had normal vision.
Procedure  A trial began with a black start marker in the bottom center of an otherwise gray screen. To proceed, the participant moved the mouse cursor (a white dot) onto the start marker. After resting there for 300 ms, a German spatial phrase appeared at a position somewhat random around the center of the stimulus region (up to ±48 mm × 20 mm horizontal/vertical), for instance, “Das Gelbe links vom Grünen.” (“The yellow left of the green.”), denoting a target item by its color and its position in relation to a reference item, which was also specified by color. Display duration ranged randomly from one to two seconds. The phrase then disappeared and a beep signaled the participant to start moving the cursor upwards within one second (if movement started too late, the trial was aborted and appended at the end). Movement onset was registered when a velocity of 20 mm/s was exceeded. At that point, twelve colored items appeared above the start marker (e.g., Fig. 1c). Thus, movement was already in progress when the selection process started. One of the twelve items was the uniquely colored reference mentioned in the phrase, one was the target, and one was the main distracter, which had the same color as the target but provided a worse match for the spatial term, according to a spatial template described below. The participant had to select the item which in his or her opinion best matched the preceding phrase (participants could select any item). If time since movement onset exceeded two seconds, the trial was aborted and appended at the end. Participants were instructed that there were no incorrect responses, that items were not obstacles, and that response time was limited such that they had to respond promptly. After 13 practice trials, each participant completed 446 trials in random order (one completed eight more, to use the entire set of 5360 trials, described below).

Spatial phrases  Spatial phrases were in German and of the form “DasartGrüntgcrechts vomgrRotenref” (“Theartgreenfigto the right ofgrthereft”), where art denotes the article, which was always “Das”, tgt denotes a target component from the set {Rote, Grüne, Blaue, Gelbe, Weiße, Schwarzart}, ref denotes a reference component from the set {Roten, Grünem, Blauen, Gelben, Weißem, Schwarzen} (“the {red, green, blue, yellow, white, black} one”), and spt denotes a spatial term from the set {links vom, rechts vom, über dem, unter dem} (“left of, right of, above, below”). In all trials, the spatial phrase posed a valid description of an item in the stimulus display.

Stimulus displays  Visual items were irregular polygons with an outer diameter of 16.4 mm (1.34° v.a.), each colored in one of the six colors that also occurred in the spatial phrases (red, green, blue, yellow, white, and black). Items were combined into stimulus displays as described in the following.

We generated multiple three-item configurations of a reference item, a target item, and a distracter, differing in how target and distracter were situated relative to the reference. Their positions were selected based on a spatial template, a fit function \( f(\phi, r) \) with the reference at \([0,0]\) that indicates how well a given position, defined by angle \( \phi \) and radius \( r \), matches a spatial term.\(^1\) Fig. 1a (left panel) shows a plot over Cartesian coordinates for “right of”. The shape of the spatial templates is inspired by behavioral data (e.g., Logan & Sadler, 1996) which computational modeling work reproduced using similar functions (Lipinski, Schneegans, Sandamirskaya, Spencer, & Schöner, 2012).

Targets were placed in a region where \( f(\phi, r) > 0.6 \) and where the outer radices of reference and target were separated by at least 0.5 mm (0.04° v.a.; Fig. 1a, center panel). Within this region, targets were centered on the junctions of a square grid, resulting in sixteen evenly distributed target positions. For each of the 16 target positions a separate set of distracter positions was determined (out of which one distracter was used per trial, paired with the respective target). These were obtained with the same method as the targets, but the general region for distracters was constrained by \( f(\phi, r) > 0.4 \) (green outline in Fig. 1a, center panel), and as an additional constraint distracters’ fit had to be at least 0.03 lower than target fit (min. border-to-border distance again 0.5 mm; see Fig. 1a, right panel). Hence, the shape of distracter regions differed between target positions so that distracter numbers differed as well, varying from 16 to 25 (mean 20.9) per target position. Colors for each three-item set were randomly picked, with target and distracter being colored alike.

A set of 335 different three-item configurations was obtained for each spatial term, differing between terms only in orientation. We thus arrived at 1340 configurations, each of which was presented at four different positions on the screen, such that the target item of each configuration appeared once in each of four different on-screen target locations (black X’s in Fig. 1b). These were arranged in a square around the cen-

\(^1\) In polar coordinates the function is given by

\[
f(\phi, r) = e^{-\frac{(\phi-\phi_0)^2}{2\sigma_\phi^2}} \cdot e^{-\frac{(r-r_0)^2}{2\sigma_r^2}} \cdot (1 + e^{-\beta(|\phi-\phi_0| - \theta_{\text{flex}})})^{-1},
\]

where \( \phi \) denotes polar angle, \( r \) is the radius, \( \phi_0 \) is the mean of a Gaussian function over angle, \( \sigma_\phi \) is its standard deviation, \( r_0 \) and \( \sigma_r \) are analogue parameters for a Gaussian over radius, \( \beta \) is the steepness of a sigmoid function over angle, and \( \theta_{\text{flex}} \) is the separation of its inflection point from the mean of the Gaussian over angle. We used \( \phi_0 = 1.05, r_0 = 0 \text{ mm}, \sigma_\phi = 47 \text{ mm}, \beta = 25, \) and \( \theta_{\text{flex}} = 1.45 \). Parameter \( \phi_0 \) depended on the spatial term, with ”right of”, ”above”, ”left of”, and ”below” corresponding to \( \phi_0 = \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\} \text{ rad.} \)
We analyzed only trials where participants selected the item best matching the spatial phrase according to the fit function (hereafter called target). We refer to the straight line from a trajectory’s first point to the target item’s center as direct path.

Sharply curved trajectories were discarded from analyses, in order to consider only trajectories exhibiting graded attraction while excluding re-decisions in mid-flight and mouse-overshoots. Curvature was assessed by temporarily interpolating to a uniform segment length of 5 mm and then applying the osculating circle method (considering each vertex and its two neighbors). Trajectories exceeding a curvature of 0.1 were discarded. We chose this approach over other values such as area under curve (Hehman, Stolier, & Freeman, 2015), as these are less informative in a setup with multiple potential effect sources on both sides of the direct path.

**Trajectory preparation** Trajectories were trimmed to start with movement onset and to end with the first data point after crossing the target border. They were then translated to start at [0,0] and rotated around that point by the angle between the target’s position vector and the y-axis. Positive x-values thus denote deviation from the direct path to the right, negative values indicate leftward deviation. Trajectories’ spatial coordinates were linearly interpolated over 151 equidistant time steps to enable averaging (combining position data from identical values of elapsed proportion of total movement time).

**Statistical analysis** Mean trajectories were compared by testing for differences between x-coordinates at each of the 151 time steps using two-tailed paired-sample t-tests with \( p < 0.01 \). Since data points in each mean trajectory are highly interdependent, and given the large number of tests, the informative value of each individual t-test is limited. To remedy this, we used the bootstrapping procedure introduced by Dale, Kehoe, and Spivey (2007), providing a criterion for how many t-tests in sequence must yield significance before a difference between trajectories can be considered overall significant. A separate criterion with \( p < 0.01 \) was computed for each comparison based on 10,000 artificial experiments each.

**Balancing to isolate main effects** To obtain unbiased estimates of the individual effects of distracter, reference, and CoM position by comparing two conditions (e.g., all trials...
where the distracter was left versus right of the direct path),
the impact of the others (e.g., reference and CoM side) must
be taken into account. For this, we distinguished trials into
categories that indicated whether a potentially confounding
item was on the same or opposite side of the direct path as
the item of interest. There was a different set of four categ-
ories for each item of interest. For instance, balancing cat-
gories for the distracter effect were rs/cs, rs/co, ro/cs, and
ro/co, “r” denoting the reference, “c” the CoM, and “s” and
“o” denoting whether the respective item was on the same
or opposite side of the direct path as the distracter. Corre-
sponding categories for considering reference and CoM were
denoted analogously (“d” denoting the distracter). When com-
paring two sets of trials for one effect, balancing out the other
two then works by ensuring that each set is composed of an
equal number of trials from each relevant balancing category.
This requirement is not fully satisfied by our full set of tri-
als and some comparisons. To allow judging how imbalances
may have affected the respective data, Fig. 2 plots the dis-
tribution of trial numbers over the relevant balancing categories
alongside each comparison. To further validate that observed
effects were not due to imbalances, post hoc balancing was
conducted: We did a second analysis for each comparison,
equivalent to the one performed on the full trial set, but before-
hand randomly discarded trials from over-represented cate-
gories such that a balanced distribution was obtained within
each condition and participant. We report when this substi-
tually changed effects. For the overall mean trajectory, cat-
ergories were based on item sides relative to the direct path.
Category labels used the same letters as above, in addition to
“l” (left) and “r” (right; e.g., dr/lr/cl means that the distracter
was on the right and the other items on the left side).

Results

A total of 5245 trajectories was obtained (115 were lost due to
technical problems) and participants selected the best-fitting
item in 4710 (mean 89.82% ± 3.3 s.d.). Of these, 446
(9.47%) exceeded curvature threshold, leaving 4264 (81.3%)
for analysis. Mean movement time was 1061 ms (± 116
s.d.); noteworthy differences occurred only between upper
and lower target positions (1140 ms ± 119 s.d. and 977 ms
± 114 s.d.). Participants reported not to have noticed that
target positions were limited to four locations.

Fig. 2 shows mean trajectories for all comparisons, along
with trial distributions over balancing categories for each con-
dition. The overall mean trajectory (Fig. 2a) slightly curved
rightwards, likely reflecting kinematic bias. To provide an
idea of this bias in relation to other effects, dotted gray lines
in each panel of Fig. 2 show the mean over trials from the
compared conditions. As expected based on the task in-
structions, a strong bias toward the CoM was evident (Fig. 2b).

We report statistical test results in this form: 46/8, 5–50%,
providing the number of successive significant time steps (46)
along with the bootstrap criterion for overall significance (8),
followed by the percentages of elapsed movement time at the
start and end of the sequence (5–50%). Considering all tri-
als, there was a significant bias away from the reference side
in the first half of the movement (57/6, 1.3–38.4%; Fig. 2c)
and a significant bias in reference direction in the second half
(52/6, 66.2–100%). Assessing the effect of reference side
separately for trials with horizontal-axis spatial terms (“left”
and “right”) and for vertical ones (“above” and “below”) showed
that the bias away from the reference was driven by
the horizontal term trials (69/30, 1.3–46.4%; Fig. 2d). Note
that in these trials deviation away from the reference is con-
gruent with spatial term direction. Correcting for the over-
represented distracter-opposite trials (cs/do, co/do) by post
hoc balancing did not remove the effect (70/47, 1.3–47%;
Fig. 2f). The later bias toward the reference was driven by
the vertical spatial term trials (94/31, 38.4–100%; Fig. 2e).
Post hoc balancing showed that it was not due to the over-
represented distracter-same trials and resulted in an earlier
onset of the reference effect (102/31, 33.1–100%; Fig. 2g).

As shown in Fig. 2h, there was a sustained, significant
bias in distracter direction for the whole trial set (100/15,
33.4–100%). Assessing the effect separately by spatial term
axis showed that the effect’s early component was driven ex-
clusively by horizontal term trials (140/6, 8–100%; vertical:
86/33, 43.7–100%; Fig. 2i,j). These included a pronounced
majority of reference-opposite trials (ro/cs, ro/co), suggest-
ing that the distracter effect’s early component may in fact
be a bias in spatial term direction (i.e., away, from the refer-
ence), as reported above. Post hoc balancing indeed reduced
the distracter bias to the second half of the movement (81/8,
47–100%; Fig. 2k). Post hoc balancing the vertical term trials
left the effect largely unchanged (89/11, 41.7–100%; Fig. 2l).

Post hoc balanced vertical term trials (Fig. 2g) provide the
most unbiased estimate of the reference effect. Comparing its
onset in these trials to that of the distracter effect in the anal-
ogous comparison (Fig. 2l) shows an earlier onset of the re-
ference effect by 8.6% of movement time (equaling 91.1 ms,
based on mean movement time in these trials).

Discussion

We have described a mouse tracking paradigm in which un-
known spatial targets were specified by the task through a
relational description and demonstrated how influences from
multiple effect sources in such a setup may be disentangled.

As predicted, distracters attracted mouse paths, similar to
decision alternatives in classic mouse tracking studies (e.g.,
Dale et al., 2007). The predicted attraction toward reference
items was observed as well, for the spatial terms “above” and
“below”. Moreover, as hypothesized, a bias in spatial term di-
rection was present from early on for horizontal-axis spatial
terms. We interpret this as a spatial term effect rather than re-
pulsion from the reference based on the very early onset (note
that the spatial term was not predictive of absolute target lo-
cation in the paradigm) and in line with previous evidence
(Tower-Richardi et al., 2012). Its apparent absence in vertical
term trials is unsurprising, as it would act orthogonally to the
Fig. 2: Mean trajectories and comparisons between conditions. Trajectory plots show the distance from the direct path to the target (x=0) against the proportion of elapsed movement time. Negative values indicate leftward deviation and vice versa. Conditions compared within a panel differ with respect to the position, relative to the direct path, of either the center of mass (CoM), the reference item (Ref), or the distracter (Dtr). Blue and red circles in the top of each plot indicate the direction in which the item of interest was situated for the trajectory in the same color. Gray dotted lines represent the mean across all trajectories in the two compared conditions. Color-coded p-values and effect sizes (absolute Cohen’s d) for each time step are shown on the left and right side of each plot. The black line on the left side of each plot indicates individual significant t-tests (p < 0.01) for a given sequence of individually significant time steps (based on bootstrap criterion; see Methods for details). A bar graph below each trajectory plot shows the distribution of trials over categories relevant for counter-balancing in the respective comparison; bars correspond to trajectory lines of the same hue. (a) Mean across all trials. (b) Means by CoM side, (c) by reference side, across all trials, (d) by distracter side, only including trials with spatial terms “left” and “right”, (e) by reference side, only including trials with spatial terms “above” and “below”; (f) and (g) show the same comparisons as (d) and (e), but with a post hoc balanced set of trials. (h) Means by distracter side, across all trials, (i) by distracter side, only including trials with spatial terms “left” and “right”, (j) by distracter side, only including trials with spatial terms “above” and “below”; (k) and (l) show the same comparisons as (i) and (j), but with a post hoc balanced set of trials.
axis along which deviation was assessed. This may also explain why reference attraction is visible only in vertical term trials: If the spatial term effect impacts the entire length of trajectories as hypothesized, the two effects may cancel each other out in the late portion of horizontal term trials.

The attraction to the reference item confirms that it engages spatial attention during relational processing (e.g., Franconeri et al., 2012), even when it is unique in color. This may hint that spatial indexing (Logan & Sadler, 1996) of its position is mandatory for grounding. It further suggests that computationally relevant non-targets can impact the motor level.

The observed distracter attraction is reminiscent of reaches to average locations under target uncertainty (e.g., Chapman et al., 2010). Aspects specific to relation grounding may as well play a role, though, for instance, through locations in a neural map being differentially activated by a relational template. A hint at this interpretation is the small extent of distracter attraction compared to a mean distracter distance to the direct path of 20.59 mm, calling into question mere averaging. The latter aspect, as well as the early spatial term effect, the mandatory reference selection, and the offset time courses of reference and distracter attraction, parallel our neural process model of grounding, in which item positions are sequentially stored in distinct neural substrates to apply a concurrently active, graded relational template (Richter et al., 2017).

There is ample room for new research in the direction suggested here. One step may be to clarify how far distracter attraction is specific to relational processing. The temporal order of effects must be probed more formally. Finally, higher cognitive processes may further be unraveled through additional variations of spatial phrase structure or visual displays.

Conclusion
As participants perceptually ground spatial phrases such as “the red to the left of the green”, they attend to potential target objects (here, red ones) and typically select the one best matching the spatial relation. Mouse trajectories toward the ultimately selected target reveal transient biases in multiple directions. First, they show attraction to the alternative targets, consistent with previous evidence. Second, an attraction to the reference object (“green”) begins somewhat earlier and may reflect allocation of spatial attention. Third, a bias in spatial term direction is present from early on. Overall, this study frames motor responses as direct reflections of the perceptual grounding of spatial phrases, bringing evidence for the coupling of cognitive to sensory-motor processes to a new level.

References